ornl

OAK
RIDGE
NATIONAL
LABORATORY

UNION
CARBIDE

NUREG/CR-3114
(ORNL/TM-8614)

cy. 32

Proceedings of Workshop on Cognitive Modeling of Nuclear Plant Control Room Operators, August 15–18, 1982

Oak Ridge National Laboratory
Research Library
4500N, MS-6191
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON.
If you wish someone else to see this report, send in name with report and the library will arrange a loan.
ORNL 118(10-05)

This Work Performed For
Nuclear Regulatory Commission
Under
DOE Interagency Agreement 40-550-75

OPERATED BY
UNION CARBIDE CORPORATION
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161

Available from
GPO Sales Program
Division of Technical Information and Document Control
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

OK/NB
3-29-83

OAK RIDGE NATIONAL LABORATORY

OPERATED BY
UNION CARBIDE CORPORATION
NUCLEAR DIVISION

POST OFFICE BOX X
OAK RIDGE, TENNESSEE 37830

March 10, 1983

TO: Recipients of NUREG/CR-3114 (ORNL/TM-8614), "Proceedings of Workshop on Cognitive Modeling of Nuclear Plant Control Room Operators, August 15-18, 1982," December, 1982.

We have learned that some copies of the above report are missing approximately 15 pages. If your copy is incomplete, please return it for a replacement copy.

Laboratory Records Department
Bldg. 4500
Oak Ridge National Laboratory
Post Office Box X
Oak Ridge, Tennessee 37830

NUREG/CR-3114
ORNL/TM-8614
Distribution Category RX

Contract No. W-7405-eng-26

Proceedings of

WORKSHOP ON COGNITIVE MODELING OF NUCLEAR PLANT CONTROL ROOM OPERATORS

Dedham, Massachusetts
August 15-18, 1982

Workshop Committee:

Thomas B. Sheridan, Chairman
Massachusetts Institute of Technology

James P. Jenkins
U.S. Nuclear Regulatory Commission

Roger A. Kisner
Oak Ridge National Laboratory

Proceedings Editor:
Lorraine S. Abbott

Date Published: December 1982

This Work Performed For
Nuclear Regulatory Commission
Under
DOE Interagency Agreement 40-550-75
NRC FIN No. B0438

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37830
operated by
UNION CARBIDE CORPORATION
for the
DEPARTMENT OF ENERGY

LIST OF WORKSHOP PARTICIPANTS

Sheldon Baron, *Bolt, Beranek and Newman, Inc., Cambridge, Massachusetts*
Barbara Jean Bell, *Sandia National Laboratory, Albuquerque, New Mexico*
Leo Beltracchi, *U.S. Nuclear Regulatory Commission, Washington, DC*
Julian M. Christiansen, *General Physics Corporation, Dayton, Ohio*
William R. Corcoran, *Combustion Engineering, Inc., Windsor, Connecticut*
Renwick Curry, *National Aeronautics and Space Administration (Ames), Mountain View, California*
David E. Embry,* *United Kingdom Atomic Energy Agency, Warrington, United Kingdom*
Donald Farr,† *Babcock and Wilcox, Lynchburg, Virginia*
Carl E. Feehrer, *Bolt, Beranek and Newman, Inc., Cambridge, Massachusetts*
Joseph R. Fragola, *Science Applications, Inc., New York, New York*
Audrey M. Fullerton, *Oak Ridge National Laboratory, Oak Ridge, Tennessee*
Paul M. Haas, *Oak Ridge National Laboratory, Oak Ridge, Tennessee*
John M. Gallagher, Jr., *Westinghouse Electric Corporation, Pittsburgh, Pennsylvania*
Henry M. Halff, *Office of Naval Research, Arlington, Virginia*
Lewis F. Hanes, *Westinghouse R&D Center, Pittsburg, Pennsylvania*
Theodore R. Hatcher, *EG&G Idaho, Inc., Idaho Falls, Idaho*
James P. Jenkins, *U.S. Nuclear Regulatory Commission, Washington, DC*
Michael G. Johnson, *University of Tennessee, Knoxville, Tennessee*
Roger A. Kisner, *Oak Ridge National Laboratory, Oak Ridge, Tennessee*
William J. Luckas, Jr., *Brookhaven National Laboratory, Upton, New York*
Thomas P. Moran, *Xerox Palo Alto Research Center, Palo Alto, California*
Neville Moray, *University of Toronto, Toronto, Ontario, Canada*
Kenneth G. Murphy, Jr., *U.S. Nuclear Regulatory Commission, Washington, DC*
Fred R. Mynatt, *Oak Ridge National Laboratory, Oak Ridge, Tennessee*
William R. Nelson, *EG&G Idaho, Inc., Idaho Falls, Idaho*
Roger A. Newton, *Wisconsin Electric Power Company, Milwaukee, Wisconsin*
Donald A. Norman, *University of California (San Diego), La Jolla, California*
Charles M. Overbey, *U.S. Nuclear Regulatory Commission, Washington, DC*
Richard W. Pew, *Bolt, Beranek and Newman, Inc., Cambridge, Massachusetts*
Robert Pulliam, *BioTechnology, Inc., Falls Church, Virginia*
Jens Rasmussen, *Riso National Laboratory, Roskilde, Denmark*
William B. Rouse, *Search Technology, Inc., Norcross, Georgia*
Thomas G. Ryan, *U.S. Nuclear Regulatory Commission, Washington, DC*
John W. Senders, *University of Toronto, Toronto, Ontario, Canada*
Thomas B. Sheridan, *Massachusetts Institute of Technology, Cambridge, Massachusetts*
Arthur I. Siegel, *Applied Psychological Services, Wayne, Pennsylvania*
Perry W. Thorndyke, *Perceptronics, Atherton, California*
Michael Williams, *Xerox Palo Alto Research Center, Palo Alto, California*
Joseph G. Wohl, *AlphaTech, Inc., Burlington, Massachusetts*
David D. Woods, *Westinghouse R&D Center, Pittsburg, Pennsylvania*
John Wreathall, *NUS, Gaithersburg, Maryland*

**Now with Human Reliability Associates, Lancanshire, England.*
†Now with Science Applications, Inc., McLean, Virginia.

ACKNOWLEDGEMENTS

The workshop committee and the editor are indebted to Dr. Audrey M. Fullerton of Oak Ridge National Laboratory for the use of the detailed notes she took during the workshop. These notes were especially helpful in the compilation and preparation of the working group reports.

Thanks are also due to Ms. Katie Lawhorn and Ms. Jackie Lasch, Oak Ridge National Laboratory, who performed the keyboarding and layout of this document, as well as assisted in the copy editing.

Table of Contents

ABSTRACT ix

INTRODUCTION AND CHARGE

Introduction 3
Roger A. Kisner and Thomas B. Sheridan
A Charge to the Workshop Participants 7
James P. Jenkins

INVITED PAPERS

1. The Role of Cognitive Models of Operators in the Design, Operation and Licensing of Nuclear Power Plants 13
Jens Rasmussen
2. A Framework for Modeling Supervisory Control Behavior of Operators of Nuclear Power Plants 36
S. Baron, C. Feehrer, R. Muralidharan, R. Pew, and P. Horwitz
3. Cognitive Models and Computer Aids for Nuclear Plant Control Room Operators 52
Thomas B. Sheridan
4. Visual Momentum: An Example of Cognitive Models Applied to Interface Design 63
David D. Woods
5. Using Cognitive Modeling to Improve the Man-Machine Interface 73
Roger A. Newton, Richard C. Zyduck, and Dennis R. Johnson
6. Human Cognition 92
Donald A. Norman
7. The Role of Conceptual Models in Nuclear Power Plant Operation 107
Michael D. Williams, Thomas P. Moran, and John Seely Brown
8. Outline of a Model of Human Problem Solving in Failure Situations 117
William B. Rouse
9. Modeling Human-Computer Interaction 124
Thomas P. Moran
10. Cognitive Models Embedded in System Simulation Models 125
Arthur I. Siegel and J. Jay Wolf
11. A Rule-Based Approach to Cognitive Modeling of Real-Time Decision Making 147
Perry W. Thorndyke

REPORTS OF WORKING GROUPS

1. Survey of Development of Models as Applied to Nuclear Plant Operators (I) 159
W. B. Rouse, Chairman
2. Survey of Development of Models as Applied to Nuclear Plant Operators (II) 162
J. G. Wohl, Chairman
3. Use of Cognitive Models of Nuclear Plant Operators in Design and Operation 174
J. Wreathall, Chairman
4. Use of Cognitive Models of Nuclear Plant Operators in Regulation and Licensing 176
P. M. Haas, Chairman

CLOSING REMARKS

By Thomas G. Ryan 188
By Kenneth G. Murphy 189
By Leo Beltracchi 190
By Thomas B. Sheridan 191
By Roger A. Kisner 191

REFLECTIONS

On Problems Raised in Letter of Invitation 194
Erik Holnagel
On Working Group No. 1 196
N. Moray
On Workshop Theme Questions 197
R. Pulliam
On Models, Aviation, and Automation 199
R. Curry
On Post-Workshop Thoughts 200
J. M. Christensen

BIBLIOGRAPHY 205

ABSTRACT

This document presents 11 invited papers and the deliberations of four working groups at a Workshop on Cognitive Modeling of Nuclear Plant Control Room Operators that was held in Dedham, Massachusetts, under the sponsorship of the U.S. Nuclear Regulatory Commission. The purpose of the workshop was to review the status of "cognitive modeling" and to recommend to the NRC whether it should support research directed toward the development of a cognitive model of a reactor operator that could be useful by itself or as a part of a larger model of the human-machine system. It was the consensus of the invited papers and the working groups that some cognitive models developed for other types of systems can be adapted to the reactor operator under limited and precisely defined conditions (and, indeed, some already are being used). However, the development of a comprehensive model for the reactor operator should be preceded by an improved understanding of the task. In the meantime, the need for further applied research in operator cognition is apparent, as is the need for supporting data collection. Work in these areas could begin immediately.

INTRODUCTION AND CHARGE

INTRODUCTION

Roger A. Kisner
Oak Ridge National Laboratory

Thomas B. Sheridan
Massachusetts Institute of Technology

In an effort to facilitate the operation of nuclear power plants under both normal and non-normal conditions, the Nuclear Regulatory Commission has sponsored a number of research projects that are aimed at simplifying the operator's task and improving his performance. These projects include the addition of control room aids; more and better training, especially for handling emergencies and to provide more detailed understanding of the plant; the development of better control room procedures; and increased automation. As the research program has progressed, it has become increasingly apparent that the focus should not be on the performance of the operator in isolation but rather on the functioning of the operator within the human-machine system as an entity. It is important not only to know the capabilities and limitations of the two major components — the human and the machine — but also to understand the interactions between them and the consequences of those interactions. With the human and the machine both acting properly, the consequence is a smoothly operating plant. But with improper performance by one component, the consequence could be a serious degradation of the system that could be made worse or mitigated, depending on the response of the other component.

In contemplating this human-machine system we are awed by its complexity. The plant alone consists of many subsystems which in turn are made up of numerous large and small components. To predict the outcome of the improper performance of just one component of the plant, we must postulate all possible sequences of events that might be triggered by its failure. The increasing degree of computerization in sensing and control has not made the plant any simpler.

Even so, the greater complexity is with the human element of the system. While it is relatively easy to measure and predict the physical movements and time spans required for a particular operator to perform a specific manual task, it is far more difficult to measure and predict his cognitive behavior — that is, the sequence of mental events prompted by given stimuli and culminating in the decision to perform (or not perform) a specific manual task. Thus, we cannot anticipate with confidence how an operator will react to a given set of circumstances, especially if those circumstances are outside his experience, which is most likely to be the case in emergencies. In such situations, normal automation is inadequate and we are more dependent on operator cognition.

The problem of viewing the human-machine system as an entity is not unique to nuclear power plants. It has been studied for many years in relation to the operation of aircraft, spacecraft, and ships, as well as the operation of various process plants. However, the adaptability of this earlier research to the operation of nuclear power plants is not obvious. In particular, it is not clear that the cognitive models that have been developed to study human-machine interactions in other (usually simpler) systems are either appropriate or adequate for the study of human-machine interactions at a nuclear power plant. Moreover, it is not clear whether the cognitive sciences yet have the knowledge required to develop the needed models. While some speculative models are available, it is generally agreed that the data needed to test them are not sufficient.

In an attempt to better understand the status of the cognitive sciences and their relationship to the study of the performance of a reactor operator, the NRC sponsored this Workshop on Cognitive Modeling of Nuclear Plant Control Room Operators on August 15-18, 1982, at the MIT Endicott House, Dedham, Massachusetts. Organized by Oak Ridge National Laboratory and attended by 42 invitees, the workshop included a presentation of 11 invited papers and the deliberations (and reports) of four working groups. This document collects the proceedings of the workshop, as well as some pertinent additional material prepared by some of the participants.

Prior to attending the workshop, the participants were informed that the workshop was to be a follow-on of a study conducted for NRC by a team from Bolt, Beranek, and Newman, Inc., on the feasibility of applying a "supervisory control modeling technology to the study of critical operator-machine problems in the operation of a nuclear power plant." The result of the BBN study was a "conceptual model that incorporates the major elements of the operator and of the plant to be controlled." The essence of this model is included in these proceedings as Paper No. 2.

As stated in the letter of invitation, the objectives of the workshop were: (1) to identify the applicability, validity, and usefulness of cognitive modeling to the licensing and regulation of nuclear plants, (2) to survey available cognitive modeling techniques, including the capabilities, limitations, data requirements, structures, and theoretical constructs on which each model is based, and (3) to recommend research directions for the NRC.

Some of the specific questions that the four working groups were to address were:

1. Does the NRC need information about cognitive models for regulation of nuclear facilities and what are the major purposes of these models?
2. What reactor operator cognitive behaviors do we need to predict? Can these behaviors be listed in order of importance?
3. How would such predictions or other outputs from cognitive models be applied by the NRC in the areas of plant design, procedures, management, training, licensing, etc.?
4. What are the principal cognitive modeling approaches in use now, and how are they used in other applications or fields?
5. At what stage of development do models become useful?
6. What taxonomies apply? (Consider, for example, normative vs. descriptive, control vs. decision, mathematical vs. verbal, deterministic vs. stochastic.)
7. What data are presently available to use for generating models? What data are missing? Do we have means of collecting the needed data?
8. What sources of data can be used to calibrate or validate cognitive models of nuclear power plant operators? [For example, do we need anecdotal data, LERs (licensee event reports), simulator data, etc.?]
9. What are the current obstacles to (a) developing the status of cognitive modeling of nuclear power plant operators, (b) calibrating and validating the models, (c) putting the models to use, and (d) learning from experience?

During the two full days of the workshop the four working groups met between sessions in which several invited papers were presented. Thus the groups were able to progressively consider the contents of the papers and select the information they considered to be useful for their own reports, which were given on the final half day of the workshop on the following topics:

Groups 1,2:	Survey of Development of Models as Applied to Nuclear Plant Operators
Group 3:	Use of Cognitive Models of Nuclear Plant Operators in Design and Operation
Group 4:	Use of Cognitive Models of Nuclear Plant Operators in Regulation and Licensing

In the proceedings that follow we present first the 11 invited papers in the order that they were given. We then present the working group reports which were given orally at the workshop but largely documented later. These reports are followed by the closing remarks made by three members of the NRC and two members of the workshop committee. A final section includes "reflections" on the workshop of four participants and one nonparticipant.

We believe that these proceedings will be valuable both to the NRC as it attempts to "get a handle" on what cognitive research is needed and to the participants as they attempt to understand the scope of the problem and to direct their cognitive abilities to its solution.

A CHARGE TO THE WORKSHOP PARTICIPANTS

James P. Jenkins
U.S. Nuclear Regulatory Commission

What are the compelling reasons for this workshop?

What are the origins in NRC concern for cognitive explanations of operator behavior, models about cognition, and predictions of cognitive performance?

What are the expectations for this workshop?

These three questions are at the root of your sponsor's interests for this workshop. We have a viewpoint of our own within the NRC and it has been stated many times: We need to understand and ultimately to predict and account for the variance in operator behavior attributable to cognitive error to no less extent than we need to account for total human performance error. The watershed event in the NRC's involvement in human performance prediction and modeling was the conclusion about remote and proximate contributing human factors to the accident at Three Mile Island Unit 2 in March 1979, and, in retrospect, to all other incidents which have had human error as a contributing or common cause.

If you reflect on the specific actions taken since TMI-2, you will perceive a continuing concern for the subject matter of this workshop. Let me mention some major decisions and programs to reinforce this point. In December, 1979, an NRC-sponsored Workshop on Human Reliability was held at Myrtle Beach (MB-I) to coalesce the regulatory and the research directions for human error and human factors. Two publications were specific products of that workshop — a handbook authored by Alan Swain and Hank Guttmann that describes the model called THERP (Technique for Human Error Production) and its application (NUREG/CR-1278), and a workbook for identification of man-machine interfaces that are or can be critical failure events in a fault tree or system event (NUREG/CR-2254).

In FY80 an upgraded human reliability research program was established by NRC. Sandia National Laboratories was tasked to develop the two important documents I have just mentioned and also to begin a series of research tasks that have resulted in the identification of various models to be used for maintainer reliability prediction. In other research projects the types of human failures in the operation and maintenance of pumps, valves, electrical and electromechanical instruments have been identified, and human failure rates have been analyzed. A human performance data bank also is being evaluated under this program, and an independent evaluation of human error data has been started.

In August 1981, a second workshop, titled the Workshop on Human Factors, was held at Myrtle Beach. A major purpose of that meeting was stated in the Foreword of the proceedings as follows: "In the area of human performance, which was more in line with the scope of the original (i.e., MB-I) meeting, a need was shown for the development of a taxonomy or model to structure future data gathering and the need of models or data to address the issue of cognitive behavior."

One of the first efforts following MB-II was a research task assigned to ORNL to implement a system-based plant-specific man-machine model. Under subcontract to ORNL, Bolt, Beranak and Newman, Inc. (BBN) developed a concept which you will hear about during this workshop. And in April, 1982, a decision was reached that a symposium — I use the word in its classical sense in Greek to mean a meeting of the minds for an exchange of ideas — on cognitive models was mandatory before proceeding further. Compelling reasons for this workshop, for which we are all gathered here, were the following:

- To bring the best available thinking to bear on the problem of cognitive performance and human reliability assessment.
- To ascertain in which direction a composite research program should lead. That is, should it be data collection? analysis of available data from a new nomothetical network? alternative model development? cognitive performance prediction and model validation?
- To establish a firm basis, by way of the conclusions reached in this workshop, for the specific direction we finally choose.

When one reflects on the issues, we find not only models of cognition which span different explanations of cognitive behavior, but also many underlying concepts of cognition. The model developers have provided a range of models. To name a few, we have Rasmussen's Rule- Skill- and Knowledge-Based Performance Model; Sheridan's Supervisory Control Model; Siegel's Man-Machine Simulation Model a la Newell and Simons; OATS-Weatherall's Operator Action Tree Model; models which are properly called mathematical models, such as Shannon and Weaver's model of communications; engineering models like the BBN model; computer models as represented by the field of artificial intelligence; conceptual models like the model of information processing contained by the S-O-R* paradigm; and psychological models.

If this workshop had its genesis in academia we would be expected to have vigorous discussions about the attributes and deficiencies of the models, or the clarity of the intermediate-level construct and what is the underlying theory of cognition, or perhaps the discussion would focus on the problem of criteria and construct validation. If this workshop were focused on engineering solutions we could expect discussions on how much, how extensive or what quantifications are required for model application. However, this workshop is directed neither toward academic-only nor engineering-only. Rather, the objectives of the workshop are to provide a framework and context for our discussions. Let me repeat them:

- To identify the applicability, validity, and usefulness of cognitive modeling to the licensing and regulation of nuclear plants.
- To survey available cognitive modeling techniques, including the capabilities, limitations, data requirements, structures, and theoretical constructs on which each model is based.
- To recommend research directions for the Nuclear Regulatory Commission.

---

*S-O-R = Stimulus-Organism-Response.

In closing, I would like to offer some guidelines or standards whereby we can evaluate the models and cognitive behavioral concepts which will be discussed at the workshop. If I may call them criteria, then they are intended to differentiate among models and concepts:

(1) The degree of acceptance of the model by the using community is a major measure of the adequacy of the model. Its utility should be recognizable.

(2) The model must possess the characteristics of verisimilitude. It must show the quality of human experience.

(3) The validity of the input data is critical. Requirements for data and availability of data to be manipulated by the model are needed.

(4) If the model requires simulation, then the model must be appropriately documented with structural programming in a common language. The translation of data from a behaviorial statement to elements of a computer program must be carefully documented, along with program specifications or flow charts.

(5) The ultimate credibility of the model lies in the validity of the underlying constructs. The multi-trait, multi-method analysis for convergent and divergent validity is recommended.

(6) The output of the model should provide results at the appropriate level of detail consistent with the objectives of the user. If decisions are to be made as a consequence of modeling, then information output should be in a form which allows that decision to be made with minimum translation of the model output.

(7) The interface between the model user and the model should be easy to use. This means documentation free from error of omission and fact and easy to understand.

This is your charge. I expect you to meet it by your papers and discussions and through the conclusions and recommendations you will make.

INVITED PAPERS

Paper 1

THE ROLE OF COGNITIVE MODELS OF OPERATORS IN THE DESIGN, OPERATION AND LICENSING OF NUCLEAR POWER PLANTS

Jens Rasmussen
Riso National Laboratory, Denmark

1. INTRODUCTION

Cognitive models of the behavior of nuclear power plant operators — that is, models developed in terms of human properties rather than external task characteristics — are assuming increasingly important roles in plant design, operation and licensing. This is partly due to an increased concern for human decision making during unfamiliar plant conditions, and partly due to problems that arise when modern information technology is used to support operators in complex situations. In support of the workshop discussions, some of the problems we have identified during our work on interface design and risk analysis are described.

First, the question of *categories of models* is raised. (For a definition of cognitive models see Fig. 1.) Different types of models match the various problems met during systems design and operation, and rather than seek a general, quantitative model of human behavior, the use of different types of models is advocated, together with a conceptual framework serving to interrelate the use of the various specific models.

Next, the use of cognitive models for *system design* is discussed. It is argued that not only do we need a development of such models, but also we need a more consistent identification of the *control requirements* of the plant and a related design of the overall plant control strategy.

The use of the available cognitive models for more effective *operator training* is also advocated. Until recently the planning of operator training has been rather unrelated to the cognitive models developed, and mainly has depended upon the use of high-fidelity, large-scale simulators.

The need for using cognitive models in *risk analysis* is also emphasized. The importance of quantitative probabilistic models for risk analysis is well known, but some problems related to emergency performance are mentioned, and the problem of identifying the potential for systematic coupling of otherwise independent events (*system interaction*) due to human decisions and actions is discussed. The value of risk analysis depends upon the degree to which the models and assumptions behind the analysis are controlled and maintained during plant operation, that is, upon *risk management* by feedback control based on analysis of event reports. To serve as reference for this feedback control, the models used in the risk analysis — including the cognitive models of operators — must be explicitly stated.

Finally, the *sources of human performance data*, that is, event reports, incident analysis, experiments, and training simulators are mentioned, and the need for a consistent framework for data analysis based on cognitive models is discussed.

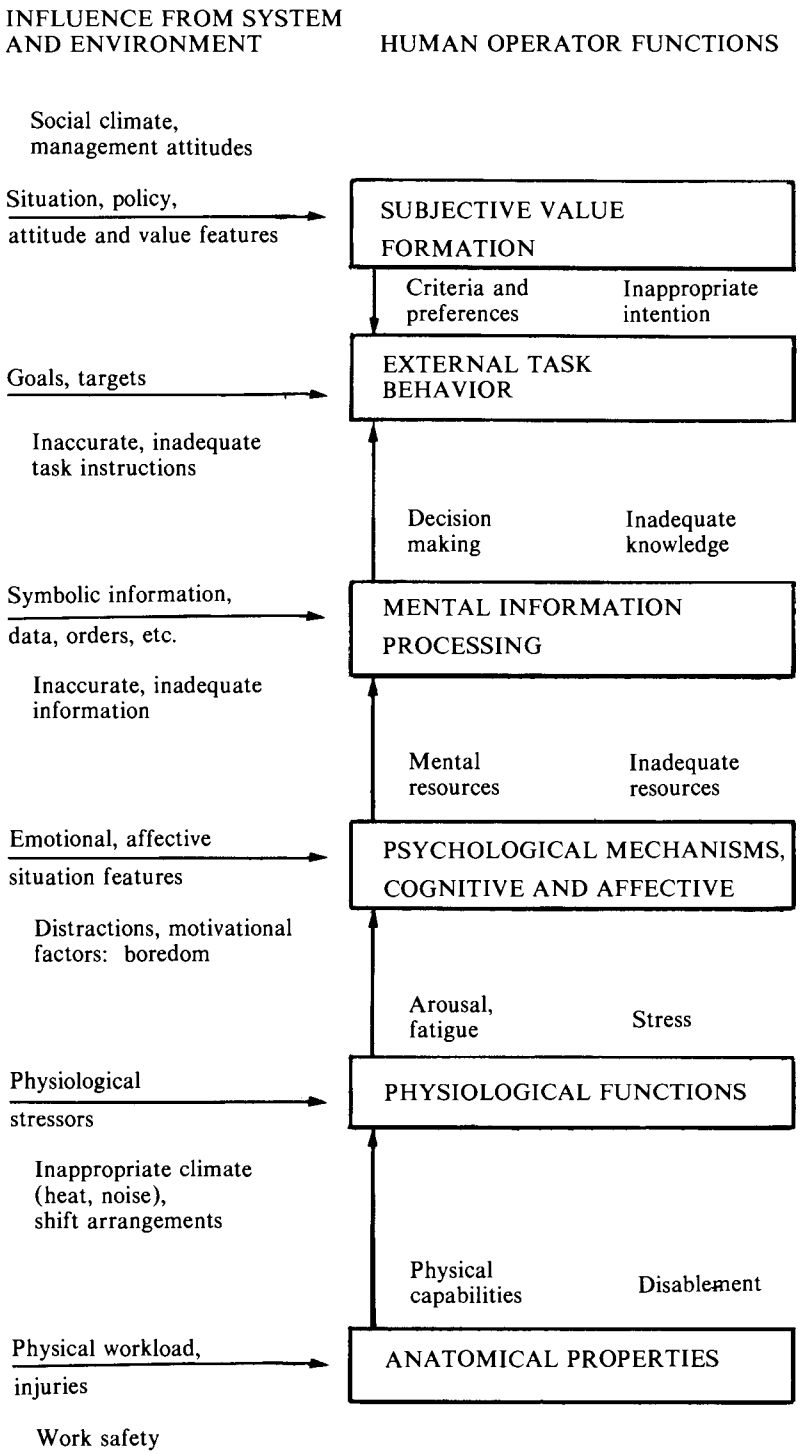


Fig. 1. Definition of Cognitive Model. Human behavior and resource/demand conflict can be modelled at several levels. Here cognitive models are taken to be models at the levels of information processing and psychological mechanisms, in contrast to behavioristic models in external task terms.

2. CATEGORIES OF MODELS OF HUMAN BEHAVIOR[1]

A discussion of models of human behavior raises immediately the distinction between qualitative and quantitative models. Frequently, qualitative models are considered merely to be premature, descriptive models which after further work will develop into or be replaced by proper quantitative models. However, this is not necessarily the case. In several respects the two kinds of models have different and equally important roles for analysis and prediction of performance. This difference in significance is related to the distinction between categories of behavior and the members of such categories, that is, the specific behavior in particular situations. Bateson[2] discusses this distinction in detail with reference to Whitehead and Russell's logical types: "...there is a deep gulf between statements about an identified individual and statements about a class. Such statements are of *different logical types*, and prediction from one to the other is always unsure" (p. 46). The fact that "the generic we can know, but the specific eludes us" (p. 45) has different implications, depending upon the purpose of the modelling effort.

For systems design, qualitative models will serve important purposes if they are able to predict the category of behavior that will be activated by different possible interface configurations and display formats. The model will then support the choice of an interface design which will activate a category of behavior having limiting properties compatible with the functions allocated to the human operator. In a way, research on human performance to support system design should not focus on modelling actual performance in existing environments, but rather on *possible* performance in optimal, future systems, as has been discussed by Sloman[3] in a philosophical context. Qualitative models identifying categories of behavior and the limiting properties of the related human resources will serve designers a long way in the design of systems which allow humans to optimize their behavior within a proper category.[4]

Compare this with Norman's arguments[5] for the importance of considering the proper mental image for design of "friendly" systems and the need for a profession he calls "cognitive engineering." The distinctions between models of categories and of particulars have different implications, depending also on the cognitive level of behavior considered. At the skill-based level we are considering highly trained people, similar to experimental psychologists' "well trained subjects" who have adapted to the particular environment. In this domain, models of optimal human performance are mainly models of the behavior of the *environment*, as seen through the man. Therefore, generic quantitative models of human performance in well-structured tasks can — and have — been developed at this level of performance. At the level of knowledge-based behavior, however, we are dealing with individual reactions to unfamiliar situations, and models will be more a question of qualitative matching of categories of system requirements with human resources. For unfamiliar tasks, these resources depend on a specific person's subjective preferences, experience, and state of training. There is a long tradition within vehicle control to use quantitative models for systems design and performance analysis, such as the models based on optimal control theory. During recent years attempts have been made to extend these models to higher level human decision making to conform with the increasing levels of automation in aviation, and to transfer such models for process control applications. Whether this approach is fruitful depends on the nature of the human task. The optimal control part of the model may not be needed if the manual acts are no longer an integral part of the control task, but merely a general interface manipulation skill. In that case, an independent

development of a decision model may lead to a more direct approach. What we need is not a global quantitative model of human performance, but a set of models which is reliable for defined categories of work conditions, together with a qualitative framework describing and defining their coverage and relationships.

To be useful, both the qualitative and the quantitative models must reflect the structure underlying the mental processes: the internal or mental *models*; the kind of *data* dealt with by the processes; and the *rules* or *strategies* used to control the processes. In addition, the models must reflect the limits of human capabilities so that human "errors" are also modelled properly.

This question of also modelling errors properly leads directly to the issue of using analogue parallel processing models versus the sequential digital models of human information processes offered by the artificial intelligence (AI) community. Can holistic human perception, for instance, be properly modelled by the sequential "production rule" systems? In the present context of models for system design and evaluation, the fundamental question appears not to be whether a model is implemented for experimental evaluation by means of one or another physical information processing system but whether or not there exists an intermediate level of concept, actually a *theory*, linking the actual human performance and the particular computer implementation of a model under study, that is, a framework having a one-to-one correspondence to human mechanisms, their processing limitations and related failure mechanisms.

A map of different cognitive levels of human behavior illustrating the relationship among various more specific models is shown in Fig. 2.

3. THE ROLE OF COGNITIVE MODELS IN SYSTEMS DESIGN[6]

In our work, our concern is with the timely development of models of human performance which can be useful for the design and evaluation of new interface systems. For this purpose, we do not need a single, integrated, quantitative model of human performance but rather an overall qualitative model which allows us to match categories of performance to types of situations. In addition, we need a number of more detailed and preferably quantitative models which represent selected human functions and limiting properties within the categories. The role of the qualitative model will generally be to guide overall design of the structure of the system (including, for example, display formats), while selective quantitative models can be used to optimize the detailed designs.

In many cases, the use of quantitative models for optimizing a design can be replaced by experimental evaluation. Unfortunately, however, it is the categories of performance for which experimental evaluation is most feasible — that is, skilled- and rule-based performances — which are most readily modelled quantitatively. A major difficulty is the modelling of the knowledge-based control of performance during unfamiliar situations, as well as the interaction among the different levels of performance, depending upon the state of training. In particular, it is very difficult to study experimentally, or analytically, the interference from overlearned routines during situations calling for knowledge-based responses. Several problem areas for research can be identified.

The first problem we meet in design of interface systems based on modern information technology is the tradition from the one-sensor—one-indicator technology that the operator task is expressed in terms of actions on the system, the state of which the operator is supposed to "figure out for himself" from readings of a number of physical variables and his

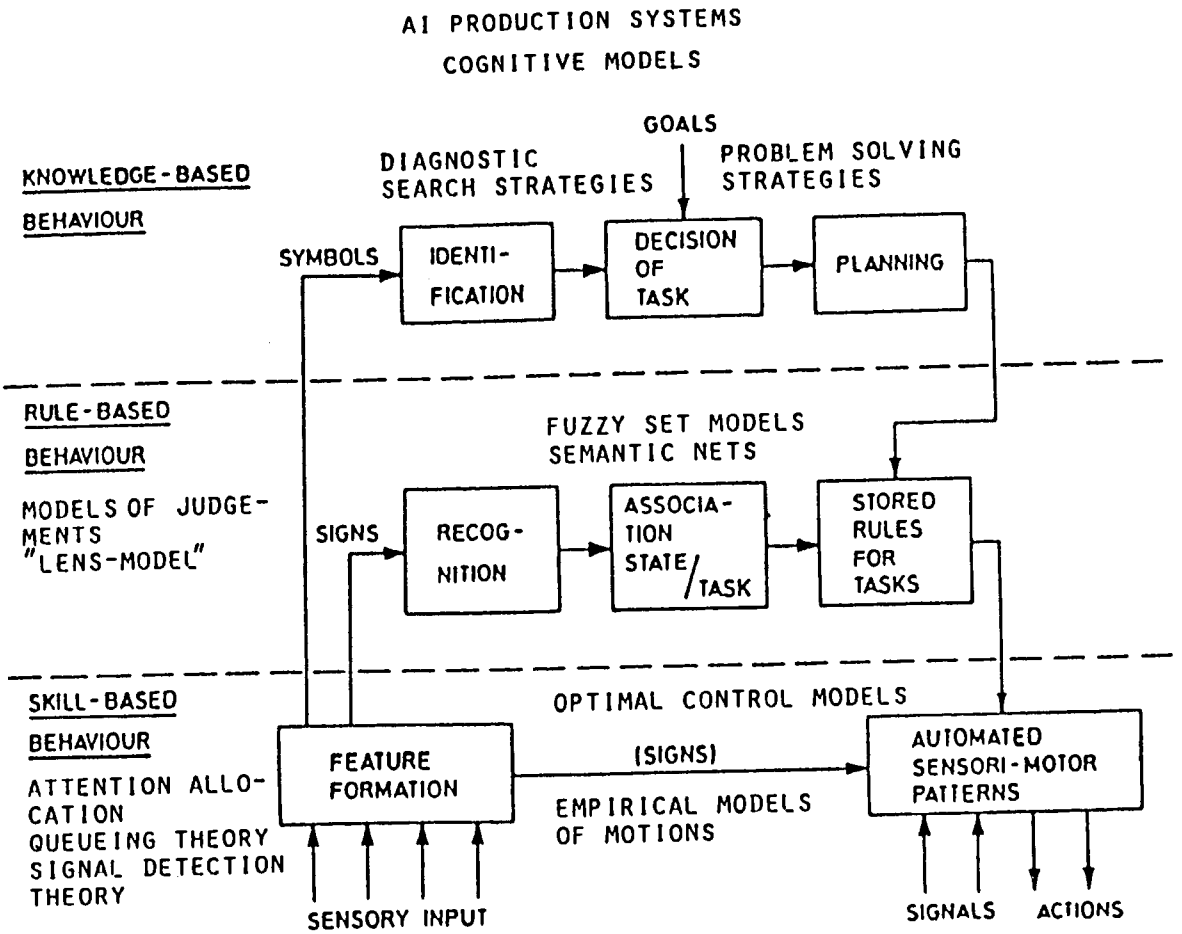


Fig. 2. Map of Human Information Processing at Three Levels of Behavior. Examples of models are indicated to illustrate their interrelation.

training in system fundamentals. However, if computer technology is to be used to optimize man-machine communication, information presentation must be structured according to the nature of the control tasks the operator is supposed to perform. To do this properly, it is necessary to design the hierarchy of functions called for in the control of complex systems as one consistent whole — regardless of whether the individual functions are automated or allocated to operators. In a supervisory control task, the operator will have to face tasks at several levels in the hierarchy of control functions; that is, the concepts used in a proper description of the various tasks will vary in level of abstraction between physical implementation and overall system purpose[4,7,8] (see Fig. 3).

In order to plan the formats of data presentation and the integration of measured data needed to derive the related variables, it is necessary to have a *formalized description of the categories of control tasks* at the various levels of abstraction. An attempt to develop such a description is given by Lind.[9] In addition, information on the subjective human preferences or performance criteria which will control the selection of strategy in a given environment is necessary for design.

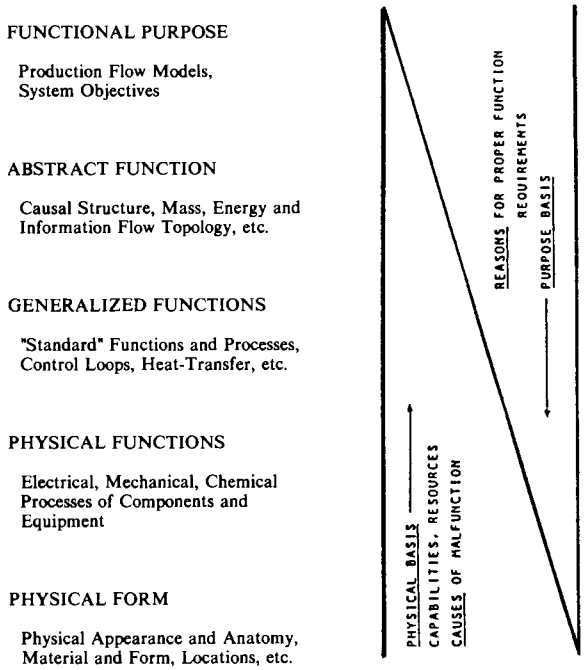


Fig. 3. Levels of an Abstraction Hierarchy. The properties of a technical system can be described at several levels of an abstraction hierarchy representing the physical structure and design intentions in different proportions.

During design of the process plant itself, the functions of the system and its physical implementation are developed by iteratively considering the plant at various levels of abstraction and in increasing degree of detail (see Fig. 4).

During this design process, the physical system is identified, that is, the implementation of those causal structures depending on mass and energy relations. However, as the degree of physical detail increases during the design process, so does the number of degrees of freedom in functional states. Therefore causal links by means of control paths relating desired states with necessary control actions must be introduced to constrain the possible operational states.

In this way, the desired states of functions and equipment will be identified during design at different levels of abstraction, and the necessary information or control constraints will be identified in terms of the conceptual framework related to these levels. In general, a skilled designer will immediately be able to identify suitable and familiar control system concepts. A consistent systems design including operator control functions can, however, be performed more systematically by means of a generalized decision model and a consistent formulation of the control requirements.[6]

The system's control requirements are derived from the necessary relations between the actual states, the desired states or changes of states, and the required actions on the system. This means that planning of control actions involves the rational decision sequence of

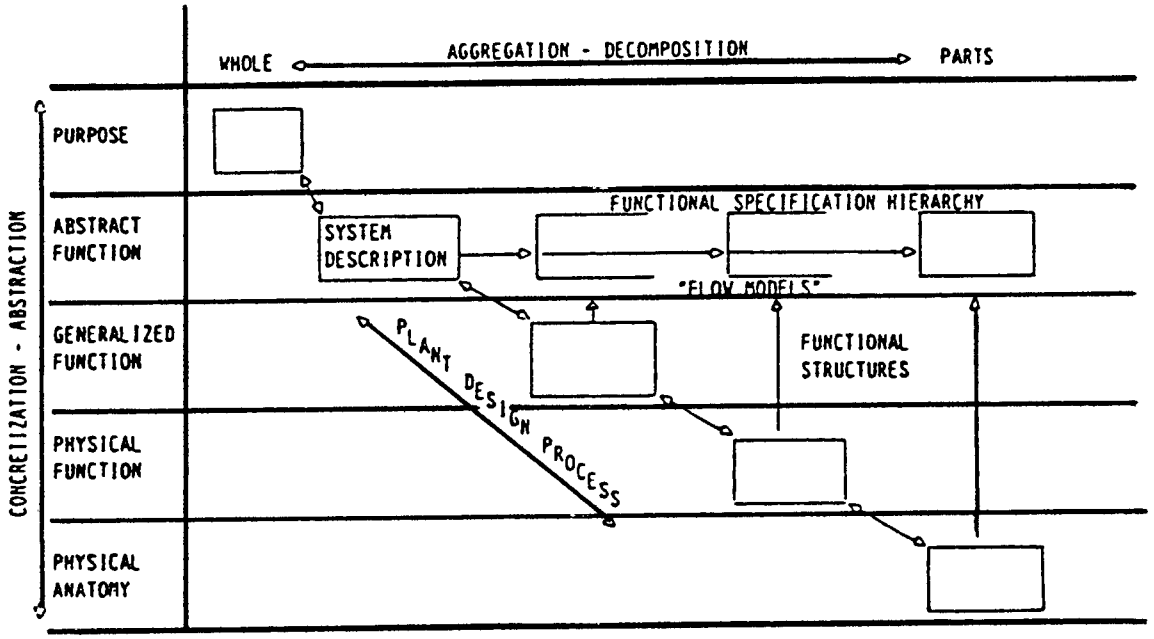


Fig. 4. Plant Design by Considering It at Various Levels of Abstraction. The design of a physical process plant can be viewed as an iterative transformation up and down an abstraction hierarchy coupled to decomposition into parts and components.

Fig. 5, covering state identification, goal evaluation, and prioritizing, in addition to the planning itself. Depending upon the control task allocation, the decision sequence — or parts of it — will be performed by the designer himself, the plant operator or the process computer. The conceptual framework within which decisions are taken will usually depend upon the background of the person, that is, whether he is the designer or operator, and upon the immediate context of the decision. However, to have a consistent overall design and to be able to formalize the decision functions to be performed by the computer, ad-hoc decisions throughout the design process should be replaced, or at least reviewed, by considerations based on a uniform description of the necessary constraints and the related control requirements which are expressed in a suitable language. For this purpose, we consider a transformation of the desired functional states and the necessary conditions, supplies, and constraints emerging during the various phases of design specification into a uniform description of specified functional states at the level of energy and mass flow structure — the abstract functional level of Fig. 3. The result is a consistent hierarchical description of target states and intended functions, that is, a goal or specification hierarchy[10] as shown in Fig. 6.

To design the control system, including the role of the operators, the functional specification hierarchy is taken as the top level in a design hierarchy in which the decision functions needed are identified in device-independent terms before tasks are allocated as operator or automatic functions, depending upon the respective resource/demand characteristics. This systematic design depends on proper cognitive models as well as on consistent identification of the control strategies (see Fig. 7).

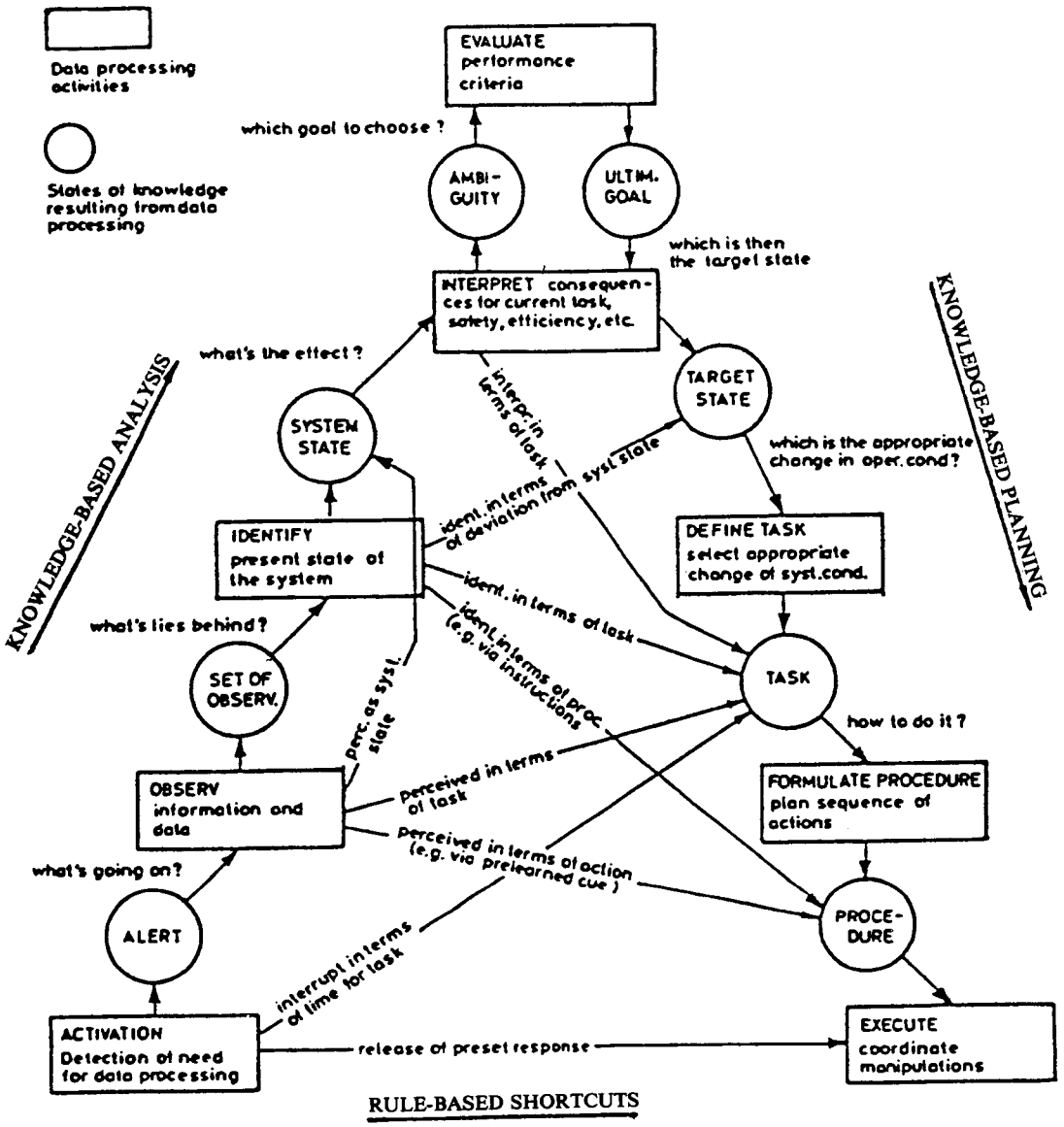


Fig. 5. Schematic Map of the Sequence of Mental Activities Used Between Initiation of Response and the Manual Action. Rational, causal reasoning connects the "states of knowledge" in the basic sequence. Stereotyped mental processes can bypass intermediate states.

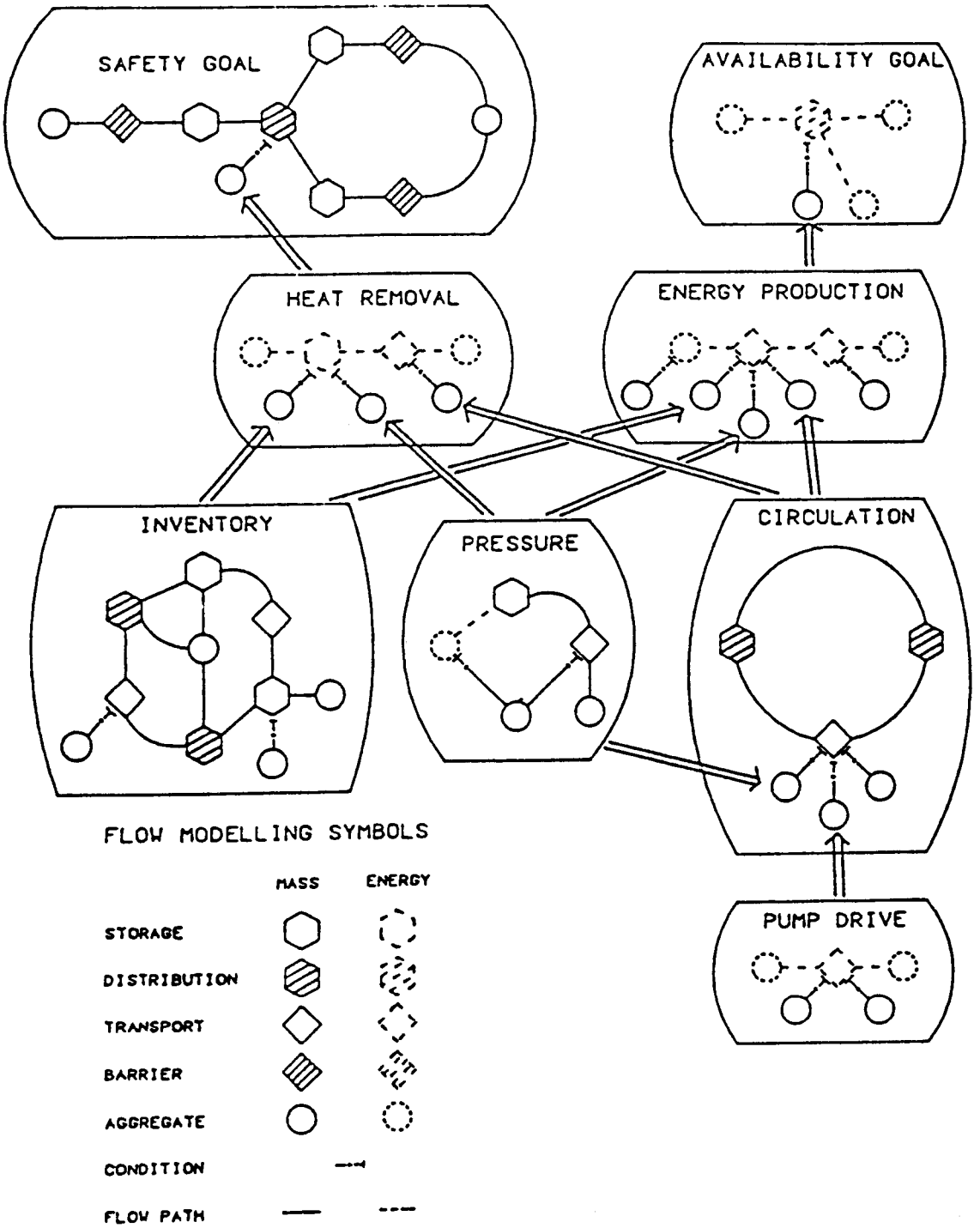


Fig. 6. Goal or Specification Hierarchy. In a complex process plant there is a many-to-many relation among purposes, functions, and equipment which can cause "systems interaction" problems. (From Lind, ref. 9.)

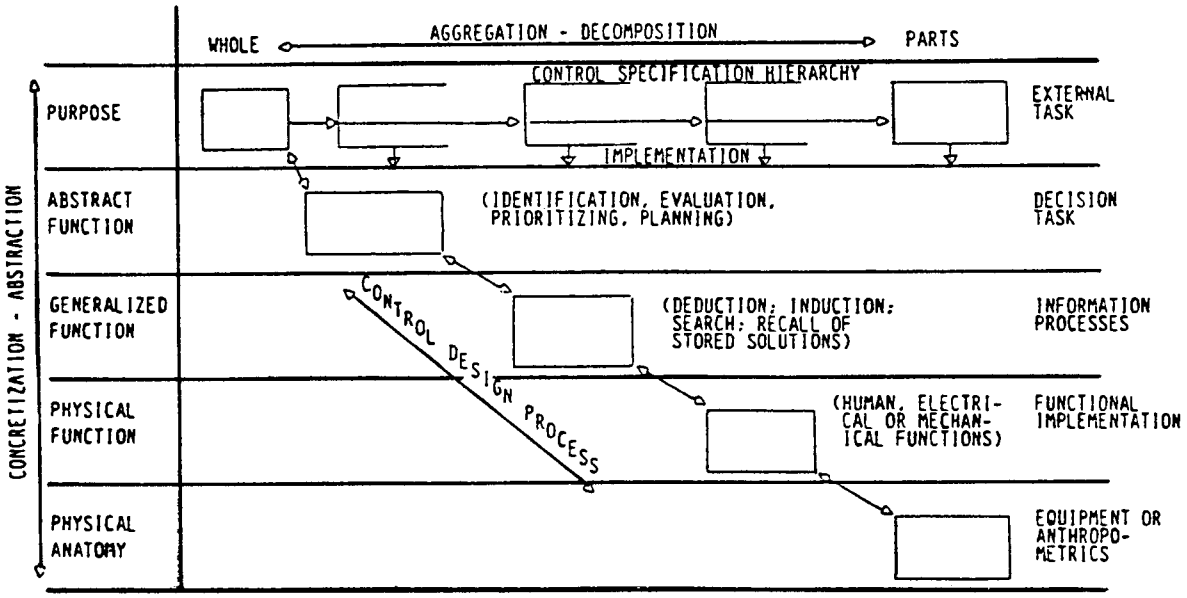


Fig. 7. Control System Design Process. A consistent design of an overall plant control strategy implies an orderly identification of decision tasks and strategies before tasks are allocated operators and automatic equipment by demand/resource considerations.

4. THE ROLE OF COGNITIVE MODELS IN OPERATOR TRAINING

In the present context, operator training means development of the necessary resources for coping with situations at all the three levels of behavior shown in Fig. 2, that is, supplying the operator with a proper repertoire of possible behaviors for both routine and unexpected situations. Until recently, training of industrial operators has not been based on models of human behavior compatible with those discussed for system design. However, explicit use of models for matching categories of system requirements and human resources to plan training programs has been made by Rouse[11] and his group, and this approach has proved promising.

Distinctions between categories of behavior similar to those of Fig. 1 have been proposed previously. Fitts[12] distinguishes between three phases of learning a skill: (1) the early or cognitive phase, (2) the intermediate or associative phase, and (3) the final or autonomous phase. If we consider that in real-life tasks there may be various degrees of training depending on variations and disturbances, the correspondence with the three levels in the present context is clear.

The strategies for decisions and the representations of the environment (that is, the plant), the rules, and the behavioral patterns forming the operator's resources at the three levels can be expected to be developed effectively by different methods of training, and the extensive use of large-scale training simulators as the main tool of training within the nuclear industry should therefore be considered.

In the skill-based domain, learning is based on behavioral patterns of movements stored during performance of the task under the control of either higher level cognitive functions or of an external teacher with his "hand on the bicycle saddle." The process computer has led to the development of "high-fidelity" dynamic training simulators which generally are considered important for training of aircraft and space-vehicle control during emergency situations, since the task is a time-space multivariable control task. However, it does not appear to be evident that high-fidelity dynamic simulators are necessary for training of skill-based behavior in most process industries. Transfer of "process-feel" seems to be rather effective within time constants and gain characteristics of the same ranges and simple dynamic simulators could serve that purpose.

Training within the rule-based domain typically depends on rehearsal of the cues and rules of the game. It is important to know the locations of sources of cues and of knobs and switches, but static mockups or magnetic board displays have proved efficient[13] for training by exercises or "talk-throughs" (Swain, private communication). Unfortunately, people are moldable and learning is not restricted to the formal training periods. The plant itself continues their training very effectively and, ultimately, the rule-based behavior will depend upon interface characteristics rather than formal training. It is therefore important to support rarely used rules and to arrange the interface in a way so that operational optimization of frequently used rules does not violate risky, but latent conditions.

Operators have no chance for developing empirical rules for infrequent occurrences, and work instructions cannot be preplanned for all possible situations. Furthermore, one of the recent "lessons learned" from TMI has been that it cannot be assumed that an instructional system for major accidents will take care of the less critical situations and thus help operators to cope with these. Consequently, in modern large-scale industrial installations, reliable operation to an increasing degree depends upon the operator being able to perform in the knowledge-based domain, that is, to form the necessary rules ad-hoc in the actual situation. Very little is apparently known about the content and means of training operators for knowledge-based performance; the dynamic and visual verisimilitude of expensive simulators is not related to the necessary knowledge of operators for performance in this domain, and training in the theoretical, physical description of plant function alone is not adequate. Special training schemes aiming at general problem solving and diagnostic skills have been tested experimentally by Duncan[14] and by Rouse[15] and should be considered seriously as an addition to full-scale simulator training.

5. THE ROLE OF COGNITIVE MODELS IN RISK ANALYSIS AND MANAGEMENT[16]

The feasibility of a probabilistic risk analysis (PRA) by decomposition techniques depends on particular design features in nuclear power plants which lead to a rather highly structured course of accidental events. Figure 8 illustrates such a generic accident and is useful for discussion of the human interaction.

From Fig. 8 it is seen that the accident propagation is subdivided into several subsequent phases by several independent counter-measures, based on different physical principles. This "defense-in-depth" design philosophy makes it possible in a realistic way to achieve very low probability of an accident with moderate requirements to the failure probabilities of each phase under the condition of independence of the different countermeasures. Therefore, the failure probability of the individual phases can be verified empirically, even though this is not the case for the overall risk probability directly.

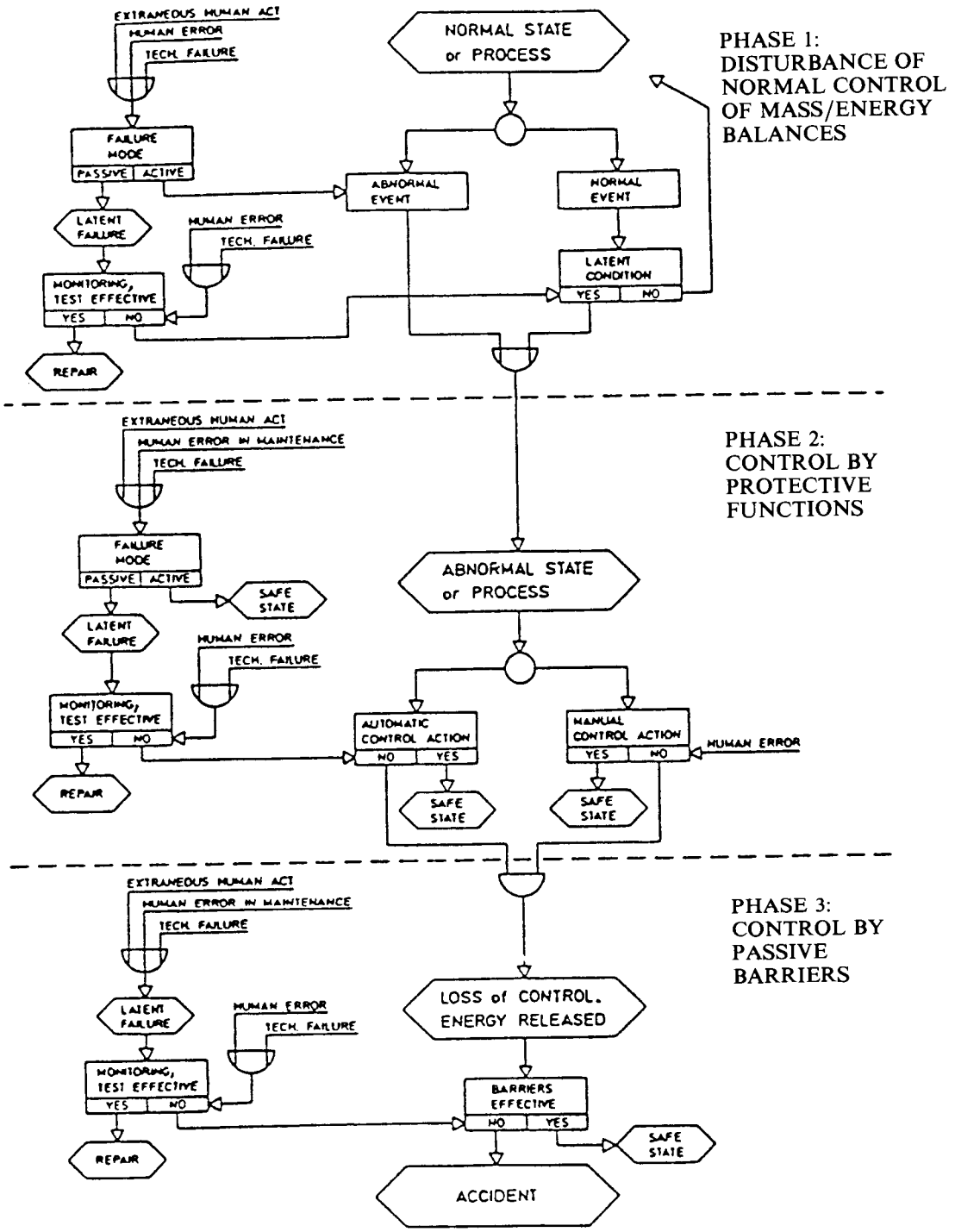


Fig. 8. Generic Accident Sequence. In nuclear power plants like other industrial processes, major accidents have a common structure of propagation due to the design philosophy applied.

With reference to the generalized sequence, typical categories of human contributions can be identified. As causes of accidental chains of events, the *simple* human errors are generally not very significant for the result of a PRA because (1) their effects are typically equivalent to technical component failures in the system the human operates, (2) their frequencies will not significantly change the overall result based on component data (considering the uncertainty of such data) or their effects are included in the data on component fault rates, and (3) human errors performed on active systems are often immediately recognized and corrected. Special consideration must instead be given the categories of human errors which influence the generic structure of the accident, such as:

- Human acts which cause couplings between different phases of Fig. 8, for example, human acts which at the same time initiate a transient and disturb the protective functions.
- Errors in the design of protective systems which affect the *capability* of the protective system to handle a subset of transients, including errors in work planning and scheduling related to maintenance and refueling periods.

In all systems based on feedback design principles, the performance is very sensitive to disturbance of the feedback path. For PRA this means that human interaction with the safety functions is a key problem. Several analytical problems can be identified:

- Estimation of the reliability of protective functions allocated to human operators. Since such functions are required under possibly stressing conditions, a meaningful quantitative reliability estimation can be made only under special assumptions regarding interface design and training.
- Estimation of the probability that, due to misunderstanding or conflicting requirements, operators will interfere with the operation of automatic safety functions during emergencies. Such interference can be caused systematically by high similarity among elements of different tasks or procedures or because the same equipment may be used for different purposes and, therefore, appear in different situations.
- Influence of human reliability on the maintenance, test and calibration of protective systems. Problems are in particular related to systems with extreme technical reliability specifications, such as redundant systems for which complex situations during work planning and maintenance may give rise to "common mode" errors.

Seen from a modeling point of view, the use of the well-known probabilistic reliability models like THERP* causes no fundamental problems when applied to the normal, well-trained task repertoire. In fact, these models are more empirical models of task sequences in normal work settings, which can also supply the necessary data, than they are cognitive models. However, to deal with the three problem areas mentioned above, real cognitive models in terms of human decision mechanisms are required.

The first problem, estimation of the reliability of human protective functions in response to emergency conditions, depends upon the development of models of decision making in unfamiliar situations under the influence of emotional stress. It is doubtful whether in the near future it will be possible to predict the course of decisions and actions

*THERP = Techniques for Human Error Rate Prediction.

taken in such a situation. Probably, reliability prediction will have to rely on boundary considerations based on the probability of error detection and correction (the feedback design principle).

The second problem, prediction of the probability that operators will interfere with the automatic protection and thus introduce a systematic coupling between a transient condition and failure of protective functions, requires an independent systematic analysis. In particular, a formal analysis must be able to identify possible decision errors which may provoke systematic interactions due to the existence of instructions for activities aiming at different purposes for the same system. An unacceptable act can be the result of a mistake caused by similiarities between the actual situation and another task context of which the act is a natural part. This similiarity may be found at any level of the abstraction hierarchy of Fig. 3; that is, the mistake may be caused by a similiarity in terms of location or appearance of equipment, of the physical function, of the purpose of the systems in overall plant goals or in terms of action sequences. A way to specify the kinds of systematic interference ("systems interaction") caused by the existing many-to-many mappings among purposes, functions and equipment, which should be taken into account in an analysis, can be to base the analysis on a proceduralized search in a consistent and documented description of plant properties at the various levels of the abstraction hierarchy (see Fig. 6).

The feasibility of such an analysis to identify potential for deterministic "systems interactions" violating the assumptions for probabilistic risk analysis depends not only on a model of human decision making, but also on a consistent description of plant properties at the levels of purposes, functions and equipment. The concepts and tools developed for CAM/CAD* should be carefully considered for this purpose.

Finally, the third problem, that of test and maintenance activities on safety systems, should be considered. The basic analysis of performance during these tasks is not posing serious modeling problems. This is, however, the case with the more complex interactions caused by "common mode errors" and errors due to inappropriate decisions taken during work planning and scheduling, for instance, for particular periods like refueling and major overhauls. The decisions in this work condition are much less constrained than the decisions related to control of a causal, physical system, and prediction of performance relies on the development of quantitative models of decisions in management organizations, which is a long-term project.

At present the only solution seems to be to support work planning by *normative models* and analysis of the potential for deterministic conflicts caused by the many-to-many mapping between equipment, functions and purposes mentioned above, and then to control the risk level obtained by feedback based on analysis of event reports.

Explicit formulation of the models used for risk analysis, including the models of operator performance, is important for consistent *risk management*. Figure 9 illustrates how the risk imposed by an industrial process plant, for instance, a nuclear power plant, is controlled in two ways: first by a plant construction based on a risk analysis; and second by risk management (RM), that is, administration of the preconditions of the risk analysis which act as requirements for plant construction and operation. In addition, through the

*CAM/CAD = computer assisted manufacturing/computer-assisted design.

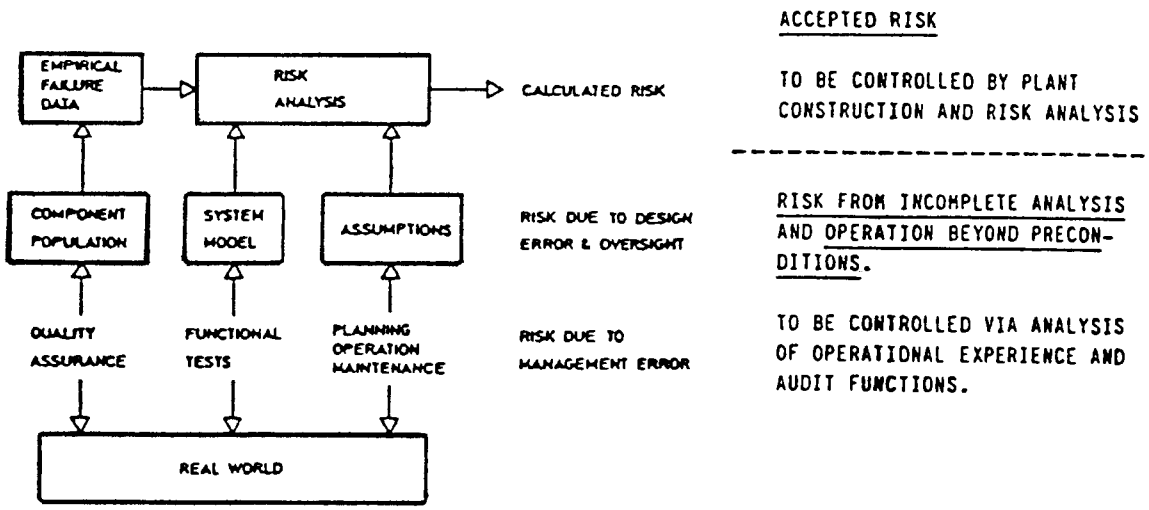


Fig. 9. Control of Risk Imposed by an Industrial Process Plant. The risk contributed by the operation of an industrial plant is composed of an accepted risk identified and analyzed in advance and a risk due to incomplete analysis and insufficient regard to preconditions for the risk analysis. The latter risk is to be controlled by risk management functions comprising quality control, functional tests, inspection, training and instruction of personnel and including analysis of operational experience as a feedback link.

plant lifetime, the preconditions for risk analysis can serve as references for inspections, tests and analyses of operational experience. Decisions made from systematic analysis of abnormal event reports can lead to risk management by means of a "feedback" control function serving to maintain the designer's safety design targets and to reveal oversights and design errors.

The result of a PRA is a calculated risk figure which, if accepted, covers the *accepted risk*. If not accepted, the design has to be modified until acceptance has been achieved. Due to incompleteness and errors during performance of a PRA, however, an *additional risk* may exist, which is not included in the accepted risk. Contributions to this additional risk can also originate from the fact that the real plant and its operation may *depart from the PRA preconditions* because:

- Components employed do not belong to the populations providing the PRA failure data.
- The real plant does not correspond with the models of the plant used for PRA.
- The real plant is not operated and maintained according to assumptions made in the PRA.

After the calculated risk has been accepted, the PRA assumptions, models and data sources are to be used as requirements and references for construction, modification and operation during the lifetime of the plant, that is, as *references for the risk management (RM) functions.*

Some important means for control in RM in order to make sure that the plant is kept in agreement with these references are:

- Quality control.
- Functional tests and inspections.
- Training of operators.
- Issuing instructions for operation, etc.

In order to close the loop from theory to practice, that is, from the PRA to the plant in existence, RM should comprise *analysis and evaluation of failures and abnormal occurrences* by comparison with the references provided by the PRA for deciding whether the abnormalities are included in the accepted risk or whether they indicate circumstances overlooked in the PRA or flaws in risk administration requiring adjustments of practice.

The use of the PRA, including its preconditions as references for RM, obviously requires explicit and user-oriented *documentation of the PRA*, including preconditions, models and data sources. Particularly, documentation of the coverage of the analysis and search methods, that is, what has been included in the search for risk contributions, is important for evaluating operating experience as a basis for RM decisions in the feedback control. The well-documented model and description of the risk identification strategies of the PRA are necessary to decide whether an individual occurrence falls within the accepted risk and, therefore, should contribute to statistical verification or whether it indicates oversights or operational problems which call for special precautions. In this respect, documentation of coverage is considered more important than attempts to reach high degrees of completeness, which may lead to undefined boundaries depending on the creativity of the analyst.

The major part of the human decision making and administrative functions involved in operations management is *not accessible to formal analyses* with the present state of the PRA art. However, errors of management may be significant sources of common mode errors and, therefore, can be important candidates for risk management by feedback control. This feedback control should depend not only on the formal analysis of abnormal event reports by *authorities*, but also on the systematic in-house analysis of log-books, repair reports and similar sources by the *plant staff itself.*

In conclusion, to have a risk analysis of a nuclear installation and a risk management by means of inspections and event reports — which together form an integrated system — it is necessary to have an explicit formulation and documentation of the *predictive or descriptive models* used for risk analysis and of the search strategies used for identification of the risk potential which is analyzed. Concerning the models of operator performance, this involves formulation of empirical models of the task performance in frequent rule- or skill-based tasks and cognitive models applicable for knowledge-based decisions during unfamiliar situations. The more complex and unstructured tasks are represented in the risk analysis as assumptions and preconditions, as, for example, the decision making in management and task scheduling which is unaccessible for analytical assessment. To have a consistent risk control, these assumptions and conditions must be interpreted as *normative models* for planning of plant operation and maintenance. By means of inspections and analysis of event reports, it is then the aim of risk management to evaluate the adequacy of the descriptive models used for the analysis and the fulfillment of the normative models used for operation management.

6. SOURCES OF HUMAN PERFORMANCE DATA[17,18]

There is a growing need for human performance data for design of man-machine interface systems based on new control room technology and for quantification and prediction of human performance in high risk systems and situations. Several different sources of data are at hand, each of them with particular features with respect to problems of data collection and the quality of data it is feasible to collect. The following is a attempt to summarize the features of data collection and analysis as we have met them during a number of cases. The purposes are to provide a basis for the coordination of future analyses and to interrelate results from different sources.

6.1. Data Sources and Data Types

Data sources may belong to either of two categories: (1) nuclear power plants or (2) simulators of nuclear reactors. Within each of these categories, one may distinguish several different types. Here the following distinct sources of data will be considered.

Routine event reports or plant events. Examples of these are the Licensee Event Reports (LERs) which are standardized reports about incidents in US nuclear power plants. The raw data in plant reports are normally checklists and free text comments and are concerned only with the incident in question.

Special human factors post-incident studies of events or plant interviews. These represent a more thoroughgoing analysis of an incident by human factors (HF) specialists and technical specialists. The raw data include, in addition to the raw data from the plant events, interviews with plant personnel, expert assessment of critical parts of the incident, special checklists, computer logs, time line printouts, etc. This kind of studies can follow real plant incidents or simulated incidents on training simulators.

Training simulators. The raw data available from training simulators are normally computer logs and various automatically generated recordings of the operator's performance, as well as the instructor's evaluation thereof. This may be supplemented by checklists (for the instructor), debriefing interviews and discussions based on replays of critical situations, and possibly the operator's self-evaluation.

Research simulators. This category includes experimental studies of operator performance during selected, simulated tasks. A research simulator normally simulates a typical plant rather than a particular plant, and the control console will not be a replica of any particular control room. Research simulators are quite often used to study experimental control rooms. The raw data available from a research simulator include the raw data available in a training simulator, but the recording of the data is normally more flexible, to honor the requirements of various special purpose investigations. In addition to this, research simulators may provide data about operator verbalizations and comments, including operator-experimenter dialogues, tape recorded during the experiment, as well as data from self-confrontations, that is, the operator's retrospective comments made during a replay of the experiment.

6.2. Data Analysis

Just as the types of raw data may vary from one source to another, so may the purpose of the analysis of the raw data depend on the context. In plant events, the purpose is to

identify the characteristics of the situation and of the event that adequately account for what occurred and to identify possible needs for improvement of work planning or instructions.

In plant interviews, the purpose of the analysis is to identify the critical decision sequence which led to the observed performance; this is not radically different from the purpose of plant event analysis, although the emphasis may be put on a understanding of human performance rather than the correction of specific work conditions.

In training simulators, the purpose of the analysis is, of course, to improve the training by improving the feedback the instructor can give to the operator. And in research simulators, the purpose is either to gather data about a particular problem or to evaluate a specific hypothesis or assumption.

This means that the way in which the raw data are analyzed depends upon their type as well as the purpose. Fortunately, this does not lead to, in this case, four completely different types of analysis, but rather several modes of analysis which have a considerable overlap and which are based on the same conceptual background. An important benefit to be gained from a common analytical frame of reference will be the possibility of cross-checking results and the availability of data from all sources for the research on operator performance models. One may, in fact, suggest a common description of the analysis along the lines described below, where each analysis typically has a number of discrete steps with intermediate results which can be characterized as follows. (See also Fig. 10.)

Raw data. This is the basis from which the analysis is made. Some examples of various types of raw data have been mentioned above. The raw data may be regarded as *performance fragments* in the sense that they do not provide a coherent description of the performance, but rather the necessary building blocks or fragments for such a description.

Intermediate data format. This represents the first stage of processing of the raw data. In this stage the data are combined and ordered along a time line to provide a coherent description of what actually occurred. It is thus a description of the *actual performance*, but given in the original terms, that is, as a professional rather than an expert description. The language used is the language from the raw data, rather than a refined, theoretically oriented language.

The step from the raw data to the intermediate data formats is relatively simple, since it basically involves a *rearrangement* rather than an interpretation of the raw data. Hence, special translation aids are not required.

Analyzed event data. In this stage the intermediate data format and the raw data have been transformed into a description of the task or performance using formal terms and concepts. These concepts reflect the theoretical background of the analysis, typically a combination of an information processing theory and a theory for decision making. The description of the performance is still ordered along a time line which is specific to the situation in question. The transformation has, however, changed the description of the actual performance to a *formal description* of the performance during the specific event.

The step from the intermediate data format to the analyzed event data may be quite elaborate, since it implies a theoretical analysis of the actual performance. The translation is one from operator task terms to formal terms. The emphasis is also changed from providing a description to providing an explanation as well. Special translation aids (tools,

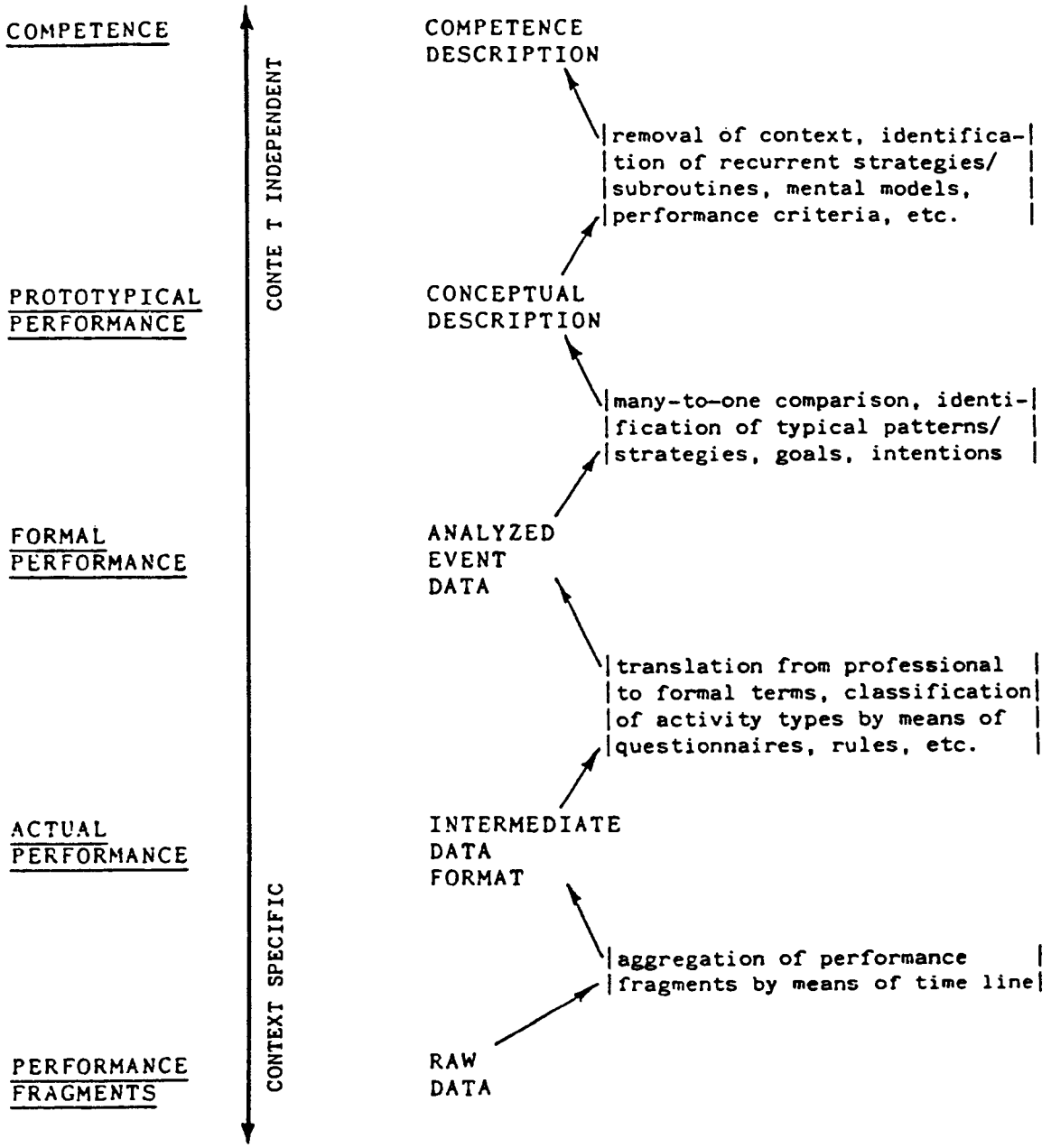


Fig. 10. Illustration of the Several Levels of Generalization in Analysis of Human Performance Data.

methods, and concepts) are therefore required, and the analysis depends on the model framework adopted.

Conceptual descriptions. At this stage of the analysis, the description is no longer specific to a particular event but rather aimed at presenting the common features from a number of events. By combining formal descriptions of performances one may end up with a description of the *generic* or *prototypical performance.* The prototypical performance may still be described as a sequence of activities ordered along some time line, but this is more a *time axis* than a time line referring to a actual situation. On the other hand, a description of the performance in a specific event may be seen as a example or a variation of the prototypical performance. Thus, generic descriptions of human error mechanisms are, in fact, descriptions of typical deviations from the prototypical performance. The step from the formal to the prototypical performance is again one which is quite elaborate, and depends heavily upon the models adopted.

Competence descriptions. This is the final state of the analysis which combines the conceptual description with the theoretical background. The description of *competence* is concerned with the basic concepts, such as mental models, decision strategies, performance criteria, preferences, problem solving strategies, etc., which in a given situation are combined to produce the performance. The description of competence is context-free; it is a description of the behavioral repertoire of the operator independent of any particular situation — though, of course, still restricted to a certain class of situations. As soon as a context is provided, the description of the competence can become a description of the prototypical performance and, pending further information, a description of the typical performance. The competence description is thus essentially the basis for performance prediction during system design.

As before, the step from the conceptual description to the competence description may be quite elaborate and require that the analyst has a considerable knowledge of the relevant theoretical areas, as well as a considerable experience in using that knowledge. It is not so much a question of knowing particular tricks and tools, as of being able to consider the conceptual description in a broad theoretical context.

6.3. Application of the Results of Analysis

The various types of situations (data sources) will normally be analyzed until a level of analysis is reached which is appropriate for the situation. This means that the results at the various levels of analysis may have different applications, both with regard to situations of different types and with regard to situations of the same type. The application of data from training simulators at the different levels of analysis is illustrated on Fig. 11. It is the intent to demonstrate that the analysis necessary to derive generic model data from research will also serve to evaluate various aspects of the training program. Similarly, Fig. 12 illustrates how results from different data sources can be useful together for specific applications at the various levels of analysis.

Since the different data sources are generally studied by separate groups having their own particular purposes, transfer of results and mutual evaluation of models can be difficult, if data analysis is based on ad-hoc models developed separately. The conclusion is therefore that a common conceptual framework and a set of compatible models should be sought to make it possible to derive generic model data from the various programs and to facilitate transfer of results among the various human factors programs within authorities, industry, and research institutions.

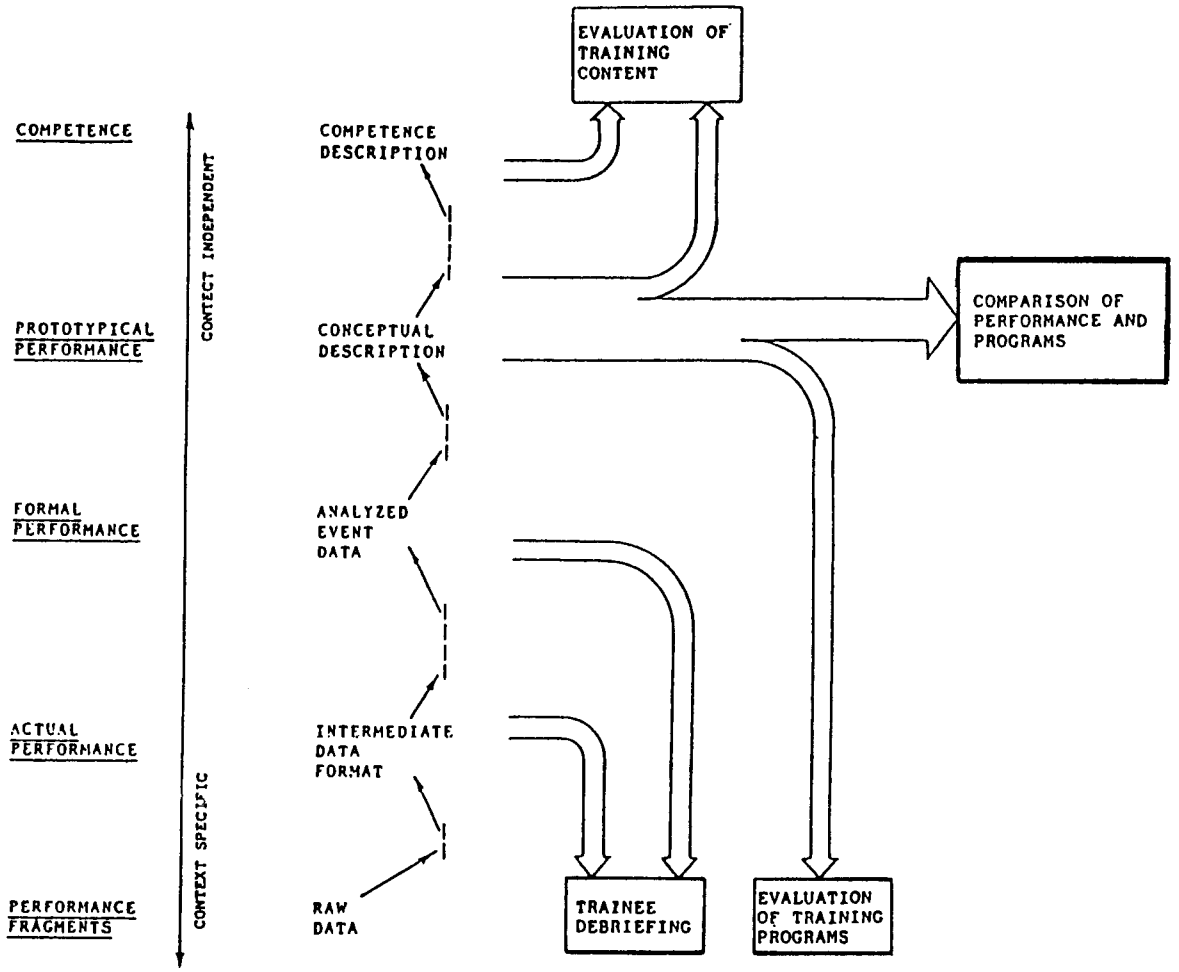


Fig. 11. Illustration of the Possible Application of Results in the Analysis of Performance Data from Training Simulators.

Competence		Requirements for Training Content Operation Selection and Staff Organization		Evaluation of Training Content
Prototypical Performance Conceptual Description	Reliability and Risk Prediction Requirements for Interface, Instruction and Training Systems	Accident Modeling	Evaluation of Designs for Interface and Instructions	Evaluation of Training Programs
Formal Performance Analyzed Data	Requirements for Individual Sessions for Experimental Studies and Training			Trainee Debriefing
Actual Performance Intermediate Data	Corrective Measures for Plant Operation	Studies of Situations, Scenarios, Mental Workload		Trainee Debriefing
Performance Fragments Raw Data				
Form / Source	Event Reports	Incident Analysis	Research Simulation	Training Simulator

Fig. 12. Applications of Data from Different Data Sources. The table illustrates different applications of information obtained at the various levels of analysis of human performance data from various sources, together with the potential of transfer of results.

REFERENCES

1. J. Rasmussen, "Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models," (to be published in IEEE Trans. Syst. Man, and Cybern.).
2. G. Bateson, *Mind and Nature*, Elsevier-Dutton Publishing Company (1979).
3. A. Sloman, *The Computer Revolution in Philosophy: Philosophy, Science, and Models of Mind*, The Harwester Press, Sussex (1978).
4. J. Rasmussen and M. Lind, *Coping with Complexity*, Riso-M-2293, Riso National Laboratory, Denmark (1981); paper presented at the European Annual Conference on Human Decision and Manual Control, Delft, 1981.
5. D. A. Norman, *Steps Toward a Cognitive Engineering: Systems Images, System Friendliness, Mental Models*; paper presented at Symposium on Models of Human Performance, ONR Contractor's Meeting, La Jolla, CA (UCSD), June 19, 1981.
6. J. Rasmussen and M. Lind, *A Model of Human Decision Making in Complex Systems and Its Use for Design of System Control Strategies*, Riso-M-2349, Riso National Laboratory, Denmark (1982); paper presented at the American Control Conference, Arlington, VA, June 14-16, 1982.
7. L. P. Goodstein and J. Rasmussen, "Man/Machine System Design Criteria in Computerized Control Rooms," abbreviated version in *Proc. of ASSOPO 80 - an IFIP/IFAC Symposium*, Trondheim, Norway, June 16-18, 1980.
8. L. P. Goodstein, *Computer-Based Operating Aids*, Design 82, Birmingham, UK (September 22-23, 1982).
9. M. Lind, "Generic Control Tasks in Process Plant Operation," *Proceedings of the European Annual Conference on Decision Making and Manual Control*, Bonn, Germany, 1982.
10. M. Lind, *Multilevel Flow Modelling of Process Plants for Diagnosis and Control* (to be published in 1982).
11. W. B. Rouse, p. 214 in *The Man-Machine Interface and Human Reliability: An Assessment and Projection*, 1981 IEEE Standards Workshop on Human Factors and Nuclear Safety, August 30–September 4, 1981, Myrtle Beach, South Carolina.
12. P. M. Fitts and M. I. Posner, *Human Performance*, Brooks/Cole Publishing Company (1962).
13. K. D. Duncan, "A Simulator and Training Technique for Diagnosing Plant Failures from Control Panels," *Ergonomics* **18**, No. 6, 627-41 (1975).
14. K. D. Duncan, "Training for Fault Diagnosis in Industrial Process Plant," in *Human Detection and Diagnosis of System Failures*, Jens Rasmussen and William B. Rouse (Eds.), published in cooperation with NATO Scientific Affairs Division, Plenum Press, New York and London.
15. W. B. Rouse and R. M. Hunt, "A Fuzzy Rule-Based Model of Human Problem Solving in Fault Diagnosis Tasks," *Proc. of the Eighth Triennial World Congress of the International Federation of Automatic Control*, Kyoto, Japan, August 1981.
16. J. Rasmussen and O. M. Pedersen, *Formalized Search Strategies for Human Risk Contributions: A Framework for Further Development*, Riso-M-2351, Riso National Laboratory, Denmark.
17. E. Hollnagel, O. M. Pedersen, and J. Rasmussen, *Notes on Human Performance Analysis*, Riso-M-2285, Riso National Laboraory, Denmark (March 1981).
18. E. Hollnagel and J. Rasmussen, *Simulator Training Analysis: A Proposal for Combined Trainee Debriefing and Performance Data Collection in Training Simulators*, Riso-M-2301, Riso National Laboratory, Denmark (July 1981).

Paper 2

A FRAMEWORK FOR MODELING SUPERVISORY CONTROL BEHAVIOR OF OPERATORS OF NUCLEAR POWER PLANTS

S. Baron, C. Feehrer, R. Muralidharan, R. Pew, and P. Horwitz
Bolt Beranek and Newman, Inc.

1. INTRODUCTION

The importance of modeling the human-machine system has long been recognized, and many attempts have been made to estimate the operator's effect on system performance and reliability. Traditionally, in the nuclear industry these efforts have focused on reliability modeling and have involved estimates of human "error rates" which, together with event tree and fault tree analyses, have led to numerical predictions for the frequency of occurrence of particular incident scenarios. Reliability models have made major contributions to analysis, and considerable effort is still being devoted to obtaining data concerning human error for such models.

The development of reliability models has been aimed at providing the means for exploring the physical consequences of specific classes of human error. However, the total impact of human performance on system operation and the adequacy of existing design and operating standards cannot be adequately captured or assessed by simple error probabilities, or even by the combination of such probabilities. The behaviors of relevance are supervisory in nature, with a substantial cognitive component. They include planning, monitoring, control, detection and diagnosis of failures, etc. Moreover, these behaviors are dependent upon, and intricately entwined with, the system itself (that is, the plant design, the displays, the procedures, and the training methods).

The broad requirements for a model of such human supervisory control are extensive and suggest that a highly sophisticated computer model will be needed. This model must address human cognitive and psychomotor behavior in both continuous and discrete tasks. The model must be capable of generating meaningful estimates of crew performance and workload. Moreover, the crew model must interact with the elements of the system model in a closed-loop fashion, so that the temporal interaction associated with accomplishment of the various activities engaged in by the crew is affected by the system variables (and changes in them), and conversely.

Fortunately, the need for computer models of this type has been anticipated to some degree. In recent years, and in a parallel development concentrated largely in the aerospace industry, important strides have been made toward the development of models for the performance of the human component of a system. This research has matured beyond the simplest "human-as-servomechanism modeling" that characterized its early years to the point where tasks involving multivariable control, monitoring, and decision making have been modeled successfully (see, for example, refs. 1-5). Recently, efforts have focused on models for supervisory control in multitask environments that involve a broad range of operator behaviors and interactions between operators.[6–8]

The approach to modeling human supervisory control that is emerging from this effort has several features that suggest that it can provide quantitative performance data for setting of *standards* relating to the impact of training, computer aids in the control room, procedure specifications, etc. These quantitative results appear to go well beyond any that can be derived from task analysis. Furthermore, the approach treats the human cognitive processes and activities in ways that can account for human error, but in a considerably more sophisticated and casual manner than by simple data statements concerning that error.

The purpose of this paper is to provide a brief overview of the approach employed in developing such supervisory control models; of some proposed specializations and extensions to adapt them for the nuclear power plant case; and of the potential utility of such a model. A more detailed presentation of this material is given in ref. 9.

2. THE PROPOSED SUPERVISORY CONTROL MODELING APPROACH

There have been several approaches to modeling human performance in the control of complex systems and each has advantages and disadvantages with respect to quantitative analysis (see refs. 10, 11 for reviews). The reliability, network and HOS (human operator simulator) models[12–15] are adept at dealing with procedural activities and discrete actions and attempt to synthesize molecular human performance data concerning these activities. They are much less capable of dealing with human cognitive activities and do not attempt to deal at all with dynamic, closed-loop behavior of the person-machine system. On the other hand, control-theoretic models (as described in refs. 1-5, for example), have emphasized the goal-oriented, continuous, closed-loop interaction of the operator with the environment (in an information-processing and decision-making sense, even when continuous control is not required), but have largely ignored discrete actions and procedures.

The supervisory control modeling framework that is suggested as a basis for developing a model for a nuclear power plant operator/crew is an attempted symbiosis of some of the features of several previous modeling approaches. Very briefly, the approach begins with top-down, control-theoretic, holistic models of monitoring, detection, decision making and control and imbeds them in a closed-loop simulation framework that includes a model of the system. This permits external events as well as the decisions of the operator model to trigger contingent response and activities. Those activities of a discrete nature are treated in a manner that is analogous to the treatment of subtasks in the network approach — that is, a task completion time is associated with the activities.* It should be emphasized that discrete activities and the corresponding completion times are generally aggregated at a much greater level than in bottom-up models; during the time the operator is engaged in any of these activities it is assumed that s/he is "locked up" and cannot perform other functions or tasks.[7]

In addition to the holistic subtask models that are the starting point for these supervisory control models, the approach retains the formalized goal structures associated with the control-theoretic models. In particular, the selection of activities to which processing resources are assigned is based on the maximization of expected net gain (the difference between the expected gain for an action and its expected cost, including the potential cost of delaying other actions).

*As a rule, it is assumed that there are no errors in execution of discrete actions or procedural steps. If such errors are of interest, they are introduced directly (that is, preprogrammed) and the subsequent responses of the model are observed to evaluate the consequences.

The proposed approach also incorporates procedures and rule-based activities in the model. In general, the procedures have the form: IF (SITUATION), THEN (ACTION); or IF (SITUATION), THEN (SITUATION). These procedures are based on information available in manuals and/or from operator interviews. In the model, the determination of whether a particular situation obtains is based on the human's (model's) assessment as determined from the information-processing models. Thus, the assessment reflects human information-processing capabilities and limitations. If more than one situation is deemed to be "true" at a given time, then a decision is made as to which procedure to execute, based on the perceived consequences of execution or deferral and the goals of the operator. Procedures involving continuous behavior maintain their own time scale, whereas those involving discrete actions use input data for completion times. This view of the model suggests that it may be considered to be a complex version of the production system concept employed in Artificial Intelligence (AI) and in the qualitative problem-solving models.[16,17]

A final point of convergence between the proposed approach and that of more psychologically oriented methods is the use of preliminary task analysis for initial definition of procedures and goals and subgoals. It is likely that the kind of task and protocol analysis utilized by cognitive psychologists and AI specialists would be an important component of premodeling analysis for the proposed system-oriented supervisory control model of the nuclear power plant operator.

An analysis of fundamental human operator functions in a variety of tasks related to system operation suggests a basic commonality that is summarized in the following list:

1. Monitoring the display variables, according to an attention allocation strategy dictated by the needs of the task.
2. Assessing the present situation, based on the monitored information, the operator's inherent knowledge of system operation, and a knowledge of alternative situations.
3. Deciding to take (or not take) some action and making this decision on the basis of the assessed situation, the basic task objectives, the procedural and other means available to effect this decision, and the expected consequences of the various actions.
4. Acting to implement the decision, either in terms of communicating the intent of the decision, or by directly observing or controlling the system in an appropriate manner.

This verbal description can be given somewhat more substance by casting it in a block diagram or flow chart form, as shown in Fig. 1. As drawn, the "system" comprises the upper portion, while the "operator" comprises the lower. The system is affected both by the direct actions of the operator and, perhaps, by the actions of others (for example, other crew members). The system displays include not only an indication of the system states, but also procedural information provided by manuals (and/or advanced displays) and the communications channels for obtaining relevant information from other operators, or from the system itself (for example, an auditory alarm). The portion of the diagram representing the operator is deliberately simplified to show the basic structure of functional processing. Thus, it should be recognized that various pathways can interconnect the several

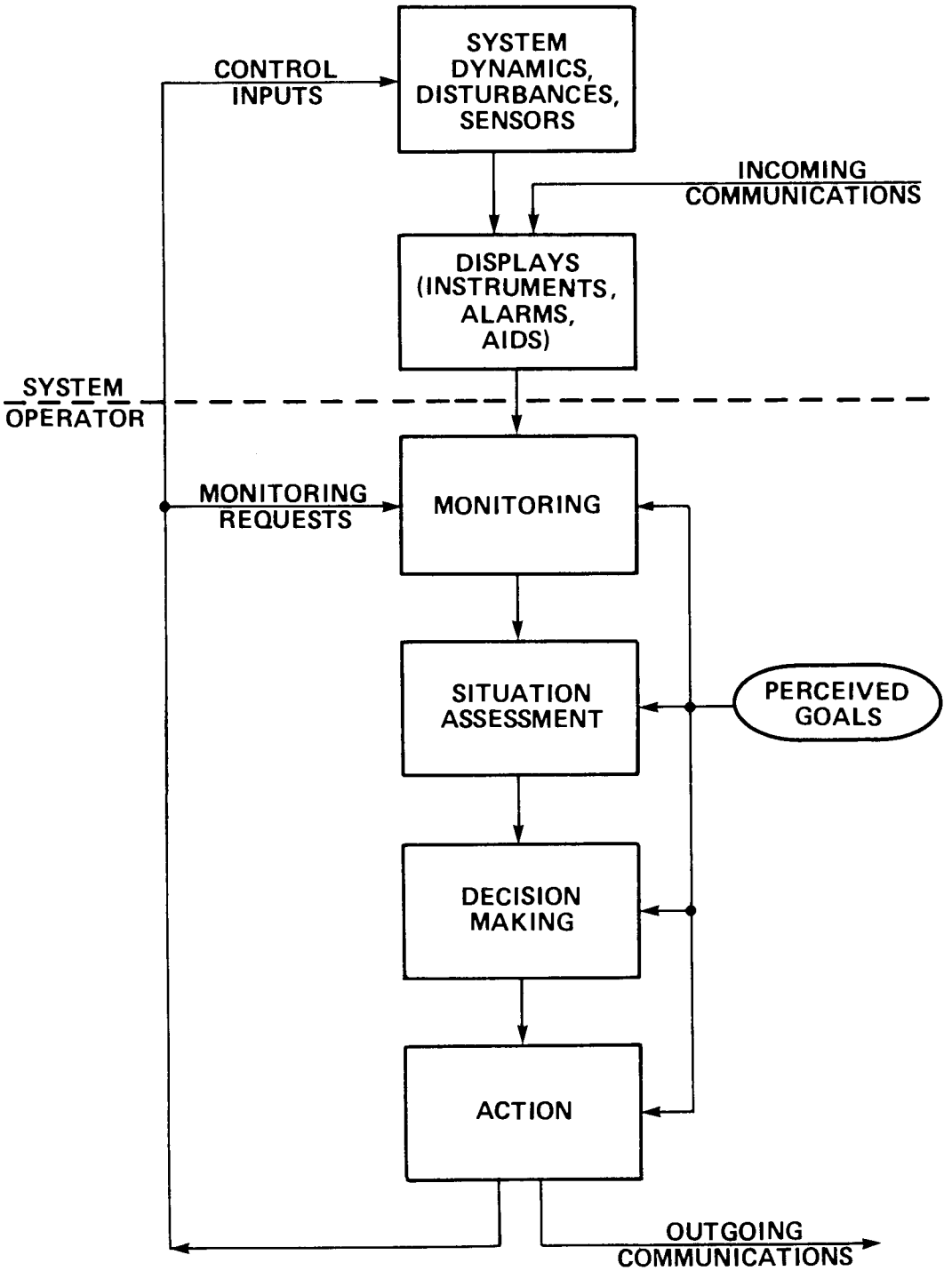


Fig. 1. Supervisory Control Model Structure.

functions shown and allow the result of one function (for example, decision making) to affect the way(s) in which another is carried out (for example, monitoring).

The generic framework and structure has been used recently to develop specific supervisory control models for an en route controller of (multiple) remotely piloted vehicles; the air crew of a commercial aircraft executing the procedural and control activities associated with descent, approach and landing; and the commander of an anti-aircraft artillery team confronted with the problems of assessing the threat situation and selecting an appropriate mode for the weapon systems in response to this situation.[6–8]

It is important to note that the models developed within this framework are not designed to deal with the "knobs and dials" questions of human engineering. Rather, they focus on the cognitive and procedural aspects of the operator's interaction with the system. Thus, they are intended to address questions concerning the availability of appropriate information for decision making and control instead of, say, the size of the numbers on a display or the viewing angle, etc. In particular, the models are suited to an examination of the effects on closed-loop performance and operator workload of plant dynamics and controls, alternative information sets, priority and/or goal structures, different mental conceptions of the system, and variations in procedures.

3. A CONCEPTUAL MODEL FOR THE NUCLEAR POWER PLANT OPERATOR

3.1. Overview

A block diagram illustrating the major facets of the closed-loop system of interest is presented in Fig. 2. Here, the control room operators are shown to communicate with the plant and its control, protection and safety systems through two interfaces. One of these, the "information interface," is intended to include all the visual displays (including manuals), auditory warning, and other sources that can convey information to the operators concerning the processes and status within the nuclear plant and the control, protection and safety systems. The other interface, the "control interface," includes all of the actuators, such as push buttons, switches, control handles, communication devices, etc. that can be employed by the operators to affect the plant behavior.

We have indicated in the operator block that more than one member of the crew can be considered in this fashion. We do not intend to be specific with respect to different crew members in our subsequent discussions, but we do believe that it is important and feasible to model multioperator crews. To specialize the model to a given member of the crew, we must restrict his/her task appropriately and define the portion of the system s/he interacts with and the procedures applicable to his/her role. Then, one must provide for communication between the crew members, something that has been done in the previous supervisory control model (see refs. 7-8).

In subsequent sections, we indicate the nature of potential models for the plant/environment and for the operator. More details are given in ref. 9.

3.2. Mathematical Model of the Plant Environment

It is necessary to include a mathematical description of the nuclear power plant/environment in the supervisory control simulation model. The degree of detail and

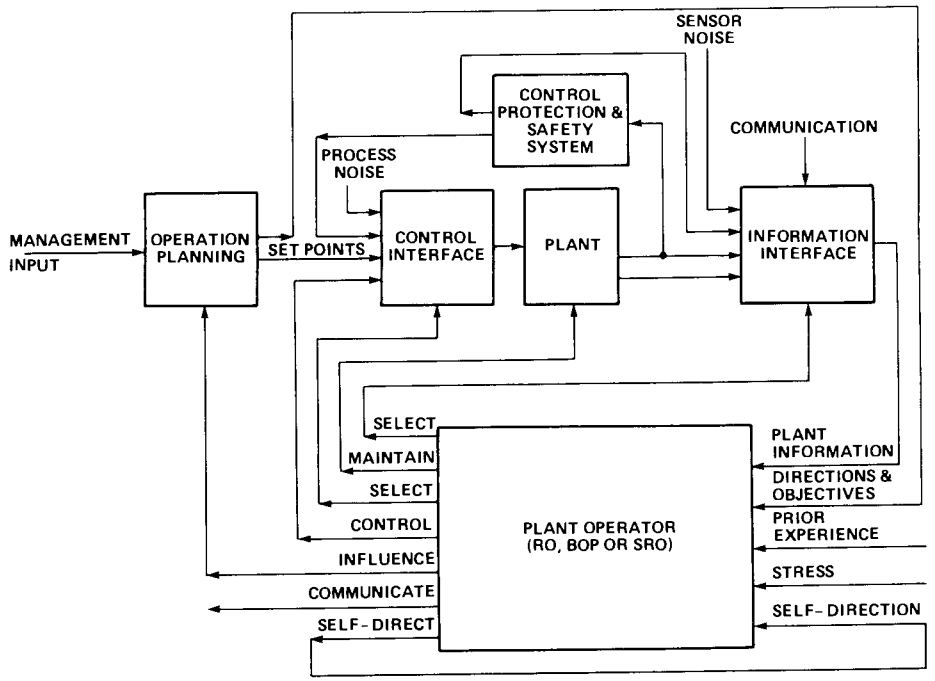


Fig. 2. Overview Block Diagram of the Nuclear Power Plant Operator in His Environment.

the level of complexity of the plant model will depend on the specifics of the issues to be addressed. However, in general, the plant model must include those factors that are needed to perform a closed-loop system analysis, as listed below:

- A state-variable description of the processes, including automatic control and engineered safety features.
- A description of potential disturbances (known and unknown).
- A description of the instrumentation and display information provided for the crew (including information content, alarm set points, instrument or sensor noise, update rates and failure nodes).

The model must provide for simulation of normal and off-normal conditions of interest. It is probably unnecessary to simulate accurately the very fast responses of the plant that are beyond the ability of the operators to control. Finally, though a "First Principles" model is desirable, empirical models can also be used.

The state-variable description of the plant will consist of differential (or difference) equations describing the time evolution of plant states (such as temperatures and pressures in various parts of the plant) in response to both control and disturbance inputs. In ref. 9, a reasonably complex plant model, including models for the reactor, primary coolant pressure and feedwater control systems, is derived; this model involves 30-40 state variables. The information available to the operators on the various displays, alarms, etc. must also be defined, along with an indication of the relative location of various displays and controls.

3.3. Model for the Human Operator

The conceptual model for the nuclear plant operator is illustrated in Fig. 3. In what follows, we discuss briefly the elements of this model. It is always difficult to "parse" human behavior, and one might well imagine breakdowns different than those shown and discussed below. Certainly, major activities like situation assessment transcend the boundaries indicated. However, we believe that, on the whole, it is important, for clarifying thinking about behavior, to adopt some structure. The one we have selected here has proven useful before.[7–8]

3.3.1. Display Processor

The display-processor portion of the model has two functions: it implements the "conscious" observation decisions of the operator by selecting the appropriate displayed quantity, and it accounts for sensory and processing limitations associated with observation.

It is generally agreed that humans are severely restricted in the number of activities they can perform simultaneously. We assume that the operator is a single-channel processor of information. In terms of observing displayed information, we will assume that only one information "display" can be "attended to" at any time. The selection of a particular source of information is governed by goal-oriented processes: thus, it will depend on the purposes for which information is being gathered.

As illustrated in Fig. 3, the operator processes both visual and auditory information coming from the various displays and the communication channels. There are several limitations associated with the acquisition of displayed information. Obtaining visual information requires fixation of the display, and this may involve a significant amount of time in a large control room. In addition to the acquisition time, there will be a perceptual noise associated with observation of visual information; this noise may be based on the observation noise model used in the Optimal Control Model.[1]

Auditory information is treated in a different manner than visual information. Briefly, we assume that auditory information is acquired instantly and that the correct message is stored in a short-term memory buffer (but is not "processed"). At the same time, the event detector (see below) is "notified" that a message is waiting to be processed. The subsequent processing of this message will, in general, depend on the nature of the message (alarm or communication) and on the time elapsed since its occurrence. If the message has a sufficiently high priority, the operator's activities will be interrupted and the auditory display will be selected for processing. The information in the buffer may also disappear or degrade in reliability with time, depending on the nature of the message. This treatment of auditory information has the advantage that alarms are priority interrupts, but that they may be missed or "unattended to" during times of high activity or workload stress (for example, when there are many alarms).

3.3.2. Information Processor

The information-processing portion of the model incorporates important extensions of previous models for this application. It uses a basic predict/correct logic based on a mental model of the plant. The logic is analogous to that of the Kalman filter model used and validated in previous models for operator control and monitoring.[1–5] However, assumptions on linearity and on the mental model have been eliminated. The logic provides a

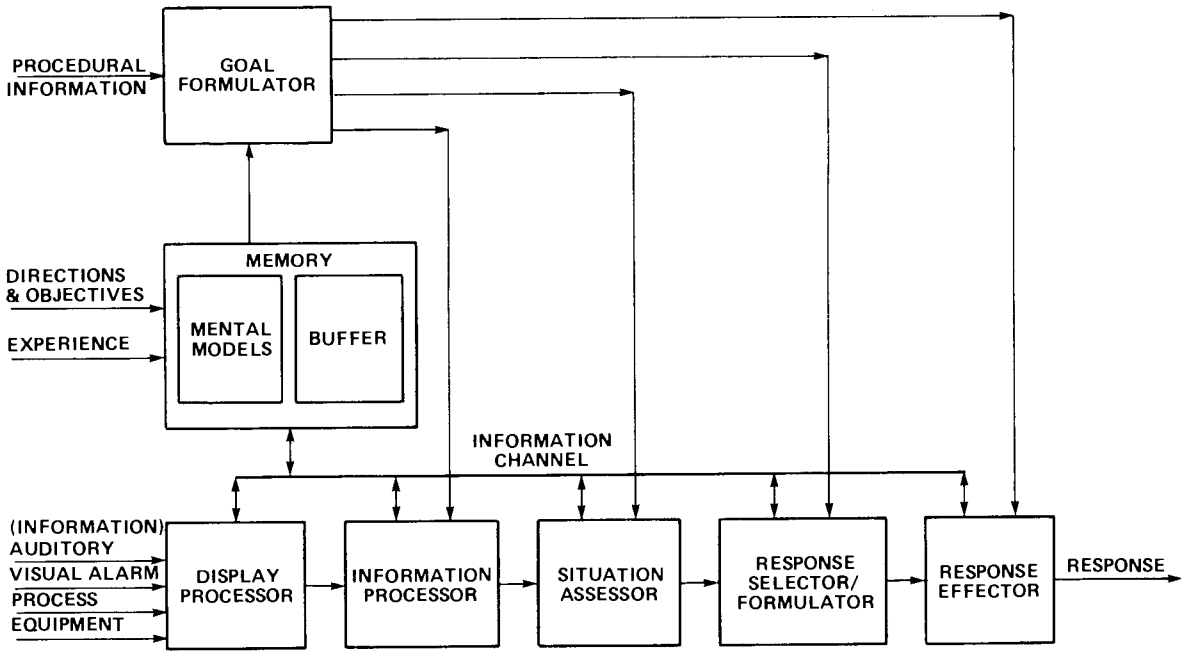


Fig. 3. Conceptual Model of the Nuclear Power Plant Operator(s).

plausible algorithm for state-estimation and prediction. The mental models may be used for fast-time prediction (or "thought experiments") to evaluate alternative options for action. Thus, some of the planning functions described in Sheridan[18] and in Kisner, Fullerton and Frey[19] can be incorporated. Furthermore, the effects of different mental models (as might result from different training regimens or different roles) on overall performance may be analyzed.

A discrete-event detector is also part of the information-processor portion of the model. The envisioned event detector is more complex than those employed in the models described in refs. 3-8 in that it is expected to employ a range of detection algorithms that vary with respect to computational sophistication. The simplest algorithms involve "attending to" and processing an out-of-tolerance condition, whether annunciated or not. The more complex algorithms are expected to be drawn from previous detection models (for example, refs. 4-5) or from other system-theoretic approaches to failure detection (for example, refs. 20-21). These methods all employ the mental models of the plant to generate expectations about plant behavior. The differences between the two, called residuals, are tested statistically, in any of several ways, for systematic deviations. The range of information-processing algorithms incorporated for event detection will allow for the modeling of most or all of the monitoring functions described by Kisner, Fullerton and Frey.[19]

3.3.3. Situation Assessor

The situation-assessor portion of the proposed model is also an expansion of the previous models. As with the event detector, a hierarchy of algorithms is contemplated for situation-assessment activities associated with the diagnosis of a problem. At the lowest level is a template-matching scheme which checks symptoms against a template that is part of a procedure residing either in a manual or in memory. The more sophisticated assessment algorithms would employ alternative mental models corresponding to the hypothesized situations and, by comparing the observed data with model outputs, would select the most likely hypothesis. The range of algorithms provided for situation assessment should provide the capability for systems-oriented modeling of both the symptomatic and topographic search strategies discussed by Rasmussen.[22]

The model of the diagnostic process viewed *in toto* could exhibit or predict known human errors under appropriate conditions. For example, under time stress, a diagnostic decision may be forced before evaluation of an adequate number of template elements has been accomplished — thus leading to an error. Or, at the opposite end of the spectrum, operator biases or uncertainties may lead to an unwillingness to abandon a given hypothesis or to make a decision without "more data," that is, to cognitive lockup.

3.3.4. Response Selector/Formulator

The "procedures" are the means by which the operators organize and carry out their monitoring, situation assessment and control responses so as to accomplish their objectives. Procedures exist in "manuals" or they reside in memory, having been learned through training and/or experience. We allow both highly structured formal procedures and other less structured, but goal-directed, responses (which we will also call procedures, for convenience).

A formal procedure is a specific sequence of tasks or actions (perhaps of length one) together with the "situations" that trigger those actions. Example procedures of this type are Immediate Action Procedures and Safety Injection React Procedures.

Informal procedures are those responses that are to be formulated and employed when the existing situation does not "fire" a known formal procedure. They are included partly in an attempt to model some knowledge-based behaviors and partly for modeling convenience. Examples of informal procedures are: selection of control laws; monitoring to detect anomalies during normal (unalarmed) conditions; situation diagnosis when procedural diagnostics fail to resolve a situation; and message decoding (the process of "reading" the auditory buffer).

Major decision making at several levels takes place in the procedure selector/formulator block. The decisions are primarily concerned with continuation or termination of a response or a procedure and with selection among competing or alternative procedures or options to achieve a given goal. In all but the immediate-action case, the operator is confronted with competing alternatives. The choice he makes is assumed to be rational and would be computed in the model on the basis of an expected net gain (ENG) calculation, as was done in the supervisory control models described in refs. 6-8. The ENG from a particular action is obtained by subtracting the cost of that action from its expected gain. The expected gain itself is the difference between the expected cost of

events when no action is taken and the expected cost of events that may arise after the action is taken. The rational choice is to select that action which has the greatest ENG. In evaluating these alternatives, the model would have to "predict" the consequences of various actions using its mental model.

3.3.5. Response Effector

The selection and execution of a procedure or other response results in an action or a sequence of actions. Three types of actions are possible: control, observation and communications. The control actions include any continuous manual control inputs to the plant (such as varying boron concentration or auxiliary feed flow) and discrete control settings (set points, switches, etc.). Monitoring requests result from requirements for specific information; these requests are input to the display processor which selects the particular requested information source. We note that verifying that a variable is within limits may not require an actual instrument check if the operator already has a "confident" internal estimate of that variable. Communications are verbal requests or responses resulting from interactions with other crew members. Mostly, they involve requesting or transmitting state or situation information.

Associated with each action is a time to complete the required action. (It is possible to allow for a probabilistic distribution of action times.) When the operator decides to execute a specific procedure, it is assumed that he is "locked in" to the appropriate mode for a specified time. For example, if the procedure requires "checking" a particular instrument and it is assumed that it takes t seconds to accomplish the check, then the display processor will not attend to other information for that period, nor will another procedure be executed. The one exception to this rule is that certain high-priority auditory alarms can interrupt the activity. Note that the time required to obtain information would depend on the action required (eye movement, walking to another part of the plant, communications, etc.).

3.3.6. Memory

The various processes considered in the model require information (event and situation lists, templates, mental models, etc.) that is presumed to reside in the operator's memory. It is not our intent to model the storage and retrieval of information as a process, at least initially. Rather, memory will be viewed essentially as a "storehouse" of information that can be referred to, added to, or modified.

Limits on short-term memory of auditory information and other detected events are included in the model. In addition, "memory" of the state estimate will degrade with time following an observation. However, the time and "mental effort" required to retrieve information from memory have not been considered. It may be possible to estimate these quantities based on psychological considerations, but we suspect their impact on problems of interest to be of secondary importance, as compared to the short-term memory limitations that will be incorporated.

3.3.7. Goal Formulation

The goals of the operator play a very significant part in determining the assessments, decisions and control strategies of the modeled operator. In previous applications of the

proposed approach, it has been largely possible to employ a unitary goal structure. In this application, a more hierarchical goal structure is probably required. An example of such a structure was included in the discussion.

4. CONCLUDING REMARKS

It is interesting and useful to note that the model proposed is capable of generating examples (or analogs) of the skill-, rule-, and knowledge-based behaviors of Rasmussen's qualitative model. Skill-based behavior is evidenced when an "alarm" or a "event" is detected that dictates an immediate (and assumed to be well-known) response. Many of the procedural activities, including several associated with detection and diagnoses, will be examples of rule-based behavior. Finally, the model will exhibit knowledge-based behavior as it ascends the hierarchy of detection and assessment algorithms to those that use mental models of the plant. Such behavior will be evidenced as well in planning and decision-making functions, which are also based on mental models and on the complex goal structures to be incorporated in the model.

It should also be mentioned that there should be no difficulty in using the model to simulate manual control as, for example, in a startup. We have not discussed this portion of the problem because it was deemed to be of secondary interest at this time and because of its closer relation to past work.

The model could provide a number of interesting and useful outputs. Because it is a simulation model that includes the plant, it can provide time histories of all plant variables included in the model. Activity time lines showing operator actions would be available to determine periods of excessive demand or activity, etc. These time lines would be significantly different from those traditionally employed in human factors analysis in that they would be generated in a closed-loop fashion.

Of course, the simulation results can be analyzed to derive a number of important operator performance metrics. Thus, it would be possible to compute error rates, task completion rates and frequencies via Monte Carlo simulations. It must be emphasized that the errors that occur are *predicted* by the model and arise as a result of the interaction of cognitive limitations and task-related parameters. Thus, for example, the presence of time-stress will be predicted, as will the errors or "near-errors" that result from that stress.

Also available would be a number of model outputs that could be extremely useful for diagnostic purposes. Some of these would be difficult, if not impossible, to measure experimentally, but all could be examined in the model. Many would provide insights into operator "margins" that are analogous, in some sense, to plant margins. Table 1 gives a sample of some unusual performance measures that might be available from the model and the relationships of the measures to some human factors issues.

A brief analysis of the data requirements for the model suggest that much of what is needed may be found in the literature. Plant model data of the type required is available in several sources. Because the operator model draws on previously validated submodels for many tasks, there exists appropriate operator data for these tasks. Most of the remaining data requirements probably can be satisfied by premodeling task analysis and through operator interviews. Naturally, as the model evolves, we do expect unmet data requirements will emerge, but we do not expect that these will be more demanding than for alternative approaches.

Table 1. Some Potential Performance Measures

Human Factors Issues	Major Focii in Model	Performance Measures
• Crew Complement, Organization, & Management • Increase in number of personnel • Centralized vs. decentralized control • Preassignment of crisis locations	Memory Situation processor Procedure selector/formulator	Accessibility of signals in aggregate memory Number and variety of hypotheses considered in situation assessment Appropriateness of procedure selected/formulated Probability of interruption
• Procedure Design & Accessibility • Lengths of diagnostics (size & content of templates) • Memorized vs. written procedures	Information processor Situation assessor	Accuracy of diagnosis Probability of interruption
• Gross Characteristics of Control Room Layout	Display processor Response effector	Accuracy of estimates Response execution time "Lockup time"
• Display Design • Filtering of secondary and tertiary alarms • Scaling • Extent and nature of instrument redundancy • Raw vs. higher order data (+ prediction)	Memory Information processor Situation assessor Procedure selector/formulator	Accuracy of estimates Number of elements contained in memory Number and variety of hypotheses formulated Content of event detector Appropriateness of procedure selected/formulated
• Training • Thoroughness of plant knowledge	Mental models Goals formulator	Accuracy of estimates Accuracy of diagnosis Appropriateness of procedure selected/formulated

In summary, the proposed model seems feasible to develop, embodies important features of a variety of approaches to human performance modeling, and should prove useful for analyzing a wide range of human factors problems of concern to regulators and designers. We close this paper with a brief discussion of some of the potential applications areas for the model.

We have identified six areas within NRC's mandate that we believe would benefit substantially from a supervisory control modeling capability of the type we envision: risk assessment, procedure evaluation and certification, control room design certification, training and licensing requirements for operators, crew organization and post-incident analysis.

Risk Assessment. There are two basic approaches to estimating the effect of human error on the probability of particular accident scenarios. One of these starts by breaking down the role of the operator into elementary actions (which may include instrument monitoring, procedure lookup, and communication, as well as control actuation) and estimating the probability of faulty execution. A complementary approach, little used in practice for reasons of expense, is to run as many trained operator crews as possible through a given accident scenario on a simulator and tabulate the results. This has the advantage of being more realistic, particularly with respect to common-mode failures or high-level errors such as misdiagnosis of the problem, but it is difficult, time-consuming, and expensive to collect enough data in this way to enable one to draw reasonable conclusions.

The development of a closed-loop simulation model of the type proposed would provide a very important tool for performing risk analyses. It would make possible, at reasonable cost, the sort of Monte Carlo calculations (in which identical operators are run many times through identical transients) that are difficult or impossible to perform with real operators. To be sure, the results obtained with the numerical model would have to be validated against real human performance, and might, in fact, differ in some degree. Nevertheless, experience with similar models obtained in other areas of application give one the expectation that a sufficiently high level of realism is obtainable to make possible the sort of risk assessment studies that we are contemplating.

An added advantage of using a model, rather than a human subject, to perform risk assessment studies lies in a model's ability to shed light on the critical question "how close did the crew come to having a major accident?" In reviewing performance, it is not possible to estimate reliably how close human subjects came to making a mistake — they either make it or they don't.* To the extent, however, that a numerical model "becomes confused" when, for example, its situation assessor comes up with two or more equally likely hypotheses, or input data are received that appear to be inconsistent with a currently hypothesized situation, a detailed examination of the operation of the program during the run will reveal a potential source of difficulty.

Procedure Evaluation. The important effect that alternative procedures can have on the overall safety of a plant is well recognized, but, paradoxically, exceedingly difficult to verify quantitatively. The use of human subjects suffers from the major methodological deficiency that subjects retain information about the procedures they have been taught, and cannot easily be "reprogrammed" to forget the old set and use only the new. Precisely such reprogramming is made possible by the use of a model, and numerical experiments involving both subtle and radical changes in procedures are thus made possible.

---

*Detailed data analyses and interviews conducted soon after the performance of tasks may provide insights into available margins of safety, but they are rarely accomplished.

Diagnostic procedures are often written in imprecise terms (for example, pressure increasing *rapidly*). An appropriate model can simulate the effects of different interpretations of such statements and evaluate the need for more precision.

The supervisory control model we have described is sophisticated enough to allow experimentation with alternative high-level goals as well. The inevitable tension, for example, between the often conflicting goals of safety and availability of the plant can be mapped into such models, and the effects of a shift in emphasis on operator performance can be examined.

Control Room Design. Much of the human factors work currently sponsored or contemplated by NRC centers around the issue of improvements in control room design. A difficulty encountered in evaluating alternative proposed improvements stems from the fact that there is little reliable quantitative data on which to base such evaluations. While such highly subjective (and idiosyncratic) considerations as, for example, a preference for analog over digital displays may not be taken into consideration by the proposed model, its emphasis on information processing and situation assessment, along with its ability to simulate the effects of cognitive overload and attention sharing, make it an excellent tool for guiding the thrust of research directed at improved control room design. In particular, proposed operational aids could be examined with regard to their impact on both specific operator functions and overall operator/system performance.

Training and Licensing of Operators. A great deal of attention has been given to the question of just what a licensed reactor operator should know in order to do his job effectively. Although a *de facto* consensus on this point has been arrived at over the years within the nuclear power industry, the Three Mile Island incident has thrown the validity of traditional training courses and certification requirements into considerable doubt. Since any model necessarily contains within it a body of "knowledge" of plant characteristics, operating procedures, and the like, which in its human counterpart are the result of training, it provides an excellent laboratory for experimenting with the effects of that training on performance. In addition, the numerical simulation may be used as a normative model for performance based on training, and, as such, may prove useful in devising suitable measures of performance to be used in licensing of reactor operators.

Crew Organization. Several alternative structures for crew organization are possible. The optimum organization will depend on the plant characteristics, the control/display design, the procedures and training adopted, and the interaction between operators. Because of this interdependence, it is a very difficult systems analysis problem to determine the best organization. Although the supervisory control model is not likely to examine social questions regarding crew organization, it can be used to study the impact of alternative structures on the individual and collective information awareness of the crew and on the distribution of workload.

Post-Incident Analysis. An important contribution to safe operation is made by the careful analysis of operational incidents occurring at reactor sites in this country and abroad in order to derive lessons for the future. During such post-mortems, simulators are extensively used to model the behavior of the plant under the presumed emergency conditions. In the case of TMI, a mockup of the control room of the plant was constructed, and operators were "walked through" the incident. Such techniques have shed light on the part played by human factors in the occurrence and evolution of the incident. It is perhaps

obvious that possession of a model for the operating crew would enable one to perform much more extensive post-incident analyses that might help to answer the many "what if" questions that inevitably arise in such circumstances. In addition, the data collected from such incidents would serve as extremely valuable input for the refinement and elaboration of the model itself.

In summary, although it certainly cannot be a universal panacea, we believe that the application of the supervisory control model to the crew of a nuclear power plant would provide the NRC with data otherwise obtainable, if at all, only from an extensive series of simulations using actual crews. Where this is not possible, or where effects such as changes in operator training are to be modeled, we believe that the supervisory control model could be extremely valuable.

REFERENCES

1. S. Baron, "A Model for Human Control and Monitoring Based on Modern Control Theory," *Journal of Cybernetics and Information Sciences* (Spring, 1976).
2. D. L. Kleinman and R. E. Curry, "Some New Control Theoretic Models for Human Operator Display Monitoring," *IEEE Trans. Syst. Man & Cybern.*, SMC-7, No. 11 (November 1977).
3. W. H. Levison and R. B. Tanner, *A Control-Theory Model for Human Decision Making*, NASA CR-1953 (December 1971).
4. E. G. Gai and R. E. Curry, "A Model of the Human Observer in Failure Detection Tasks," *IEEE Trans. Syst. Man & Cybern.*, SMC-6, 85-94 (1976).
5. P. Wewerinke, "A Model of the Human Observer and Decision Maker," *Proc. of the Seventeenth Annual Conf. on Manual Control*, Pasadena, CA (1981).
6. R. Muralidharan and S. Baron, "DEMON: A Human Operator Model for Decision Making, Monitoring and Control," *Journal of Cybernetics and Information Science*, Vol. 3 (1980).
7. S. Baron, G. Zacharias, R. Muralidharan, and R. Lancraft, "PROCRU: A Model for Analyzing Flight Crew Procedures in Approach to Landing," *Proc. of Eighth IFAC World Congress*, Tokyo, Japan (August 1981).
8. G. L. Zacharias, S. Baron, and R. Muralidharan, *A Supevisory Control Model of the AAA Crew*, BBN Report No. 4802, Bolt Beranek and Newman, Inc., Cambridge, MA (October 1981).
9. S. Baron, C. Feehrer, R. Muralidharan, R. W. Pew, and P. Horwitz, *An Approach to Modelling Supervisory Control of a Nuclear Power Plant*, Report No. 4811, Bolt Beranek and Newman, Inc., Cambridge, MA (July 1982).
10. R. W. Pew, S. Baron, C. E. Feehrer, and D. C. Miller, *Critical Review and Analysis of Performance Models Applicable to Man-Machine Systems Evaluation*, BBN Report No. 3446, Bolt Beranek and Newman, Inc., Cambridge, MA (March 1977).
11. W. B. Rouse, *Systems Engineering Models of Human Machine Interactions*, North Holland, New York (1980).
12. Nuclear Regulatory Commission, *Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants*, NRC-Report WASH-1400 (NUREG-75/014) (October 1975).
13. A. D. Swain and H. E. Guttman, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*, Draft Report NUREG/CR-1278, Office of Nuclear Regulatory Research, US NRC, Washington, D. C. (April 1980).
14. A. I. Siegel and J. J. Wolf, *Man-Machine Simulation Models*, John Wiley and Sons, Inc., New York (1969).

15. D. B. Wortman, S. D. Duket, D. J. Seifert, R. L. Hann, and G. P. Chubb, *Simulation Using SAINT: A User Oriented Instruction Manual*, AMRL-TR-77-61, Wright Patterson Air Force Base, OH (July 1978).

16. R. J. Wherry, "The Human Operator Simulator-HOS," in *Monitoring Behavior and Supervisory Control*, T. B. Sheridan and G. Johannsen (Eds.), Plenum Press, New York and London (1976).

17. A. Newell, J. C. Shaw, and H. A. Simon, "Elements of a Theory of Human Problem Solving," *Psychological Review* **65**, 151-166 (1958).

18. R. B. Wesson, "Planning in the World of the Air Traffic Controller," *Proc. of the Fifth International Joint Conference on Artificial Intelligence*, MIT (August 1977).

19. T. B. Sheridan, "Understanding Human Error and Aiding Human Diagnostic Behavior in Nuclear Power Plants," Preprints of *Conf. on Human Failure Detection and Diagnosis of System Failures*, Roskilde, Denmark (1980).

20. R. A. Kisner, A. M. Fullerton, and P. R. Frey, "A Taxonomy of the Nuclear Plant Operator's Role," *Enlarged Holden Progress Meeting*, Friederichstaadt, Norway (June 14-19, 1981).

21. J. E. Potter and M. C. Suman, *Thresholdless Redundancy Management with Arrays of Skewed Instruments*, AGARDograph No. 224 (1977).

22. A. S. Willsky, "A Survey of Design Methods for Failure Detection in Dynamic Systems," *Automatica* **12** (1976).

23. E. Y. Chow, *Failure Detection System Design Methodology*, LIDS-TH-1055, MIT, Cambridge, MA (October 1980).

24. J. Rasmussen, "Models of Mental Strategies in Process Plant Diagnosis," Preprints of *Conf. on Human Detection and Diagnosis of System Failures*, Roskilde, Denmark (August 1980).

Paper 3

COGNITIVE MODELS AND COMPUTER AIDS FOR NUCLEAR PLANT CONTROL ROOM OPERATORS

Thomas B. Sheridan
Massachusetts Institute of Technology

INTRODUCTION

This paper reviews what is usually meant by a cognitive model of a control room operator in a nuclear power plant. It emphasizes the idea of internal (that is, mental) representation of external events and the use of such representation for the cognitive steps of attending, recognizing or learning, assessing and deciding.

The internal models of real control room operators of nuclear power plants surely include, in some form: piping and instrumentation diagrams; logic or fault tree diagrams; procedures; verbal rules of thumb (such as, "never let the pressurizer go solid"); images of the control board; visual or verbal stipulations of thresholds and limits of variables; diagrams (for example, the pressure-temperature saturation curve); a few abstract formulas (such as the perfect gas law); or recollections of scenarios experienced or imagined.

As external observers/analysts/modelers, we not only make assertions like those above about the operator's mental models but we also use more abstract or mathematical generalizations about such internal models, in effect, public models of the operators' private models. These take forms such as the Kalman filter observer or predictor, utility theory, the relative operating characteristic, etc. These ideas help us develop normative models — what is optimal or ideal under given specific constraints and performance/safety criteria. They serve as norms to evaluate descriptive models of "what is" (as best we can infer it) and to develop more coherence between procedures, control room design, control room supervisor practice, training, and licensing.

As computers play an increasingly important role in nuclear power plants, especially as cognitive aids to human supervisors of highly automated control systems, it is important that the software and computer interface characteristics be compatible with the operator's internal model. Specific examples discussed in this paper are in the monitoring and prediction of the plant state and in the detection and diagnosis of failures. Current trends in SPDS (safety parameter display system) and failure detection/location systems will be discussed in this regard.

WHAT IS A COGNITIVE MODEL?

A cognitive model is a formal representation, in mathematics, words or graphical symbols, of a person's mental events (attention, recognizing, learning, remembering, assessing and deciding) that is suitable for describing in a general way his or her observed mental behavior and for predicting future such behavior.

It is evident that the cognitive or mental aspects of behavior cannot be separated neatly from sensory and motor aspects, since independent unobtrusive measurements cannot be made between sensors and brain and between brain and muscles. Nevertheless, in tasks for which the required sensing and motor behavior is "easy" (for example, there is no difficulty in seeing or hearing the displays, or in operating the hand controls per se), we may attribute a person's response time, inaccuracies or errors and style of behavior to "cognition."

Fifty years ago psychologists avoided constructs like "cognition" or "mind." They sought to emulate the operational physics of the day. All that could be known was the stimulus and the response. "Behaviorism," as fashioned by Watson and Skinner, was commonly accepted. Introspection was taboo. It was meaningless to hypothesize variables intervening between stimulus and response because they could never be dealt with explicitly and in a vigorous experimental fashion. For all practical purposes they didn't exist.

Computers, along with the inherent human drive to explain what one is conscious of, have changed that. Computers make it much easier to represent complex attention, learning, memory, decision, pattern recognition and "artificial intelligence." One starts with a set of input and output data and some programmed structure or transfer relationships to map the inputs into the outputs. Then a computer can be set to trying systematically to adjust parameters in the model to "optimize" with respect to least squares deviations of the data from the model — or some other criterion. And presto, a cognitive model!

It's easy and fun — perhaps a bit too much so. The trouble is that a great variety of model structures may give "reasonable" fits to the data. Thus the modeler is obliged to invoke "Occam's razor" — Sir William of Occam's principle of least complexity in explanation. A simpler model which does the same job is better.

Few modelers purport to be explaining the mechanisms of mind, which would imply that real brains work just like the computer programs. They seek primarily to find a computerized structure or model which is:

1. capable of characterizing or describing a broad class of input-output data, data devised under a variety of conditions, and therefore a model which is robust in predicting outputs for new inputs,
2. simple to explain in words and symbols,
3. efficient in computer time and memory,
4. (sometimes) fits in with some broader explanatory theory or criterion of what is ideal.

Models in general are sometimes classed as "descriptive" or "normative," and this also holds for cognitive models. A descriptive model is based on the first three objectives above, with the emphasis on the first. A normative model adds emphasis on the last objective. It seeks to explain how things *should* work, what *should* happen if behavior is optimal with respect to the given theory or criterion. Examples of descriptive models are Skinnerian conditioning and the McRuer crossover model of manual control. Examples of normative models are Bayesian probability estimation and the Kleinman-Baron-Levison optimal manual control model.

It is natural for engineers and scientists to aspire to mathematical and computer models of the control room operator's cognitive behavior. However, we may have to be satisfied, for some time, with models that are at least partially verbal, that is, partially qualitative. A verbal model, of course, may be said to "fit" a broad class of events, but there is no vigorous test of goodness of fit.

A simplest quantitative model has mutually exclusive and collectively exhaustive categories of events according to which simple numerical counts and correlations are made. A well-organized table of data is such a model. Swain's human error models are of this type.

As a boolean logic or simple algebraic relations fit the data, we move upward in the hierarchy, and finally we get to complex logical structures and high-order differential or difference equations. The more parameters, however, the more arbitrariness, and the greater the universe of models which are equivalent in describing or predicting.

Finally, in a special category are those models which seek to fit not only the input and output data, but also claim to explain just what the process is by which actual input determines actual output — such that a probe could be put anywhere on the actual process and a corresponding variable could be found in the model. Such models are common for the physical plant. We are rather far from that in modeling cognitive behavior of human operators.

WHY COGNITIVE MODELS OF CONTROL ROOM OPERATORS OF NUCLEAR POWER PLANTS NOW?

Since the Three Mile Island accident formal concern for the human operator has greatly increased. At first there was a burst of compulsive enthusiasm for transferring all the military/aerospace human factors experience to the nuclear power plant control rooms. More recent is the realization that control of nuclear power plants is essentially different from the control of airplanes and fast-responding industrial and military systems, though it may be similar to large slow-responding ships.

Probably that area of man-machine system theory that has been most successful (with respect to both sophistication and usefulness) has been manual control theory. With this theory, a low-order linear differential equation model can provide an excellent prediction of operator response for simple, well-defined, man-in-the-loop continuous control tasks. Unfortunately there are few such simple and continuous manual control tasks in the nuclear plant control room. Steam generator water-level control probably comes closest, but even that amounts to somewhat intermittent coordination of at least two levels of control valves for each of several steam generators while observing and integrating information from displays of water inflow, steam outflow and level for each. Mostly in nuclear power plants the operator's control is a matter of following a procedure, more or less, for discrete operation of pump and valve controls while monitoring a great many variables of a highly nonlinear, slow-responding dynamic system with a number of automatic control loops built in and multiple control criteria.

It seems evident that the operator has little difficulty reading those displays he chooses to attend to and read. The question is which one he attends to and when. The same is true of the controls: his neuromuscular reaction time is not a critical factor. Thus the sensory

and motor characteristics of the operator are not critical. What seems to be critical is how the operator attends, learns or recognizes, remembers, and decides — those "internal" experimentally less accessible aspects of behavior we choose to call "mental" or "cognitive."

Experimental laboratory research on these components of cognition is hardly new. There has been some success in isolating them by carefully synthesizing the experimental conditions. However, to combine all these factors into a simple cognitive model and to apply such a model to a complex real-world situation is a bit heady — just now.

For present purposes we may assume, in conformance with the writings of a number of investigators in this area, that a more or less ordered series of cognitive steps occurs:

1. *attending* to stimuli (including deciding what to attend to and doing so),
2. *recognizing* familiar patterns, *learning* or acquiring new information and *remembering* it (three very closely related constructs),
3. *assessing* the environmental situation with respect to a remembered structure of knowledge or "internal mental model" of reality and the values (or criteria or objective functions) associated with it, plus *updating* of this "internal model,"
4. *deciding* what response to make, if any.

I make no claim that this series of steps constitutes a model of cognition, that the steps are mutually exclusive or uncoupled, or that they are collectively exhaustive, that is, that cognition includes nothing else. Surely, however, these are mostly what we mean by cognition.

If we could develop a sufficiently robust model of these activities with reference to nuclear power plant operation, we might then be more able to understand how operators make errors and how design of control room instrumentation, layout, and procedures contribute to such errors. Then we could design the hardware and software at the man-machine interface to reduce error and improve performance. We could train and license operators to ensure proper cognitive behavior as calibrated against such a model.

WHAT MUST BE INCLUDED IN THE CONTROL ROOM OPERATOR'S MENTAL MODEL OF REALITY?

We might return for a moment to the real world of the control room operator and ask, leaving aside the jargon of the cognitive modeler, what does the operator think about his task? What are his images of the reality surrounding him? Some of these surely are:

1. *hardware configuration* of the plant and control panel (from process and instrumentation diagrams, books, classroom instruction and direct observation of equipment),
2. *dynamic behavior* of the plant variables (from immediate or earlier simulator or actual control room experience in observing transient instrument behavior and inferring how plant variables are changing),
3. *nominal operating procedures* (from reading and using procedures documents, plus learned rules of thumb such as "never let the pressurizer go solid"),

4. *control criteria or objectives* (follow procedures, keep the plant running to generate megawatts, don't damage equipment, maintain coolant inventory and chemistry, maintain containment isolation, don't damage the core, be known as a good operator, etc.),

5. *abstract event relations* (PV = RT and the saturation line in PT space, $h = \int qAdt$, E = IZ, some fault trees, some event trees, etc.),

6. *administrative procedures and management policies* for assigned duties, shift turnovers, tagging, communications, storage and use of emergency equipment, coke machine and washrooms.

These are the mental images the trainers seek to implant. This is the information the NRC inspectors test for. Therefore this must be the substance of the operator's internal model, represented as shown to the left in Fig. 1. There must be another part of the cognitive process which *uses* the internal model to attend, recognize, learn, assess, update and decide how to respond, as shown to the right.

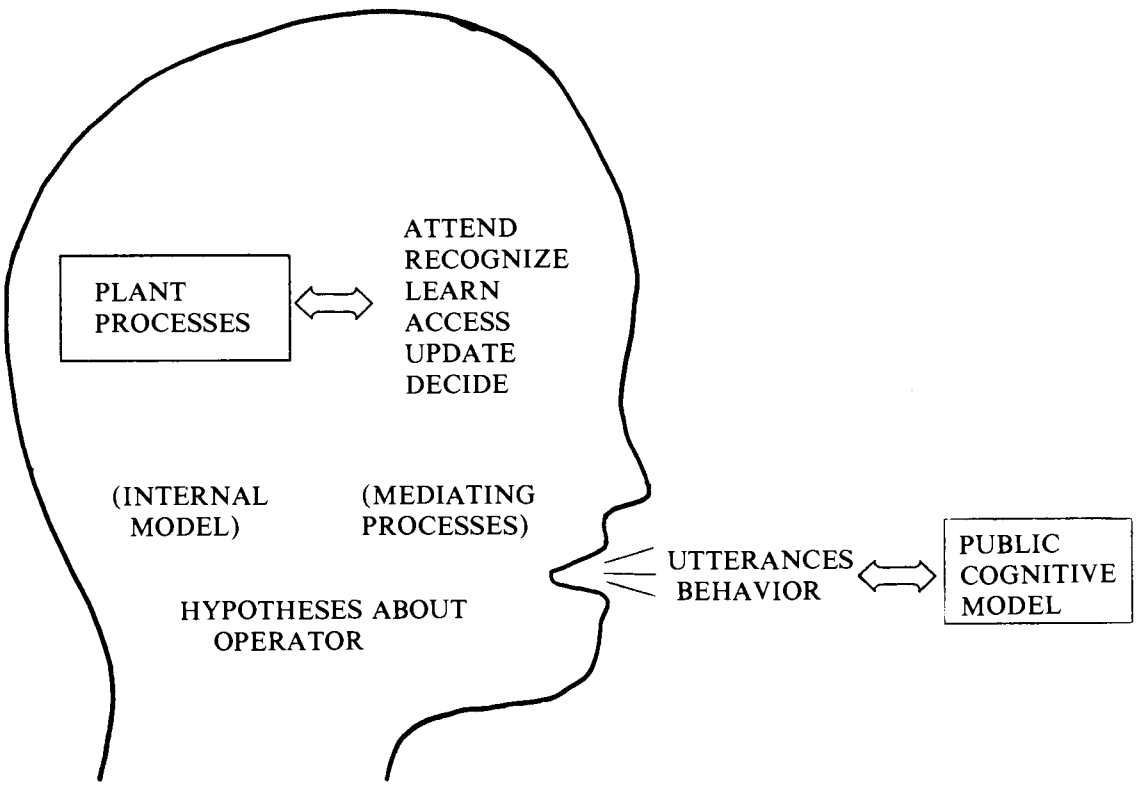


Fig. 1. Relationship of Operator's Internal Mental Model and Processes to Observer's Public Cognitive Model.

Operator attention behavior has been the subject of a number of modeling efforts. We know that people cannot consciously attend very well to more than one new thing at a time, though such a statement must be qualified by our knowledge that people can coordinate many unrelated "overlearned" activities. People can plan ahead several steps in a dynamic process, but can get themselves too committed in doing so and not be responsive to new unexpected opportunities.[1]

Differential recognition is good for only a few stimulus patterns that are coded by only one or two dimensions,[2] but as dimensions and redundancy become large (for example, as with faces and voices), the number of recognizable items becomes large.

Though the origins of the internal model go back to the Greeks, it really is the Kalman filter (or "observer") of control theory which has made explicit use of this idea. Here a dynamic computer model of a controlled plant is driven by the same input that drives the plant, and any discrepancy between model and actual plant outputs is used to bias the model input. Intermediate state variables as measured by the model are then used to perform optimal control. Man-machine modelers have had considerable success in adopting this idea to manual control systems.[3] Computer scientists and psychologists collaborating in the field of "knowledge representation" have devised knowledge structures from hierarchical trees and lists which have arrows to related items on other lists. In the nuclear plant, which has a slow, complex type of man-machine control task, the best use of the internal model still eludes us, though we take as an article of faith that some such internal model is an important ingredient.

As to human decision making, we know about various human biases and risk aversions that occur in risky choice behavior.[4] We also know that people try to confirm incorrect hypotheses and do not make best use of negative information.[5] We have fair evidence on many reasons why people make errors.[6,7]

THE MODELER'S ABSTRACT MODEL OF THE OPERATOR'S MENTAL MODEL

The discussion above has characterized a cognitive model as a formal representation of the activities of mind and suggested that this is useful for predicting human error and for designing hardware and software at the operator interface. The last section further suggests that the operator carries in his head a variety of mental images which are the subject matter of any cognitive model.

Now it is necessary to confront a problem of compound modeling:

1. Hypothetically the operator has internal to *his* head a model *of the plant*, its hardware and software and how it operates (Fig. 1). This internal representation plays a key role in each of the processes of attending, recognizing (or learning), assessing (and updating) and deciding.
2. The engineer and/or human factors analyst seeks to represent, within *his* overall cognitive model of the operator, the operator's internal model. This must be an explicit publicly understandable representation.

Thus we have a compounding — to model (in mathematics or diagrams or words) what we hypothesize is the operator's internal model of his external reality. In other words a model of a model: a public model of the operator's private model.

In order to achieve this feat, it would be helpful to have good hard empirical data regarding the operator's mental model. How to get such data? In what form? Earlier it was argued that the inputs and outputs from the operator's cognitive processes are not available directly and explicitly, that necessarily sensory functions prefilter the cognitive inputs and motor functions postfilter the cognitive outputs. Further, inputs and outputs do not occur in a neat continuous stream as in continuous control systems, so that conventional signal and control theory seems inappropriate except in a limited way for a very few specific tasks like steam generator water level control, synchronizing the turbine generator to the line voltage and bringing the reactor critical.

One approach to getting such data on cognitive function and the internal images or models of operators is the so-called "verbal protocol." This is a generic name for a class of techniques which in effect have the operator tell in words what he or she is thinking while performing operating tasks. One embodiment of this is the "talk-through," now being executed in many power plants as part of the NRC-mandated control room review. In effect, an operator steps through a written procedure, telling one or more observers what he is thinking and doing, while the observer notes for each step the objectives and criteria, the decisions made, the displays and controls used, and the consequences of error. The operator's utterances are the data most proximate to his mental model. The observer's entries and any subsequent analysis and graphing become the public model of the operator's private internal model of his external reality.

At Public Service of New Hampshire, some far-sighted computer engineers who are developing a computer-based alarm and graphic display system for the Seabrook plant have asked the operators to produce a set of abbreviated diagrams and alpha-numeric alarm-response checklists which for them constitute the most important information about the plant. The computer engineers will then make these diagrams computer-accessible and in part active (symbols and numbers change color, flash, etc.) in response to plant dynamics. These graphics, produced out of the operators' heads, constitute another representation of their mental models.

The modeler seeks to represent the operator's mental model in a form which is clear and public and unique, that is, in a form which is simple and efficient of graphic or computational resources and for which there is no ambiguity regarding what is meant. As noted earlier, the intention is robustness and predictability. Thus the form of the model is important since some forms are not amenable to characterizing certain attributes. The same is true of the occasional need to make the model fit into some broader theory.

Table 1 lists a variety of popular forms of model representation, with an indication of what theoretical attributes they embody. As often as not, in attempting to model a complex phenomenon like cognitive activity the modeler will resort to some combination of these modes, perhaps using a block diagram to show the logical order in which input-output transformations occur.

Table 1. Some Forms of Representation Useful for Cognitive Models

	Attributes								
	Metric Which Can Accommodate:								
Types of Modeling Approaches	Nominal	Ordinal	Interval	Continuity in Time	Probabilistic Events	Nonlinear Transforms	Feedback Loops	Value (Good-Bad)	Memory
Verbal descriptions	X							X	
Categorization		X						X	
Graphs and diagrams		X		X		X	X		X
Interpretive structural modeling		X						X	
Fault/logic trees		X							
Event/decision trees		X			X			X	
Markov graphs		X			X		X		X
Fuzzy set theory			X			X			
Classical and bayesian statistical inference			X	X	X				X
Information theory			X	X	X				
Queuing			X	X	X				X
Game theory			X	X	X		X	X	X
Estimation theory			X	X	X	X		X	X
Control theory			X	X	X	X	X	X	X
Multidimensional scaling			X					X	
Multiattribute utility			X		X	X		X	
Computer simulation and artificial intelligence (which can include any or all of the others)			X		X	X	X	X	X

Thus the cognitive modeler must infer from the operator's verbal utterances and overt stimulus-response behavior what are his or her mental images and associated processing of the type described in the previous section. Then the modeler must decide how best to characterize this by means of techniques such as those listed above. The degrees of freedom are many.

COMPUTER AIDS AND SUPERVISORY CONTROL: THE "EXTERNALIZED INTERNAL MODE"

Increasingly the computer is finding its way into nuclear plants for on-line process monitoring and logging, for generation of alarm displays and associated response procedures, for storage and display of reference graphics, for failure-detection and location, for post-accident analysis and for control. Mostly these functions are what would be called "cognitive" were the operator to do them.

By having the computer aid the operator by "off loading" many monitoring, logging, failure-detection and control tasks, the operator becomes a supervisor of multiple lower level intelligent systems.[8] His supervisory functions are to (1) *plan* use of the computer aids, (2) *teach* or initialize the various programs to be used, (3) *monitor* to ensure that

they are functioning properly once executed, (4) *intervene* to redirect the computer if it completes the assignment or to take over direct manual control in the event of failure and (5) *learn* from experience.

One of the most promising aspects of the computer aid for the nuclear power plant control room is the implementation within the computer of a model of all or parts of the plant. Such a model can be used for a variety of functions (ref. 8), including estimation, prediction, control, sensor validation, failure detection and location, and training and testing of the operator. An ideal is to run such a model in real time in synchrony with the plant, but this is not always practical or necessary.

My colleagues and I have implemented a system which includes a 47th order model of a fossil steam plant including all major systems modeled physically and including major feedback control loops. Based on discrepancies between key model variables (selected on the basis of state-variable power bonding) and measured variables of the actual plant, we can detect and locate failures of either sensors or components.[9] We have experimentally evaluated several alternative displays and output data formats for this system, believing that it is important that the interface for such a computerized aid be compatible with the operator's cognitive model.[10]

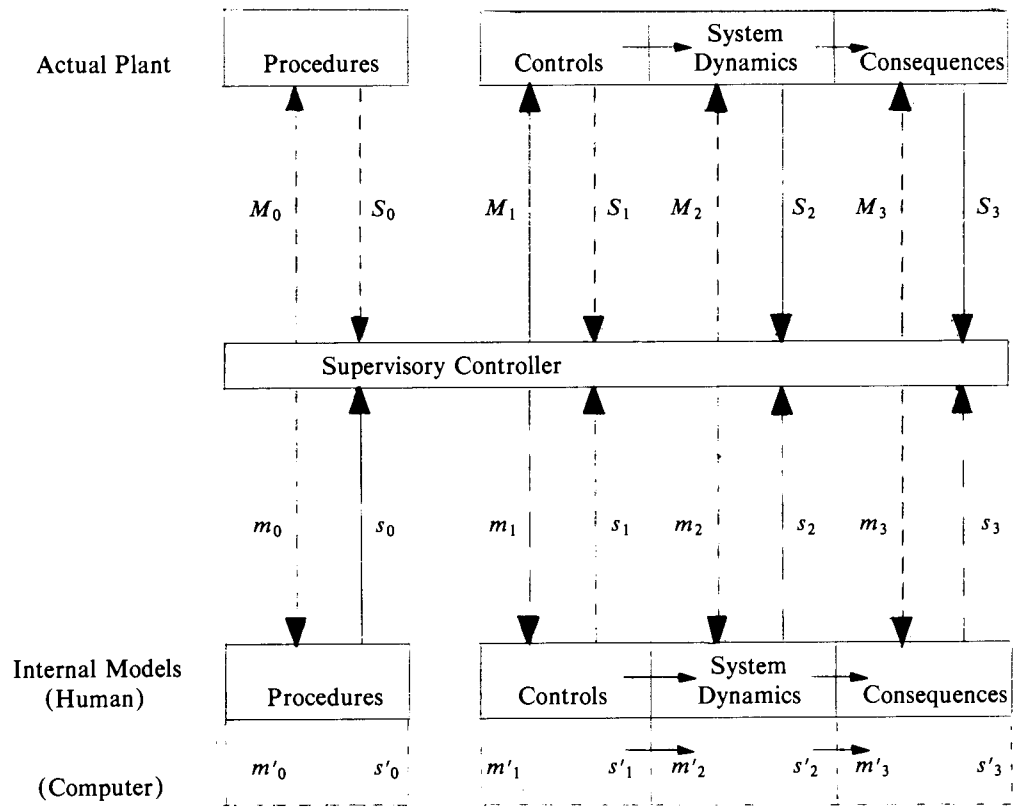
The SPDS (safety parameter display system) has been much talked about within the nuclear power industry. A few most critical plant variables are to be presented to the operator in a form from which he can make a quick appraisal of the overall health and safety of the plant, or generally where the plant is abnormal. Because it is on its way to becoming a requirement, many vendors have solutions to offer. Unfortunately we have as yet no solid human engineering experimental basis for what such a display should include or how it should be designed or used. In most embodiments the SPDS will be computer-based. As asserted above, it is most important that the SPDS be consistent with the operator's mental model of the plant.

One might generalize and assert that the operator's active cognitive function mediate between the actual plant (its procedures, its controls, its system dynamics, and the consequences of control actions and other inputs which reveal themselves in the displays) and a corresponding set of elements which constitute the internal mental model. This is shown in Fig. 2. If an "externalized cognitive model" is now computerized, it appropriately takes its place alongside the others, as shown at the bottom of the figure. Obviously it is most important that the two forms of model and the actual plant be consistent with each other in terms of images, coding, communication, etc. Discrepancies in consistency, if not caused by failure of one, are almost certain to produce human error at some time. This diagram can be extended to include additional operators who must communicate with each other, and additional computers, on or off line.

CONCLUSIONS

Summarizing the above arguments, and in conclusion, I make the following assertions:

1. Operating a nuclear power plant is neither a sensory nor a motor skill. It is a cognitive skill.
2. Engineering designers, procedure writers, operator trainers and regulators could benefit from having even simple cognitive models of control room operators that can be fit to simulator data and are sufficiently robust and predictive.



1. Procedures

M_0	Motion to consult a procedure (or to modify it).
S_0	What is read in procedure.
m_0	Effort to remember procedure (or to modify memory).
s_0	What is remembered in procedure.

2. Controls

M_1	Motion to activate a specific control.
S_1	What is observed as to which and how control is activated.
m_1	Making assumption of move to activate specific control.
s_1	Decision of what would be observed if m_1.

3. System State

M_2	Motion to observe system state display.
S_2	What is observed from system state display.
m_2	Alteration of $m_1 \rightarrow s_2$ model to conform to $M_1 \rightarrow S_2$ experience.
s_2	Decision of how system state would change if m_1.

4. Consequences

M_3	Motion to observe alarms, consult with other persons.
S_3	What is observed from alarms, other persons.
m_3	Alteration of $m_1 \rightarrow s_3$ model to conform to $M_1 \rightarrow S_3$ experience.
s_3	Decision of what consequences would change if m_1.

Fig. 2. A Framework for Modeling Human Diagnostic Behavior. M and m are manipulated variables for real and model plants, respectively; S and s are corresponding sensory variables; m' and s' relate to a computer-based model. (From ref. 7).

3. No satisfactory model of cognitive behavior of the nuclear power plant control room operator is available. Very limited models exist for specialized tasks.
4. It is likely that efforts will mount to develop useful cognitive models of the control room operator. These will start out being qualitative, special purpose and patchy. Fits to simulator data will occur in time. The idea of an "internal model" representing the operator's mental image of the environment has potential as a component of the cognitive model.
5. The computer will play an increasing role as a cognitive aid for monitoring, remembering and recall of data, failure detection and location, and analysis and control in both normal and emergency conditions. This can include an additional "internal" model of the plant.
6. It is of utmost importance that an operator's internal mental model of the plant and its behavior be consistent with that of other operators, with the externalized plant and process models embodied in process computers, and especially with the actual plant. Having available some validated cognitive models of the operator will assist in ensuring that this is true.

REFERENCES

1. M. K. Tulga and T. B. Sheridan, "Dynamic Decisions and Workload in Multi-Task Supervisory Control," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-10, No. 5 (1980).
2. G. A. Miller, "The Magical Number Seven Plus or Minus Two; Some Limits on Our Capacity for Processing Information," *Psychology Review* **63**, 81-97 (1954).
3. D. D. Kleinman, S. Baron, and W. N. Levison, "A Control Theoretic Approach to Manual Vehicle Systems Analysis," *IEEE Transactions on Automatic Control*, AC-16, No. 6, 824-832 (1971).
4. A. Tversky and D. Kahneman, "Casual Schemas in Judgments Under Uncertainty," in *Progress in Social Psychology*, M. Fishbein (Ed.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1980).
5. W. B. Rouse, *Systems Engineering Models of Human-Machine Interactions*, North Holland (1980).
6. J. Rasmussen and W. B. Rouse, Editors, *Human Detection and Diagnosis of System Failures*, Plenum Press, NY (1981).
7. T. B. Sheridan, "Understanding Human Error and Aiding Human Diagnostic Behavior in Nuclear Power Plants," in *Human Detection and Diagnosis of System Failures*, Plenum Press, NY (1981).
8. T. B. Sheridan, *Supervisory Control: Problems, Theory and Experiment for Application to Human-Computer Interaction in Undersea Remote Systems*, MIT Man-Machine Systems Laboratory Report (1982).
9. U. Tsach, T. B. Sheridan, and J. Tzelgov, "A New Method for Failure Detection and Location in Complex Dynamic Systems," *Proc. 1982 American Control Conference*, Arlington, VA.
10. J. Tzelgov, U. Tsach, and T. B. Sheridan, "Effects of Indicating Failure Odds and Smoothed Outputs on Human Failure Detection in Dynamic Systems," *Human Factors* (submitted paper).

Paper 4

VISUAL MOMENTUM: AN EXAMPLE OF COGNITIVE MODELS APPLIED TO INTERFACE DESIGN

David D. Woods
Westinghouse Research and Development Center
Pittsburgh, Pennsylvania

1. INTRODUCTION

The growth of computer applications has radically changed the nature of the man-machine interface. Through increased automation, the nature of the human's task has shifted from an emphasis on perceptual-motor skills to an emphasis on cognitive activities (e.g., problem solving and decision making). The result is a need to improve the "cognitive coupling"[1] of person and machine. (See Norman,[2] Hollnagel and Woods[3] for discussion of the implications of technological developments for human-system interaction).

The goal of this paper is to describe how knowledge from cognitive psychology can be used to provide guidance to display system designers and to solve human performance problems in person-machine systems. The mechanism is to explore one example of a principle of man-machine interaction — visual momentum — that was developed[4] on the basis of a general model of human "front-end" cognitive processing.[a]

2. WHAT GUIDES OUR GLANCES

Parallel (that is, simultaneous) presentation of data is often claimed to be superior to serial presentation. Serial modes of data display (for example, CRT displays) provide only a narrow "keyhole" to view the total data base. This degrades user information extraction because "the human is used to having his total information system displayed and being able to sample and timeshare from his system *by a movement of the eyes and his interpretive skills*" (Pope,[5] words italicized by author for emphasis in current discussion).

However, the narrow field of view (2°) of the high-resolution portion of the retina (the fovea) constrains the amount of data a viewer can acquire in a single glance. This is no limitation under naturalistic viewing conditions because there are psychological mechanisms which convert a serial input (i.e., a succession of eye fixations) into what we commonly think of and experience as parallel data acquisition. In other words, the advantage attributed to parallel over serial data presentation is a function of human perceptual and attentional mechanisms rather than the mode of data presentation. This means user data-acquisition skills can be improved if knowledge from cognitive psychology about the above characteristics of perception and attention is applied to display system design.

Studies of scanning patterns (for example, those by Mackworth and Morandi[6]) show that people tend to look at (fixate on) "informative" areas in a visual scene or picture. But how does one identify informative data without first processing all of the available data?

[a] "Front-end" cognitive processing emphasizes processes of information acquisition as opposed to response selection and execution.

Furthermore, which data are informative will depend on the context, for example, on the viewer's task or expectations.[7] It is informativeness defined as some relation between viewer and the scene (such as importance, expectation, predictability), rather than as only a property of the scene, that is the important determinant of visual scanning patterns.

There are two mechanisms which support the ability to find and fixate on informative areas: (1) an analysis of the global properties of the stimulus,[8] and (2) a concept-driven or top-down analysis, which begins with information the viewer already possesses.

The global analysis process organizes the visual field into the psychological units (that is, figures and ground) that attentional processes operate on. Global analysis is based on figural cues (such as shape, color and location) that help define relationships across the stimulus field; in other words, it determines which stimulus elements belong together and which belong apart.[9]

A concept-driven analysis will help to determine quickly the gist of a scene. This form of analysis is based on information or knowledge that the system already possesses in the form of schemata or internal models.[b] Concept-driven analysis controls viewer data-sampling behavior through cognitive and semantic factors; for example, an object incongruent with the active schema is informative.[7]

A pattern of eye fixations is like an exploration route of the visual scene. Like any explorer, the process will be much more organized and efficient if a map or representation of the area being explored is available for guidance. Concept-driven analysis utilizes a map or schema of the available and *potentially* available information in the stimulus world to guide and control data-gathering activities.

One corollary of concept-driven behavior is that the word "observation" is an empty term unless accompanied by a notation of what guides and/or activates the observation process. Observation *looks for* something, i.e., is guided by knowledge, and, therefore, is informational only when the guiding factor is included. Even browsing activities are guided — for example, I might look for a book I would enjoy to read in the science fiction section of a library. This example also points out that the factors guiding the observation process can range from the very specific (I am looking for one book in particular) to the general (any science fiction book or even any "interesting" book). This is very often overlooked in interface design where the designer assumes, usually implicitly, that the user searches with a very specific question in mind.

The above conceptualization of visual information processing as a perceptual cycle (Fig. 1) represents a model of cognitive processing, not as a fixed series of linear stages, but as a recursive set of operations including both bottom-up/data-driven activities (that is, analyses arising from information which comes from the environment) and top-down/concept-driven activities (that is, analyses which start from information which the viewer already possesses) (c.f., Norman and Bobrow[12] and Kubovy and Pomerantz[9]). This view emphasizes that perception is an *active, inherently selective* process of data gathering, that is, *part of the cognitive processing mechanisms*, rather than a process of passive reception and transmission prior to cognitive processing.

---

[b]A schema is an integrated representation of the relationships in a scene. But schemata are not internal images; they also function as plans for obtaining more information about objects and events (c.f., Hochberg[10] and Neisser[11]).

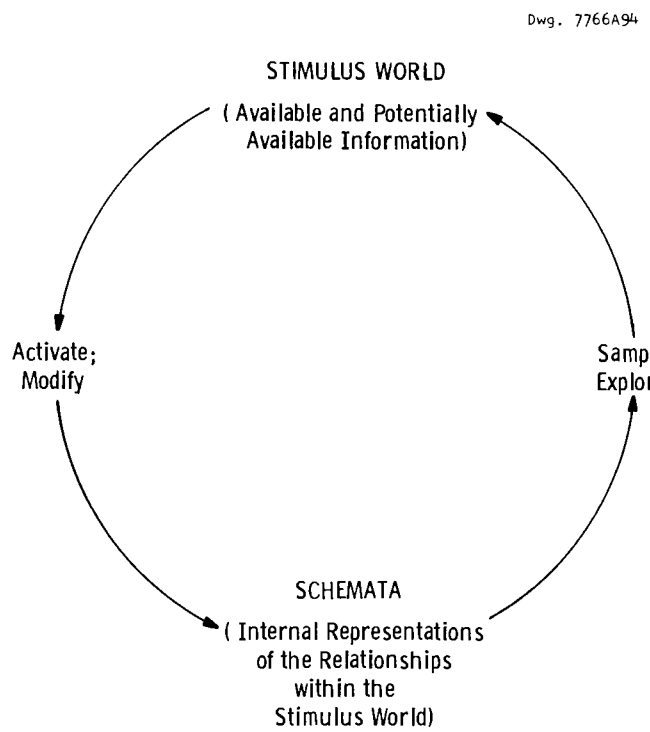


Fig. 1. The Perceptual Cycle. (Adapted from Neisser.[11])

3. VISUAL MOMENTUM

The concept of visual momentum describes the level of compatibility between the characteristics of a computer display system and the characteristics of perceptual processing and selective attention that direct our glances and help to integrate data across successive glances. Visual momentum is a measure of the distribution of attention. When visual momentum is high, the viewer's attentional mechanisms can identify highly informative areas because the display system structure provides a visual frame of reference that describes the relationships among data points as well as the data points themselves. When visual momentum is absent, there are no perceptual cues available to support the global and concept-driven processing activities for guiding information acquisition. Instead the user must rely on other, lower capacity mental processes, such as memory, to generate where to look and to determine how the data acquired from one "glance" into the data base relates to data examined in the past and to data that might be examined in the future.

When high capacity perceptual mechanisms are bypassed due to interface characteristics, there is a mismatch between the system's image of the user's cognitive skills and his actual skills. Mismatches in the man-machine cognitive system increase the user's mental workload because demands are increased on lower capacity processing systems (for example, memory) and because additional, irrelevant mental tasks are imposed on the user.[13]

One example of the effects of low visual momentum is memory bottlenecks such as the "getting lost" phenomena reported in large multiple display networks.[14] Getting lost in a

display network means that the user does not have a clear conception of relationships within the system, does not know his present location in the system relative to the display structure, and finds it difficult to decide where to look next within the system. The result is inefficient and incomplete utilization of display system data resources.

The breakdown in the viewer's attentional processes represented by low visual momentum can also be seen in what has been described as "cognitive tunnel vision."[15] Cognitive tunnel vision occurs when the user's attention is locked onto a subset of variables to the exclusion of others. This decrease in the size of the field of attention (or deviation from optimum sampling patterns) can lead to monitoring failures, especially when the unexpected occurs or when correct state identification is a function of integrating data from several sources.

In general, then, consequences of low visual momentum can include:[c]

- Cognitive tunnel vision, that is, decreases in the size of the viewer's field of attention or functional field of view.
- Impaired ability to locate "important" data.[16,7]
- Getting lost in display networks; that is, finding it difficult to decide where to look next.[14]
- Memory bottlenecks due to increases in mental workload.[13,17]
- Decreases in problem-solving performance.[18]

Low visual momentum is the equivalent of serial data presentation. High visual momentum, on the other hand, represents what is meant by parallel data presentation — not simultaneous data presentation, but support for the user to sample data based on "a movement of the eyes and his interpretive skills."[5]

4. CONVERTING SERIAL DATA PRESENTATION TO PARALLEL DATA PRESENTATION

What techniques can be used to eliminate the problems associated with serial data presentation? Figure 2 represents visual momentum as a psychological dimension of display systems that describes the level of compatibility between display system characteristics and the psychological mechanisms that support selective attention.

Based on studies of how people integrate data across successive views, there are a series of techniques available to increase the visual momentum in a display system.[19,20] The key element in all of these concepts is to provide the viewer with visual data about the location of one view with respect to another or, more generally, with data about the relationships across display frame. The goal is to help the user construct and maintain a cognitive map or schema of the data structure. It is this internal model that results in the simultaneous representation of information.

---

[c]See ref. 4 for more extensive discussion of consequences.

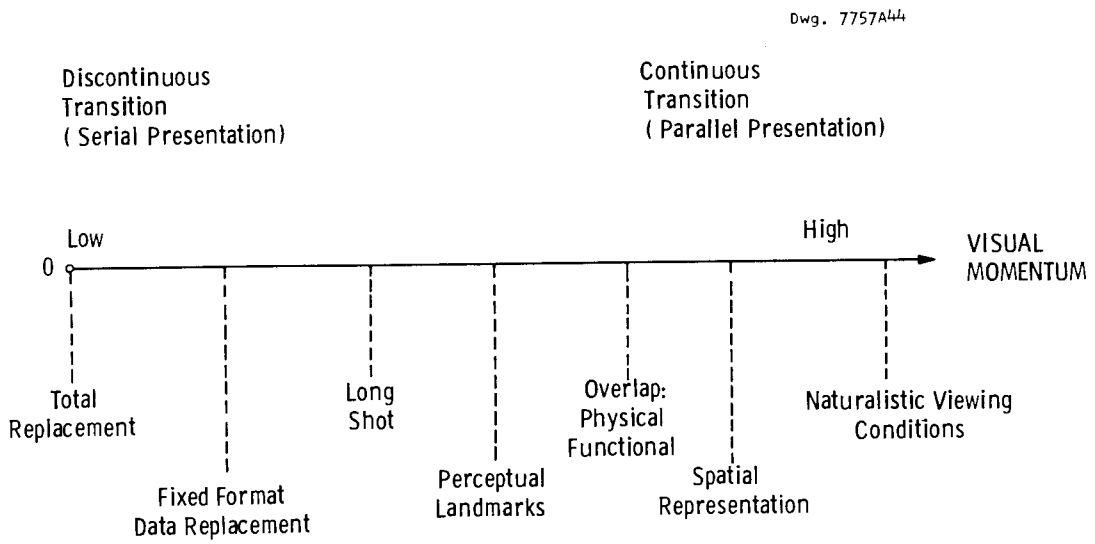


Fig. 2. The Dimension of Visual Momentum.

A long shot or establishing view provides an overview of the available data as well as summary status data. It establishes the relationships among data and acts to funnel the viewer's attention to the "important" details. Essentially, a long shot is a picture of a schema, that is, a map of the relationships between individual display elements. As a result of the pictorial representation of the schema, the viewer does not have to remember or construct a model of the data structure in his head. For this to be an effective technique, the display system structure must explicitly incorporate the set of relationships that are important to the user's tasks to be portrayed in the long shot. Merely summarizing data is insufficient for effective across-display information extraction.

Another technique to join together successive views is to provide landmarks.[21] Clear landmarks help the viewer integrate data by providing an easily discernible feature in order to establish a relative frame of reference for inter-data element relationships. Landmarks are "features that are visible at a distance and that provide information about location and orientation" (ref. 21, p. 23), that is, features recognized through the global analysis mechanism. When some feature or object is immediately recognizable in a scene, schemas can be quickly activated to guide subsequent looking behavior. "Once an object is identified in a scene, we may quickly know the kind of company it keeps" (ref. 16, p. 239).

When the subject is computer display systems, there is another type of glue to enhance comprehension across display transitions. In one sense, a display system is one large display of the entire data base that the user examines in discrete chunks or frames. From this point of view, the design of individual displays is the process of cutting the large data representation into pieces that will fit onto the available display surface. The viewer is helped to integrate the individual pieces back into the complete representation through a simple technique that overlaps the pieces. Physically overlapping displays is a standard cartographic technique to increase viewer comprehension. Just as in maps, the overlap sections

should not be presented at the same level of detail as the main portion of the display frame. Only those features needed to establish across-display relationships and to call the viewer's attention to other data and display frames should be incorporated as display overlap. In other words, display overlap can be used to support the global analysis mechanism in human perceptual function, although analysis of the viewer's decisions and tasks is required in order to effectively choose the size and content of the overlap area.

Functional display overlap can be used in addition to physically overlapping display frames. Functional overlap is a technique to present pictorially the functional relationships that cut across display frame boundaries. For example, displays of functional rather than physical form have long been used to portray electronic circuits.[22] In process control applications each display frame often presents data on the status of one system. Functional overlap occurs when each display contains not only data on the status of the system in question but also data on other systems that affect system operation and data about the functions or goals that the system is meant to support. If a system is designed to transport material to maintain inventory in some reservoir, the display should show data about system operation (for example, is there flow?) and about goal achievement (is inventory increasing?). However, functional overlap can be implemented only if the functional relationships among the data points are specified; that is, how the data relate to the user's task must be known. By identifying and incorporating relationships between data and user tasks into the display system structure, the user can more easily locate important and informative data.

The long shot, landmark and overlap techniques all increase visual momentum by providing data about the location of one view with respect to another — for example, they show the physical or functional relationships between successive views. These techniques help to establish a spatial representation or frame of reference that describes the relationship between individual displays.

Spatial organization of data (spatial coding) is a potent aid to human cognitive processing.[23,24] The priority of space as an organizing principle is so compelling that nonspatial data is often given a spatial representation to improve user comprehension. Taxonomic trees in biology or computer program flow charts are examples. Spatial organization translates the normative user internal model into a perceptual map. The user *sees*, rather than remembers, the organization of data in the system and *can move within the system just as he moves in an actual spatial layout*. Spatial representation is important because it helps make the process of finding data a high capacity perceptual function rather than a limited capacity thinking function. In other words, it reduces the mental load imposed on the user, freeing limited mental resources to deal with the underlying process or system rather than with the characteristics of the display mechanism.

The goal of the visual momentum concept is to produce display systems that support the perceptual and attentional mechanisms that make viewing real world scenes so effective. These mechanisms include the global analysis process and concept-driven analyses. The various techniques to increase visual momentum are means to put into display systems the special cues that these mechanisms utilize.

5. DATA SAMPLING BEHAVIOR

Visual momentum provides a mechanism to improve operator data sampling behavior. What is the relationship of visual momentum to other models of sampling behavior? Based on the Sampling Theorem of Shannon, an optimum frequency of data sampling can be determined.[25] Optimal sampling models have been expanded to include the effects of payoff structure,[26] the proximity to error boundaries,[27] and the data-acquisition time for each sample.[28]

Moray expanded the view of optimal sampling by considering the characteristics of the data that describe complex systems such as a nuclear power plant (ref. 15, p. 187):

- There are multiple, dynamic data sources.
- The data elements are richly interconnected (and therefore vary in degree of correlation).
- The data elements are diagnostically ambiguous (in general, the state of one variable is insufficient to characterize the state of the system).

Moray uses these observations (particularly the second) to note that operator sampling should vary as a function of the correlation between data points and the results of past samples (is system state normal or abnormal?). In other words, sampling behavior is also a function of the relationship between user tasks and the state of the data sampled.

The concept of visual momentum carries Moray's analysis further by noting two other characteristics of the available data in complex systems:

- Data elements can be indirect measures of the to-be-controlled quantity (for example, is pressurizer level always a direct measure of inventory?).
- The relevant interconnections between data points can vary with system state and operator task. (For example, in nuclear power plant normal operations, when the primary task is to control boron concentration, inventory control can be a side effect; when the primary task is inventory control, boron concentration control becomes the side effect).

This means optimum data sampling is a function of the user's ability to distinguish relevant from irrelevant data. As discussed earlier, what is relevant depends on the relationship between user expectations/tasks and the data, that is, concept-driven behavior. Visual momentum captures knowledge about the cognitive activities that support the identification of relevant data in human perception so that the display system design will support an effective distribution of attention.

6. DISCUSSION

Visual momentum is one example of principles of human-machine interaction.[29] These principles[d] are a mechanism to translate knowledge (data/models) from cognitive psychology into a form useful to designers (see Fig. 3). They describe how characteristics of the

[d]These concepts are called principles because they are not guidelines that can be rotely or directly applied; rather they are meta-guidelines which the designer can use to derive the specific guidelines to match the specific application.

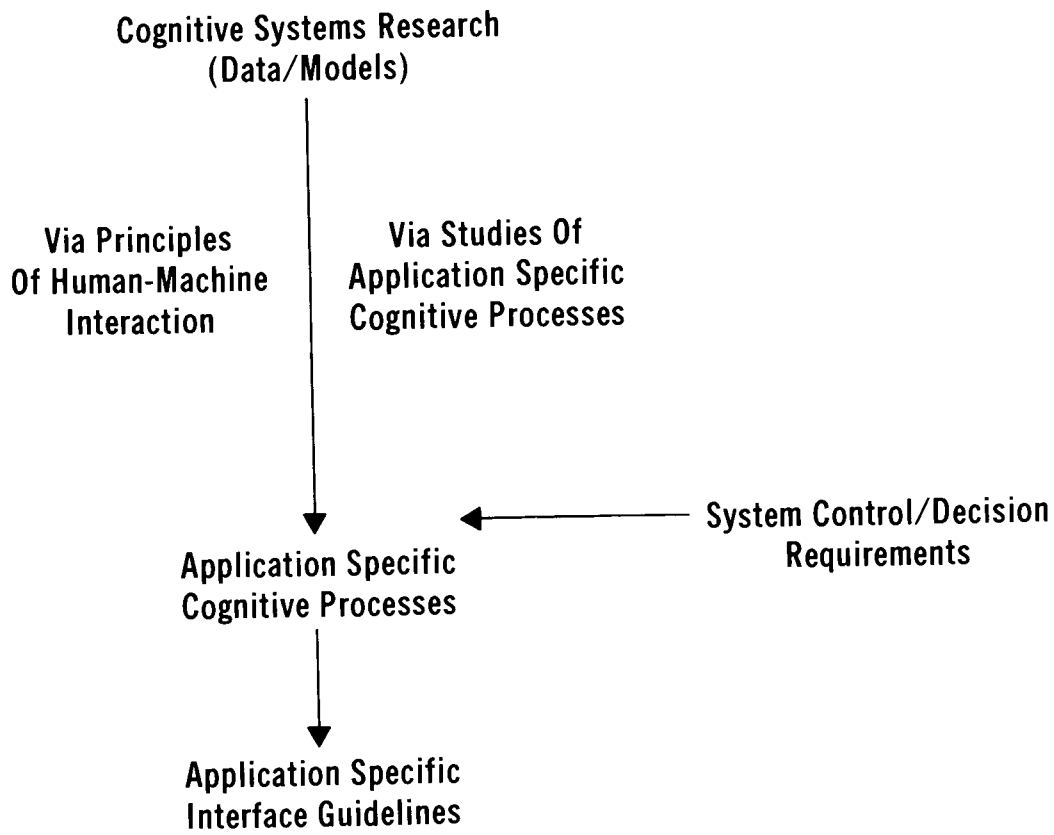


Fig. 3. Applying Knowledge from Cognitive Psychology.

interface affect or interact with the user's cognitive skills. For example, visual momentum relates several human performance problems to a single underlying mechanism and identifies techniques that can be used to help solve them. The result is that the process of developing improved forms of human-machine interaction becomes principle-driven. Knowledge from cognitive psychology is used to identify places and means that are likely to improve person-machine performance; experience becomes an evaluation/confirmation tool.

The concept of visual momentum broadens the scope of the display design process. Traditionally, display design attempts to ensure that human sensory limits are not strained. However, if the design process stops at this point, there is an implicit assumption that if the viewer can potentially see/read the data, then he will and should find all of the "right" data at the "right" time. Phenomena like the keyhole, getting lost, and cognitive tunnel vision effects, as well as events like Three Mile Island, show that this assumption is unwarranted.

Visual momentum points out that questions of information acquisition must also be considered in display system design. To accomplish this, the designer must understand the set of potential relationships among data points and how those relationships map onto user tasks (Fig. 3, system control/decision requirements). Visual momentum asserts that building these sets of relationships into the display system through spatial representations will contribute to improved cognitive performance.

REFERENCES

1. M. J. Fitter and M. E. Sime, "Responsibility and Shared Decision-Making," in *Human Interaction with Computers,* H. T. Smith and T. R. G. Green (Eds.), Academic Press, London (1980).
2. D. A. Norman, *Steps Towards a Cognitive Engineering: System Images, System Friendliness, Mental Models,* Department of Psychology and Program in Cognitive Science, University of California, San Diego (1981).
3. E. Hollnagel and D. Woods, "Cognitive Systems Engineering: New Wine in New Bottles," *International Journal of Man-Machine Studies,* in press.
4. D. D. Woods, "Visual Momentum: A Concept to Improve the Cognitive Coupling of Human and Machine," manuscript submitted for publication (1982).
5. R. H. Pope, "Power Station Control Room and Desk Design: Alarm System and Experience in the Use of CRT Displays," *International Symposium on Nuclear Power Plant Control and Instrumentation,* Cannes, France (1978).
6. N. H. Mackworth and A. J. Morandi, "The Gaze Selects Information Details within Pictures," *Perception and Psychophysics* **2,** 547-552 (1967).
7. G. R. Loftus and N. H. Mackworth, "Cognitive Determinants of Fixation Location During Picture Viewing," *Journal of Experimental Psychology: Human Perception and Performance* **4,** 565-572 (1978).
8. D. Marr, "Early Processing of Visual Information," *Philosophical Transactions of the Royal Society of London* **275,** 483-519 (1976).
9. M. Kubovy and J. R. Pomerantz, *Perceptual Organization,* Lawrence Erlbaum Associates, Hillsdale, New Jersey (1981).
10. J. Hochberg, *Perception* (2nd ed.), Prentice-Hall, Englewood, New Jersey (1978).
11. U. Neisser, *Cognition and Reality,* W. H. Freeman, San Francisco (1976).
12. D. A. Norman and D. G. Bobrow, "On the Role of Active Memory Processes in Perception and Cognition," in *The Structure of Human Memory,* C. N. Cofer (Ed.), W. H. Freeman, San Francisco (1976).
13. J. Rasmussen, "The Human as a System Component," in *Human Interactions with Computers,* T. H. Smith and T. R. G. Green (Eds.), Academic Press, London (1980).
14. G. Robertson, D. McCracken and A. Newell, "The ZOG Approach to Man-Machine Communication," *International Journal of Man-Machine Studies* **14,** 461-488 (1981).
15. N. Moray, "The Role of Attention in the Detection of Errors and the Diagnosis of Failures in Man-Machine Systems," in *Human Detection and Diagnosis of System Failure,* J. Rasmussen and W. B. Rouse (Eds.), Plenum Press, New York (1981).
16. I. Biederman, "On the Semantics of a Glance at a Scene," in *Perceptual Organization,* M. Kubovy and J. R. Pomerantz (Eds.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1981).
17. J. Rasmussen and M. Lind, *Coping with Complexity,* Riso-M-2293, Electronics Department, Riso National Laboratory, Roskilde, Denmark (1981).
18. J. B. Brooke and K. D. Duncan, "Effects of System Display Format on Performance in a Fault Location Task," *Ergonomics* **24,** 175-189 (1981).
19. J. Hochberg and V. Brooks, "Film Cutting and Visual Momentum," in *Eye Movements and the Higher Psychological Functions,* J. W. Senders, D. F. Fisher and R. A. Monty (Eds.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1978).
20. J. Hochberg and V. Brooks, "Psychological Aspects of the Cinema," in *Handbook of Perception, Vol. X,* E. C. Carterette and M. P. Friedman (Eds.), Academic Press, New York (1978).
21. J. Hochberg and L. Gellman, "The Effect of Landmark Features on Mental Rotation Times," *Memory and Cognition* **5,** 23-26 (1977).

22. L. H. Bainbridge-Bell, "The Standardization of Symbols and the Arrangement of Electronic Circuit Diagrams," *Journal of the British Institute of Radio Engineers* **13,** 339-53 (1953).

23. G. A. Miller, "Psychology and Information," *American Documentation* **19,** 286-289 (1968).

24. J. L. Bennett, "Spatial Concepts as an Organizing Principle for Interactive Bibliographic Search," pp. 67-82 in *Interactive Bibliographic Search: The User/Computer Interface,* D. E. Walker (Ed.), AFIPS Press, Montvale, New Jersey (1971).

25. J. Senders, J. Elkid, M. Grignetti and R. Smallwood, *An Investigation of the Visual Sampling Behavior of Human Observers,* NASA CR-434 (1966).

26. T. Sheridan, "On How Often the Supervisor Should Sample," *IEEE Trans.,* SSC-6, 140-145 (1970).

27. J. Carbonell, "A Queueing Model for Many-Instrument Visual Sampling," *IEEE Trans.,* HFE-7, 157-164 (1966).

28. N. Moray, *Human Information Processing and Supervisory Control,* Man-Machine Systems Laboratory Report, M.I.T. (1980).

29. J. M. Carroll and J. C. Thomas, "Metaphor and the Cognitive Representation of Computing Systems," *IEEE Trans. Syst., Man, & Cybern.,* SMC-12, 107-116 (1982).

Paper 5

USING COGNITIVE MODELING TO IMPROVE THE MAN-MACHINE INTERFACE

Roger A. Newton and Richard C. Zyduck
Wisconsin Electric Power Company

Dennis R. Johnson
Quadrex Corporation

1. INTRODUCTION

The 1979 accident at Three Mile Island has caused us to look at many items associated with nuclear power plants in a new light. In particular, the evaluation of the accident has shown that it would have been helpful to the operators if they could have more easily determined the safety state of their plant. In order for this to be possible, however, the interface between the man (nuclear plant operator) and the machine (control board instrumentation) would have to be improved.

A group of utilities from the Westinghouse Owners Group was formed in early 1980 to examine the interface requirements and to determine how they could be implemented. The products available from the major vendors were examined early in 1980 and judged not to be completely applicable. The utility group then decided to develop its own specifications for a Safety Assessment System (SAS) and, later in 1980, contracted with a company to develop the system, prepare the software and demonstrate the system on a simulator.

The resulting SAS is a state-of-the-art system targeted for implementation on pressurized water reactor nuclear units. It has been designed to provide control room operators with centralized and easily understandable information from a computer-based data and display system.

This paper gives an overview of the SAS plus a detailed description of one of its functional areas—called AIDS. The AIDS portion of SAS is an advanced concept which uses cognitive modeling of the operator as the basis for its design.

2. OVERVIEW DESCRIPTION OF SAFETY ASSESSMENT SYSTEM (SAS)

The major features of the SAS include several functional areas: (1) top-level displays of key parameters used to assess the safety status of the plant; (2) trend graphs of groups of related parameters; (3) the accident identification and display system (AIDS); (4) a message area; (5) a safety system readiness monitor; (6) a safety system performance monitor; and (7) a critical safety function monitor. One of the design criteria for the SAS was to address requirements for a Safety Parameter Display System (SPDS) as described in NUREG-0696, "Functional Criteria for Emergency Response Facilities," which describes the facilities and systems to be used by nuclear power plant licensees to improve responses to emergency situations.

The purpose of the SPDS is to assist control room personnel in evaluating the safety status of the plant. Functionally, the requirements are:

- Provide a continuous indication of plant parameters or derived variables representative of the safety status of the plant.
- Aid the operator in the rapid detection of abnormal operation conditions.
- Concentrate, in one location, a minimum set of parameters to allow timely status assessment without surveying the entire control room.
- Incorporate human factors considerations for simplicity and pattern recognition.
- Identify faulty data.
- Display information during steady-state and transient conditions.

The minimum set of SPDS parameters from which the safety status of the plant is assessed was not specified by NUREG-0696. However, it did define important plant functions which should be covered by the SPDS parameters. These functions are reactivity control, reactor core cooling and heat removal from the primary system, reactor coolant system integrity, radioactivity, and containment integrity.

The design of the SAS fulfills the functional requirements of the SPDS by providing the control room operator a means of rapidly detecting abnormal operating conditions and evaluating the safety status of the plant. In addition, the SAS demonstrates concepts that go beyond SPDS requirements — in particular, the advanced diagnostic techniques used in AIDS.

3. AIDS: SAS ACCIDENT IDENTIFICATION AND DISPLAY SYSTEM

The purpose of the SAS functional area AIDS is to identify the "most likely" accident(s) to be occurring based on the behavior of current, measurable and operating parameters. This is a concept completely opposite to the one which models the plant and tries to determine the cause of abnormal parameter response by making changes to the model, thus identifying the cause. Instead, the response of the plant parameters are examined and analyzed, in a manner similar to how an operator performs this evaluation, to determine the "most likely" accident(s). In order to accomplish this purpose, one needs to construct a function which is composed of weighted response functions of independent variables. The methodology is divided into two distinct areas: (1) the development of a weighting factor or "importance" of a parameter to an accident type, and (2) the development of an algorithm that models the perceived response to a parameter.

3.1. AIDS Functional Description

The AIDS module of the SAS calculates a weighted indicator for each of four major accidents: loss of coolant accident (LOCA), steam generator tube rupture (SGTR), loss of secondary coolant (LOSC), and inadequate core cooling (ICC). These weighted indicators are then displayed to the operator as a bar height for each major accident. A target above each AIDS bar indicates the status of the SAPs (see below) for that event.

The inputs to AIDS are parameters derived (in most cases) from redundant plant sensor inputs. These parameters and their rates of change are used to indicate accident types by a symptom-oriented algorithm based on previously analyzed simulator data or calculational experiments. For an SGTR or a LOSC, the logic includes information to allow identification of the affected steam generator. For a LOSC, the break location (inside or outside of containment) can be determined. The AIDS parameters are as follows:

- Primary AIDS Parameters (PAP)—Parameters which are positive indicators of an accident and provide unambiguous information that an accident has occurred and, in combination with other key parameters, will differentiate between the four types of accidents (refer to Table 1).
- Secondary AIDS Parameters (SAPs)—Parameters which may anticipate a possible condition, add confirmation to a key parameter's behavior, have low quality, or be a casual condition for a particular accident. In general, this information needs to be provided to the operator for evaluation, but may not give (by itself) a clear indication of what is happening (refer to Table 2).
- Modifier—A conditional event that alters or qualifies the use of a PAP or SAP before the parameter is valid as an accident indicator.

3.2. AIDS Displays

There are four supporting displays for AIDS. They are labeled, on the dedicated function-keypad, as: LOCA, SGTR, LOSC, and ICC. The top half of each display lists the PAP, SAP, and modifier status and their current values or conditions and trends, as appropriate. The current bar unit values for the PAP are also displayed.

The bottom half of each display shows the trend of the appropriate bar height for the last 30 minutes. There is a trend available for both the NORMAL and HEATUP/COOLDOWN AIDS algorithms. Which one is displayed is determined by which top-level display was selected last. Any one of the 14 trend groups can replace the AIDS bar height trend if a trend group is selected next. There is no AIDS algorithm for the COLD SHUTDOWN mode. Therefore, if COLD SHUTDOWN was the last top-level display selected and any of the four AIDS supporting displays is selected, the right side of the display will initially be blank. The left side will display only the date, time, unit designation and the message MODE: COLD SHUTDOWN AIDS NOT APPLICABLE. At this point, any of the trend graphs can be selected to be displayed, starting in the lower right side.

3.3. AIDS Methodology

As stated above, the purpose of the AIDS portion of the SAS is to identify the "most likely" accident(s) to be occurring based on the behavior of current, measurable and operating parameters. In order to accomplish this purpose, one needs to construct a "relative indicator" function that is composed of a weighted sum of response functions of independent variables:

$$B_j = \sum_{i=1}^{n} W_{j,i} Q_{j,i} \quad ,$$

Table 1. Primary AIDS Parameters*

Derived Parameter	OCA		SGTR		LOSC		ICC	
	NOR	HU/CD	NOR	HU/CD	NOR	HU/CD	NOR	HU/CD
Reactor coolant system pressure	X	X	X	X				
Pressurizer level	X[1b]		X[1b]		X[1a]			
Containment environment	X	X	X[2]					
Containment radiation	X	X						
Secondary radiation			X	X				
Steam generator level			X[3]		X[5]			
Steam generator pressure					X[5]			
Core exit temperature							X	X
Reactor vessel level							X[4]	X[4]

Modifiers:

1. The reactor coolant system average temperature is computed as a volume-weighted average of the cold leg temperatures and the core exit temperatures. An average density is then computed which is used to adjust the observed pressurizer level for change due to temperature. If the change in the pressurizer level is due to a corresponding temperature change, then this parameter is enabled for LOSC (*1a*). If the pressurizer level decrease is due to mass loss (i.e., greater decrease than can be explained by the temperature change), then this parameter is enabled for a LOCA and a SGTR (*1b*).

2. If the pressurizer level is decreasing due to a mass loss (see modifier *1*, above), then a lack of abnormal containment environment conditions is enabled for a SGTR.

3. If there is no auxiliary feedwater flow to a steam generator, then steam generator level is enabled for a SGRT. Use the highest steam generator level to determine the setpoint violation.

4. If all the reactor coolant pumps are on and the reactor vessel level reading is on scale in the narrow range, then the reactor vessel level is enabled for ICC. If all the reactor coolant pumps are off and the vessel level is at the top of the core, then the reactor vessel level is enabled for ICC. This modifier was established for use with a d/p-type vessel level instrument system, which is affected by reactor coolant system flow.

5. For LOSC, use lowest steam generator level and pressure to determine setpoint violation.

*LOCA = loss of coolant accident; SGTR = steam generator tube rupture; LOS = loss of secondary coolant; ICC = inadequate core cooling; NOR = normal algorithm; HU/CD = heatup/cooldown algorithm.

Table 2. Secondary AIDS Parameters[a]

Derived Parameter	LOCA		SGTR	LOSC		ICC	
	NOR	HU/CD		NOR	HU/CD	NOR	HU/CD
Pressurizer power-operated relief valves position	X	X	(None)				
Pressurizer safety valves position	X	X					
Pressurizer relief tank pressure	X	X					
Containment environment				X	X		
Steam generator steam flow/ feed flow mismatch				X	X		
Subcooling						X	X
Steam generator level (lowest)						X	
Source range power level[b]						X	X

[a]See Table 1 for definitions of acromyms.
[b]Enabled only if the reactor is tripped and the source range high voltage is on.

where

B_j = bar height = "relative indicator" that accident type j is in progress,

j = accident type (1 = LOCA, 2 = SGTR, 3 = LOSC, 4 = ICC),

$W_{j,i}$ = weighting factor (importance) of parameter i to accident type j,

$Q_{j,i}$ = response function of parameter i for accident type j,

and

x_i = accident indicator or parameter relevant to the identification of an accident.

For different accidents, i does not necessarily define the same parameter; i.e., if $j = 1$ (LOCA), x_1 = reactor coolant system pressure; if $j = 2$ (SGTR), x_1 = pressurizer level.

The only condition imposed on the weighting factor is that for any accident type j, the sum of the weighting factors of all parameters i must equal one. This condition is imposed so that no accident type has a greater expectancy than any other accident type. However, the weighting factor is to be constructed to reflect the importance of the parameter i to accident type j. For example, containment humidity may be used as an indicator for both $j = 1$ (LOCA) and $j = 3$ (LOSC), but it is more important for $j = 1$ than for $j = 3$. For $j = 1$, the containment humidity is x_4, while for $j = 3$, it is x_2. Thus

$$W_{1,4} > W_{3,2}$$

and

$$\sum_{i=1}^{n} W_{1,i} = 1 = \sum_{i=1}^{n} W_{3,i} \quad .$$

The weighting factors $W_{j,i}$ can be developed in a number of ways. Initially they were constructed by a method that is based on engineering judgment and experience and is described below. (These factors can then be fine tuned based upon parametric studies and test results.)

Each of the key parameters is evaluated as to whether it is:

- a direct indicator of an accident,
- an indicator of an accident,
- based on reliable instrumentation, or
- based on instrumentation that covers the range of accident conditions without changing the number of channels used.

The key parameter is given a score of 1 for each true statement in this evaluation.

The weight of each parameter is then determined by the sum of these scores divided by the number of different accidents in which it is a PAP. These fractions are then summed for all parameters in each accident and the percentage is computed. This information is summarized in Tables 3 through 6.

The response function $Q_{j,i}$ is defined as the expectancy of accident type j based solely upon the characteristics and behavior of parameter x_i. It is assumed that the response function is composed of m independent, linearly combined functions where the independent variables are $x_i(t)$, $dx_i(t)/dt$, setpoints, etc.:

$$Q_{j,i} = F_{j,i}^{(1)}(x_i) + F_{j,k}^{(2)}\left(\frac{dx_i}{dt}\right) + \ . \ . \ . + F_{j,i}^{(m)}[g(x_i)] \quad ,$$

Table 3. Primary AIDS Parameters—Normal Mode (Weighting)

Description	Weighting			
	LOCA	SGTR	LOSC	ICC
Reactor coolant system pressure	4/2 1,1,1,1	4/2 1,1,1,1		
Pressurizer water level	2/3 1,0,1,0	3/3 1,1,1,0	2/3 1,0,1,0	
Reactor coolant system core exit temperature				3/1 1,1,0,1
Reactor vessel level				2/1 1,0,0,1
Containment environment	3/2 1,0,0,1	3/2 1,0,1,1		
Containment radiation	2/1 1,0,0,1			
Secondary radiation		3/1 1,1,1,0		
Steam generator water level		3/2 1,1,1,0	3/2 1,1,1,0	
Steam generator pressure			4/1 1,1,1,1	
Total	6.17	9.0	6.17	5.00

Table 4. Primary AIDS Parameters—Heatup/Cooldown Mode (Weighting)

Description	Weighting			
	LOCA	SGTR	LOSC	ICC
Reactor coolant system pressure	4/2 1,1,1,1	4/2 1,1,1,1		
Pressurizer water level				
Reactor coolant system core exit temperature				3/1 1,1,0,1
Reactor vessel level				2/1 1,0,0,1
Containment environment	3/2 1,0,0,1			
Containment radiation	2/1 1,0,0,1			
Secondary radiation		3/1 1,1,1,0		
Steam generator water level				
Steam generator pressure				
Total	5.5	5.0		5.00

Table 5. Primary AIDS Parameters—Normal Mode (Percent)

Description	Percent (Maximum)			
	LOCA	SGTR	LOSC	ICC
Reactor coolant system pressure	32.41	22.22		
Pressurizer water level	10.86	11.11	10.86	
Reactor coolant system core exit temperature				60.00
Reactor vessel level				40.00
Containment environment	24.31	16.67		
Containment radiation	32.42			
Secondary radiation		33.33		
Steam generator water level		16.67	24.31	
Steam generator pressure			64.83	

where $g(x_i)$ is the desired conditioning of parameter x_i to obtain the proper variable; i.e.,

$$g(x_i) = \frac{dx_i}{dt} \quad ,$$

$$Q_{j,i} = \sum_{\omega=1}^{m} F_{j,i}^{(\omega)}[g(x_i)] \quad .$$

The only condition imposed on the response function is that it can vary only between zero and one. This condition is imposed so that the contribution of a parameter to the bar height of a given accident cannot be greater than the importance ($W_{j,i}$) of that parameter:

$$0 \leqslant Q_{j,i} \leqslant 1 \quad .$$

Table 6. Primary AIDS Parameters—Heatup/Cooldown Mode (Percent)

Description	Percent (Maximum)			
	LOCA	SGTR	LOSC	ICC
Reactor coolant system pressure	36.36	40.00		
Pressurizer water level				
Reactor coolant system core exit temperature				60.00
Reactor vessel level				40.00
Containment environment	27.27			
Containment radiation	36.37			
Secondary radiation		60.00		
Steam generator water level				
Steam generator pressure				

Additionally, since each functional component of the response function is independent, it is required that each function's component vary only between zero and one. This condition is imposed so that the contribution of a parameter to the bar height of a given accident is strictly non-negative. Therefore,

$$0 \leqslant F_{j,i}^{(\omega)}[g(x_i)] \leqslant 1 \quad .$$

Now, the bar height (B_j) is determined as the sum of the products of the importance ($W_{j,i}$) and expectancy ($Q_{j,i}$) of a parameter (x_i) to accident type j:

$$B_j = \sum_{i=1}^{n} W_{j,i} Q_{j,i} \quad ,$$

$$B_j = \sum_{i=1}^{n} W_{j,i} \left\{ \sum_{\omega=1}^{m} F_{j,i}^{(\omega)}[g(x_i)] \right\} \quad .$$

The response function should be constructed such that quick transients and slow changes both contribute to the expectancy independently.

$$Q_{j,i} = \underbrace{F_{j,i}^{(1)}[g(x_i)]}_{\substack{\text{Transient} \\ \text{Response}}} + \underbrace{F_{j,i}^{(2)}[g(x_i)]}_{\substack{\text{Proportional} \\ \text{Response}}} + \ldots \quad .$$

Next, expressions for the transient response and proportional response are derived. In these derivations, it is assumed that the parameter setpoint is below the normal operating value of the parameter. In a case where this assumption is invalid, the signs of some of the terms change.

3.3.1. Conditions on Transient Response

The conditions to be placed on the transient response are as follows:

(1) The transient response should contribute only when a setpoint (R_0) is violated.

$$F_{j,i}^{(1)}[g(x_i)] = H[R_0 - x_i(t)] \left\{ \cdots \right\} \quad ,$$

where H [...] is the "Heaviside" function,

$$H(x) = \begin{cases} 1 \, , & x \geqslant 0 \\ 0 \, , & x \leqslant 0 \end{cases} ,$$

and {...} represents other terms which are to be defined below. In the SAS implementation, the Heaviside function is used for both upper and lower setpoints. For the upper setpoints, it is defined as

$$H(x) = \begin{cases} 1 \, , & x \geqslant \text{setpoint} \\ 0 \, , & x < \text{setpoint} \end{cases} ,$$

where x is the parameter data. When checking lower setpoints, the logic is defined as

$$H(x) = \begin{cases} 1 \ , x > \text{setpoint} \\ 0 \ , x \leqslant \text{setpoint} \end{cases}$$

In both cases, $H(x) = 0$ reflects that the data are in a normal condition, and $H(x) = 1$ reflects that they are in an abnormal condition.

(2) The transient response should be proportional to the rate of change of the parameter

$$F_{j,i}^{(1)}[g(x_i)] = H[R_0 - x_i(t)] \left[-\frac{dx_i(t)}{dt} \right] \left\{ \cdot \cdot \cdot \right\} \ .$$

(3) The transient response should contribute only when the parameter rate of change is in one direction. This condition is imposed to prevent negative contributions to the expectancy when the derivative changes sign:

$$F_{j,i}^{(1)}[g(x_i)] = H[R_0 - x_i(t)] \ H\left[-\frac{dx_i(t)}{dt} \right] \left[-\frac{dx_i(t)}{dt} \right] \left\{ \cdot \cdot \cdot \right\} \ .$$

(4) The most recent data should be a more significant contribution to the value of the transient response than previous data. Therefore, the data should be incrementally decayed. As an example of incremental decay, assume a function $A(t)$ which has an input of a_i at $t = t_i$. Additionally, $A(t)$ is a cumulative function and requires incremental decay:

$$\begin{aligned} A(t_0) &= a_0 \ , \\ A(t_1) &= a_1 + a_0 e^{(t_0 - t_1)/\tau} \ , \\ A(t_2) &= a_2 + e^{(t_1 - t_2)/\tau} \left[a_1 + a_0 e^{(t_0 - t_1)/\tau} \right] \\ &= a_2 + a_1 e^{(t_1 - t_2)/\tau} + a_0 e^{(t_0 - t_2)/\tau} \ , \end{aligned}$$

$$A(t_\kappa) = \sum_{\rho=0}^{\kappa} a_\rho e^{(t_\rho - t_\kappa)/\tau} \ .$$

Because $A(t)$ is a cumulative function, $A(t)$ becomes cumbersome to calculate after a fairly short time of plant operation. For example, after only four hours of plant operation, the equation will require a sum of 14,400 terms for $A(t)$ when the calculation occurs every second. For a longer period of time, the equation becomes statistically infinite. Therefore, the following method is used as an approximation:

$$A(t_\kappa) = \sum_{\rho=0}^{\kappa} a_\rho e^{(t_\rho - t_\kappa)/\tau} \quad ,$$

$$\begin{aligned} A(t_\kappa) &= a_\kappa + \sum_{\rho=0}^{\kappa-1} a_\rho e^{(t_\rho - t_\kappa)} \\ &= a_\kappa + \sum_{\rho=0}^{\kappa-1} a_\rho e^{(t_\rho - t_{\kappa-1})/\tau} \cdot e^{(t_{\kappa-1} - t_\kappa)/\tau} \\ &= a_\kappa + \left[\sum_{\rho=0}^{\kappa-1} a_\rho e^{(t_\rho - t_{\kappa-1})/\tau} \right] \cdot e^{-\Delta t/\tau} \quad , \end{aligned}$$

where $\Delta t = t_{\kappa-1}$ is the time difference between two consecutive time steps, and the term in brackets is actually $A(t_{\kappa-1})$ according to the definition. The equation for $A(t_\kappa)$ becomes

$$A(t_\kappa) = a_\kappa + a_{\kappa-1} \cdot e^{-\Delta t/\tau} \quad .$$

The total response function equals the sum of the current response function $A(t_\kappa)$ and the response function of the previous time step $A(t_{\kappa-1})$ times the decay exponential $e^{-\Delta t/t}$:

$$F_{j,i}^{(1)}[g\{x_i(t_\kappa)\}] = H[R_0 - x(t_\kappa)] H\left[-\frac{dx_i(t_\kappa)}{dt}\right]\left[-\frac{dx_i(t_\kappa)}{dt}\right]\left\{ \cdots \right\}$$

$$+ F_{j,i}^{(1)}[g\{x_i(t_{\kappa-1})\}] \cdot e^{-\Delta t/\tau} \quad .$$

Now, upon close examination of the expression for the transient response term, $F_{j,i}^{(1)}[g(x_i)]$, one can deduce the following:

(a) If the derivative term $dx_i(t_p)/dt$ is continuously increasing, the transient response term is an unbounded increasing function.

(b) If the derivative term is a function which increases to a maximum and then remains at that maximum, the value of the transient response term also increases to a maximum and remains at the maximum. In other words, the transient response term saturates to a value determined by the maximum value of the derivative and the time constant τ.

(c) If the derivative term is a function which increases to a maximum value and then decreases to a constant value, the value of the transient response also increases to a maximum and then decays away to a constant value.

(d) If the derivative term is a function which is continuously decreasing from some initial positive value, the transient response increases to a maximum and then decreases to zero.

Examples of each of the above observations can be found in Table 7 and in Fig. 1. In the examples a represents the derivative term and $A(t)$ represents the transient response.

(5) The saturation (equilibrium) value of the transient response term and the time to reach saturation value of the transient response term are both a function of the time constant in the exponential term of the transient response. If the value of the time constant is small, then the time to saturation is short, the saturation value is low, and the decay time is short. If the value of the time constant is large, then the time to saturation is long, the saturation value is large, and the decay time is long. The time constant should be constructed such that:

(a) The time to reach the saturation value increases as the rate of change of the parameter decreases. A maximum time response should be established based on the accident, parameter of interest, and parametric studies.

(b) The decay time response should be longer than the accumulation time response. It should take longer to decay the transient response than it does to build up the transient response. A minimum and maximum decay time response should be established based on the accident, parameter of interest, and parametric studies.

(c) The saturation value of the transient response for all parameters and all accidents should be maintained within a range specified by a minimum and maximum saturation value. This will ensure that slowly changing parameters will derive the transient response term to one. Additionally, this will ensure that the decay of the transient response will begin from some maximum value so that the transient response is available to contribute to the expectancy after plant conditions have been stabilized.

Therefore, the time constant should be functionally dependent upon (1) the rate of change of the parameter, and (2) changes in the rate of change of the parameter:

Table 7. Examples of Cases a-d

(Assume: $t_{i-1} - t_1$ = constant = -2 sec, t = 6 sec)

Time (sec)	Case a		Case b		Case c		Case d	
	a_i	A(t)	a_i	A(t)	a_i	A(t)	a_i	A(t)
0	1	1	1	1	1	1	9	9
2	2	2.72	2	2.72	2	2.72	8	14.4
4	3	4.95	3	4.95	3	4.95	7	17.3
6	4	7.54	4	7.54	4	7.54	6	18.43
8	5	10.41	5	10.41	5	10.41	5	18.21
10	6	13.46	6	13.46	6	13.46	4	17.05
12	7	16.64	6	15.64	5.5	15.14	3	15.21
14	8	19.92	6	17.21	5.0	15.85	2	12.90
16	9	23.28	6	18.33	4.5	15.86	1	10.24
18	10	26.68	6	19.13	4.0	15.36	0	7.3
20	11	30.12	6	19.71	3.0	14.01	0	5.2
22	12	33.58	6	20.12	2.0	12.04	0	3.7
24	13	37.06	6	20.42	2	10.06	0	2.7
26	14	40.55	6	20.63	2	9.61	0	1.9
28	15	44.06	6	20.78	2	8.89	0	1.3
30	16	47.57	6	20.89	2	8.37	0	0.9
32	17	51.09	6	20.97	2	8.00	0	0.7
34	18	54.60	6	21.03	2	7.73	0	0.5
36	19	58.13	6	21.07	2	7.54	0	0.3
38	20	61.65	6	21.09	2	7.40	0	0.2
40	21	65.17	6	21.11	2	7.30	0	0.1
42	22	68.70	6	21.13	2	7.23	0	0.1
44	23	72.22	6	21.14	2	7.18	0	0.10
46	24	75.75	6	21.15	2	7.15	0	0.0
48	25	79.28	6	21.15	2	7.12	0	0.0
50	26	82.81	6	21.16	2	7.10	0	0.0

$$\tau = \psi\left[\frac{dx_i}{dt} \quad , \quad \frac{d^2x_i}{dt^2}\right] \quad ,$$

where ψ = function which determines τ. For SAS implementation, the time constant τ is actually dependent upon the transient response function:

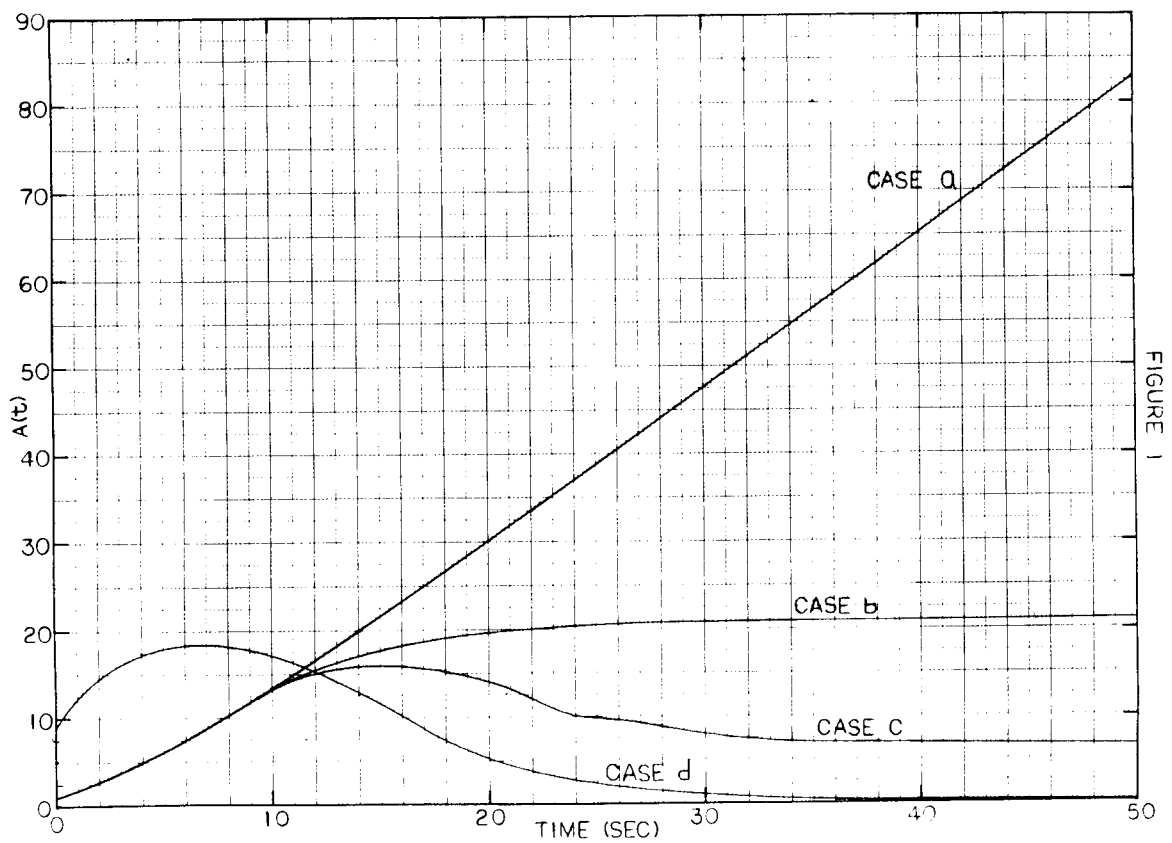


Fig. 1. Examples of Cases a—d.

$$\tau = \begin{cases} 1 & \text{if transient response} = 0, \\ 2 & \text{if transient response} \neq 0. \end{cases}$$

The transient responses should contribute only when the parameter rate of change is in one direction. If the parameter rate of change is in the direction that will contribute to the transient response, a 2-s time constant is used. Otherwise, a 1-s time constant will be used. For example, the high setpoint for the reactor coolant system pressure is 2310 psig. If the pressure increases from 2300 psig and passes the setpoint to 2350 psig, a 2-s time constant will be used in this time period. If the pressure drops back to 2250 psig thereafter, the 1-s time constant is used. In thi way, the decay of the transient response will be slower than the buildup of the transient response due to the larger time constant. This also ensures the decay of the transient response will begin from some maximum value.

(6) The transient response should be scaled so that it is dimensionless. The scaling should be constructed such that it amplifies small rates of change but does not significantly affect the magnitude (except for dimensional conversions) of large rates of change. Therefore, the scaling should be functionally dependent upon the parameter rate of change. An example of such a scaling function is

$$G_{j,i}^{(1)} = \frac{C_1}{-\frac{dx_i}{dx}+1} + C_2 \quad ,$$

where $G_{j,i}^{(1)}$ is the scaling function and C_1 and C_2 are constants determined by "trial and error" from the Indian Point simulator transient data tapes. The best constants were determined to be $C_1 = 1.0$ and $C_2 = 0.5$. Thus, for a large rate of change,

$$G_{j,i}^{(1)} = \frac{1}{1 \pm \frac{dx_i}{dt}} + 0.5$$

$$= \frac{1}{\pm \frac{dx_i}{dt}} + 0.5 \approx 0.5 \quad ,$$

and for small rate of change,

$$G_{j,i}^{(1)} = \frac{1}{1 \pm \frac{dx_i}{dt}} + 0.5 \approx 1.5 \quad .$$

Essentially, the constants chosen will not significantly affect the magnitude of a large rate of change (0.5), but will amplify a small rate of change by 3 (i.e., 1.5/0.5 = 3). Therefore,

$$F_{j,i}^{(1)}[g(x_i)] = \sum_{\rho=0}^{\kappa} H\left[R_0 - x_i(t_\rho)\right] \cdot H\left[-\frac{dx_i(t_\rho)}{dt}\right] \cdot \left[-\frac{dx_i(t_\rho)}{dt}\right] G_{j,i}^{(1)} \cdot e^{(t_\rho - t_\kappa)/\tau} \quad .$$

For simplification of notation, the following definition is made:

$$F_{j,i}^{(1)}[g(x_i)] = \sum_{\rho=0}^{\kappa} \left[\cdot \cdot \cdot \right] \quad .$$

(7) The transient response cannot be greater than one:

$$F_{j,i}^{(1)}[g(x_i)] = \min \left\{ 1, \sum_{\rho=0}^{\kappa} \left[\cdot \cdot \cdot \right] \right\} .$$

3.3.2. Conditions on Proportional Response

The conditions to be placed on the proportional response are as follows:

(1) The proportional response should contribute only when a setpoint R_1 is violated:

$$F_{j,i}^{(2)}[g(x_i)] = H[R_1 - x_i(t)] \left\{ \cdot \cdot \cdot \right\} .$$

(2) The proportional term should be dependent upon the difference between the current value of the parameter and the setpoint R_1:

$$F_{j,i}^{(2)}[g(x_i)] = H[R_1 - x_i(t)] \; [R_1 - x_i(t)] \left\{ \cdot \cdot \cdot \right\} .$$

(3) The proportional response should be scaled so that it is dimensionless. The scaling should be constructed such that if the parameter value goes below a predefined value R_2, the proportional response is driven to a value of one:

$$F_{j,i}^{(2)}[g(x_i)] = \min \left\{ 1, H[R_1 - x_i(t)] \left[\frac{R_1 - x_i(t)}{R_i - R_2} \right] \right\} .$$

Now, by summing the transient response and the proportional response, the desired response function (expectancy) will have been constructed:

$$Q_{j,i} = F_{j,i}^{(1)}[g(x_i)] + F_{j,i}^{(2)}[g(x_i)]$$

$$Q_{j,i} = \min \left\{ 1, \sum_{\rho=0}^{\kappa} \left[\cdot \cdot \cdot \right] \right\} + \min \left\{ 1, H[R_1 - x_i(t)] \left[\frac{R_1 - x_i(t)}{R_1 - R_2} \right] \right\} .$$

However, the form of $Q_{j,i}$ given above violates the condition that the response function can vary only between zero and one. The above form allows the response function to vary between zero and two. In order to meet the range condition imposed on the response function, one can relax the range conditions of the transient response and the proportional response such that they are required only to be nonnegative:

$$Q_{j,i} = \min \left\{ F_{j,i}^{(1)}[g(x_i)] + F_{j,i}^{(2)}[g(x_i)] \right\} .$$

The equation for the bar height B_j is

$$B_j = \sum_{i=1}^{n} W_{j,i} Q_{j,i}$$

where

$$Q_{j,i} = \min \left(1, \left\{ H[R_1 - x_i(t_\kappa)] \left[\frac{R_1 - x_i(t_\kappa)}{R_1 - R_2} \right] + \sum_{\rho=0}^{\kappa} H\left[R_0 - x_i(t_\rho)\right] H\left[-\frac{dx_i(t_\rho)}{dt} \right] \left[-\frac{dx_i(t_\rho)}{dt} \right] G_{j,i}^{(1)} e^{(t_\rho - t_\kappa)/\tau} \right\} \right) ,$$

$$G_{j,i}^{(1)} = \frac{C_1}{-\dfrac{dx_i}{dt} + 1} + C_2 ,$$

$$\tau = \psi\left[\frac{dx_i}{dt}, \frac{d^2 x_i}{dt^2} \right] .$$

4. CONCLUSION

The models, algorithms and weighting factors were developed and programmed into the SAS. The system was tested at a simulator using licensed plant operators under various types of accident conditions. The AIDS portion of SAS performed very well considering it was the very first time the concept was tested. In all cases, the AIDS properly identified the accident in progress in a timely manner and the operator was able to use the AIDS displays to augment and improve his response to the emergency situation.

AIDS is intended to be a model of a good operator that is continually scanning and evaluating the response of many parameters simultaneously. The AIDS displays are designed to show the factors that are involved in this process, so that the operator is not given a single answer he cannot confirm. It is the judgment of those involved in the project that the AIDS concept was successfully demonstrated during the testing program. Also identified were some areas that could use additional work. Overall, the program was very successful in demonstrating the usefulness of an advanced diagnostic technique.

Paper 6

HUMAN COGNITION

Donald A. Norman
Department of Psychology and Institute for Cognitive Science C-015
University of California, San Diego

1. INTRODUCTION

The study of human cognition encompasses the study of all mental phenomena, from the receipt and interpretation of sensory information to the final control of the motor system in the performance of action. The cognitive scientist examines all intermediary processes, including thought, decision making, and memory and including the effects of motivation, states of arousal and stress, the study of language, and the effects of social factors. The field therefore ranges over an enormous territory, covering all that is known or that should be known about human behavior. The reason I feel compelled to make this statement here is that at the meeting that gave rise to this review of human cognition, some thought it useful to distinguish between human skilled behavior and human cognitive behavior. To the cognitive scientist, such a distinction is unthinkable. If behavior involves mental mechanisms, it is by definition a *cognitive* behavior (although not necessarily *conscious* cognitive behavior).

It should come as no surprise that we are far from understanding human cognition. The task is extremely difficult, both because of the complexity of the human cognitive system and because of the extreme difficulty in making scientifically reliable observations of mental operations. Still, considerable progress has been made in the development of psychological models. We have our best understanding in peripheral areas (sensation, perception, and working memory) and our least understanding in the more central areas (consciousness, motivation, emotion, social interactions, and thought).

It is not possible to summarize the current state of knowledge about cognition with any great confidence that we know the "correct" answer about any aspect of the work. Nonetheless, we have models that provide good characterizations of certain aspects of the data and situations. Even if these models should prove to be incorrect, they do provide good approximate descriptions of people's behavior in some situations, and these approximations will still apply even when the underlying theories have changed. In the text that follows, I provide a quick description of models within a number of areas of human cognition and skill and some general theoretical frameworks with which to view human cognition. The frameworks are qualitative descriptions that provide a way to view the development of more detailed, quantitative models and, most important, a way of thinking about human performance and skill.

Figure 1 shows a sketch of the important components of the human information processing system.[1] The figure is misleading in making an extremely complex system appear simple. The human is not at all simple, which brings me to the following comments on the assessment of human performance by means of a PRA.

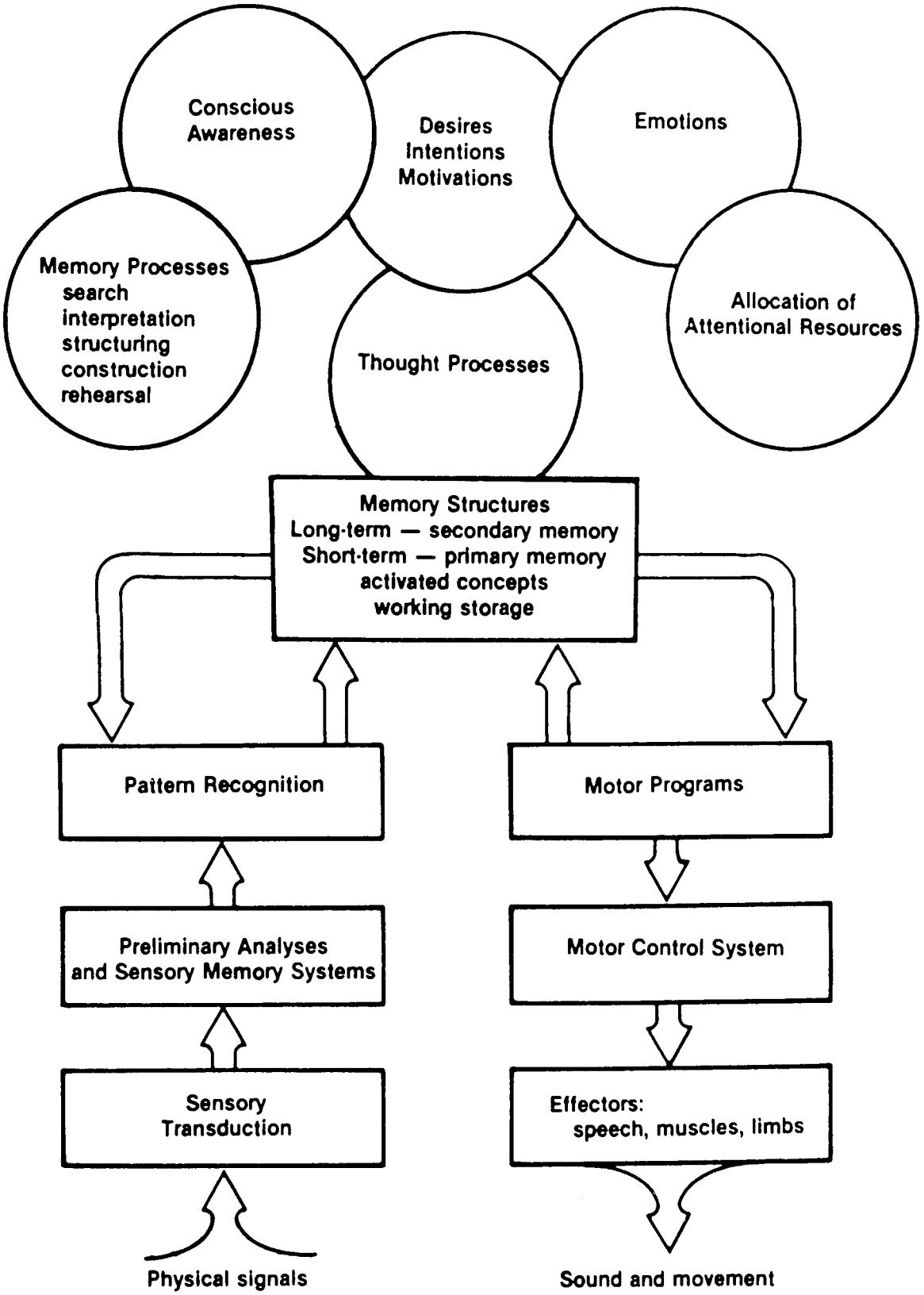


Fig. 1. Rough Model of Human Information Processing Structure. (From Norman, ref. 1).

Nuclear power plants are complex systems. They are often analyzed by means of fault and event trees to determine the sequencing of operations and to get estimations of error probabilities of the components, and therefore relative frequencies of major accidents and plant damage [the procedure of Probability Risk Analysis (PRA)]. It is obvious that human error — by operators, by plant maintenance personnel, by construction crews, and by designers — can play a significant role in the cause of avoidance of accidents. Accordingly, there has been considerable effort devoted to the assignment of human error frequencies during PRAs. The *Handbook of Human Reliability Analysis*[2] is designed to provide information about the likelihood of human error in various conditions.

My complaint with the inadequacy of assigning simple probabilities to human error is this. When the PRA is performed on the entire plant, millions of scenarios and operations are considered. Within this analysis, it is tempting to treat the human as any other component, such as a pump. Just as it is important to know the likelihood that a pump will fail, it is important to know the likelihood that a human will fail to read a meter or throw a switch. But note:

The complexity of the human EXCEEDS that of the nuclear power plant.

Moreover, we can at least look at the plant and examine the components. We have the designer's drawings and intentions and the operating characteristics of all the components. Despite all this, the full analysis of the events and fault trees is quite difficult. Now imagine doing this for the human, where we do not have design drawings, we cannot get access to the components, and we do not even know the fundamental principles of operation of some of the more essential components (such as those systems shown in Fig. 1). Actually, we don't even know what all the components and states are.

We are faced with the dilemma that one of the components of the plant analysis exceeds the complexity of the entire plant. Just as one cannot assign a simple error probability to the entire plant, it makes little sense to attempt to apply a simple error probability to the human.

Humans have many internal states, limited memory and attentional capacities, and varying motivational and physiological states. To determine the range of behavior requires consideration of these variables. Error probabilities may range from close to zero to very close to unity, depending upon the state of the person. If the state diagram of the human were known, then it might be possible to get a realistically tight distribution of error estimates. At the least, we need a model of the human that approaches in complexity that of the plant. That model should then be used in conjunction with the plant PRA to determine accident probabilities.[a]

If we need an analysis of the human, what is the likelihood that we can get one? On the one hand, this is an enormous research project, one that will take decades. On the

[a]The people who worked on the *Human Reliability Handbook* are aware of these issues and they emphasize the need to consider the state of the human in making probability assignments. Essentially, they suggest doing an informal state analysis of the human. I believe my suggestion is compatible with this philosophy, except that I urge development of a formal, quantitative model rather than the subjective rules-of-thumb discussed in the Handbook.

other hand, we may already know enough to allow us to make relatively simple, approximate models. Even if it is known that these approximate models are incorrect and highly oversimplified, if we can devise simple state diagrams of the human (with perhaps tens of states and variables), we will have made a considerable advance in our ability to make predictions. Yes, such a model will horribly complicate the task of analysis. But these tasks are already horrendously complex just for the plant. It should hardly be surprising that when we add the proper analysis of a component whose complexity is greater than that of the plant (but which is still unknown), that the complexity of the analysis might have to increase. In the next several sections that follow, I provide some of the information on human performance that will have to be considered in devising simple approximate models of the human plant operator. In a final section, the implications of the various aspects of human performance to design are discussed.

2. CONSIDERATIONS IN HUMAN PERFORMANCE

2.1 Attention

We are finite beings with finite processing capability. Among other things, this means that we cannot process deeply all the sensory information that arrives, nor follow through with much depth several trains of thought at any one time. Moreover, doing one task may interfere with the doing of another. These limitations have been studied under the topic of "attention," and the result has been to address a reasonable number of questions:[b]

What purpose does attention serve? Simplified Answer: Attention seems to be needed for understanding, for memory, for evaluation and decision making, possibly even to perceive properly. Attention does not seem to be required for some kinds of monitoring and for the performance of skilled, routine tasks.

How many things can a person attend to at once? Simplified Answer: If by "attend" is meant consciously, then the answer is "very few," perhaps only one.

Can one avoid distraction (interruption) from other events when attempting to attend to one? Simplified Answer: Usually not, but on the other hand, one cannot count on being able to note other events. Here the answer is highly dependent upon the nature of the main task and of the distracting tasks.

Do all tasks require attention? Simplified Answer: Probably not. Highly skilled performance of routine tasks can be automated, freeing it from the requirement for continued attention.

Do all tasks compete with one another, or are there some that can be performed together without interference, while others cannot? Simplified Answer: Not all tasks compete with one another for attentional resources. We believe that there are a number of resource pools from which tasks draw.

Although we still are ignorant of many features of the human attentional system, current work suggests that the human can be characterized as a limited resource system,

[b]For convenience, I provide some simple answers. Current research provides much more complex answers, but these will suffice as reasonable approximations if it is understood that the answers apply only in limited conditions and situations.

although probably with multiple resource pools. Thus, if several processing tasks all draw from the same resource pool, then once their total demands exceed the total available, they will interfere with one another, a state called "resource limitation." Performance on a task can be limited in several ways: by the amount of resources that can be applied to the processing (resource-limited processing); by the quality of the data available to the process, even when sufficient resources are available (data-limited processing); and by output considerations (output-limited processing). When a single resource must be shared among competing tasks, one gets a resource-operating characteristic, a function that relates performance on one task against performance on another, usually showing a tradeoff — as one improves, the other deteriorates (see Fig. 2). (For more information, see ref. 3.)

The relatively simple picture of resource tradeoffs is made more complex by the realization that in the human there are probably multiple resource pools, so that one task may draw from a set of resources that only partially overlaps the pools required of another task. Moreover, there may be flexibility in the use of resources, so that a process that cannot get sufficient units of one resource might be able to substitute resources from some combination of other pools. This multiple pools picture is clearly more realistic, but it poses great difficulties for the collection and interpretation of data on this topic. (This is discussed in ref. 4.)

A further complication occurs because the very nature of human performance seems to change in a qualitative manner as a person goes from novice or beginning performance of some task to expert performance. Experts seem to have less dependence upon resource limitations than do novices, although it is not clear whether this is because they use different resource pools that are more specialized (and therefore less apt to be used by competing tasks) or whether they have so specialized (or compiled) their processes that they require little or no common resources.

Consciousness plays an important role in learning and in complex decision making (which includes performance on complex tasks). However, we know almost nothing of the nature of conscious and subconscious systems, and so the exact role in attentional limitations is not clear. One possibility is that most of the attentional limitations that have been studied to date are really conscious limitations, and that although the conscious control system is a single, limited resource system, the structures used by subconscious (and expert) processes have multiple resources and are not so limited (see refs. 5-7). In summary, therefore:

- Conscious control is a limited resource system.
- Systems can be limited in quality of performance because of data limits, output limits, and resource limits.[8]
- There are multiple resources,[4] and tradeoffs among resource pools are possible.
- There are cost-benefit tradeoffs in the focus of attention.[9]
- Attentional limits do not seem to apply for tasks that are automated (or they draw from different resource pools).[5]

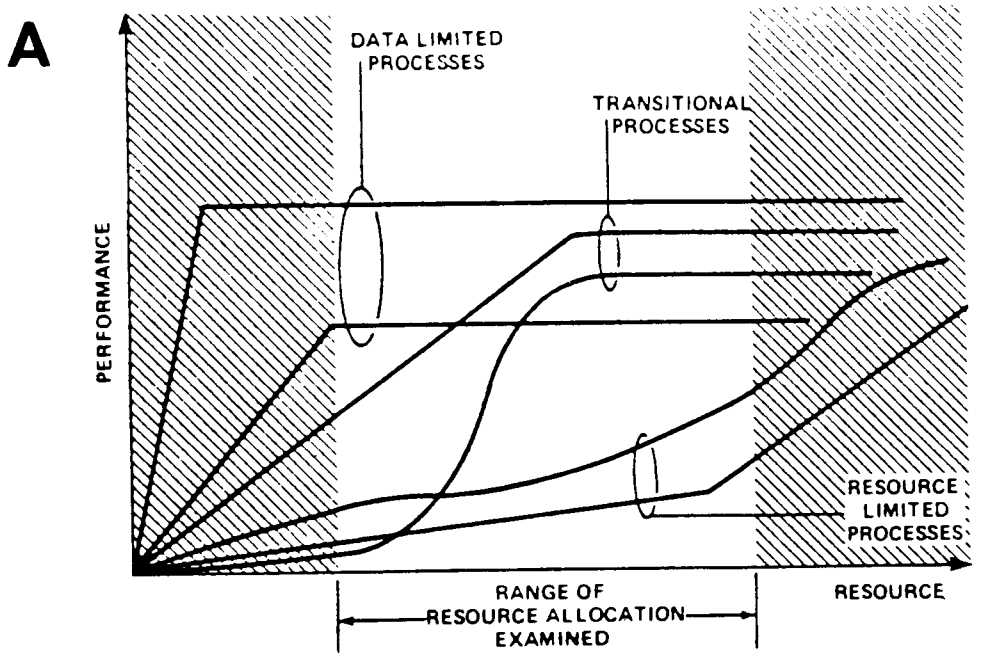


Fig. 2a. Data-Limited and Resource-Limited Behavior.

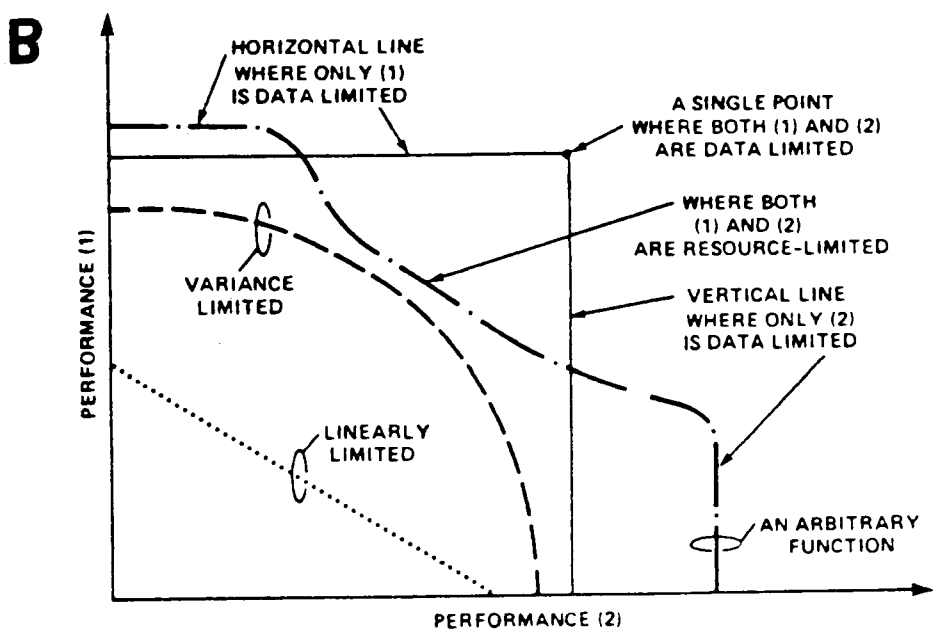


Fig. 2b. Resource Operating Characteristic. The performance on one task (1) will be affected by performance on another task (2) when the two tasks draw from the same resource pool and the sum of the resources demanded exceed what are available. (From Norman & Bobrow, ref. 3).

2.2 Arousal and Stress

The role of stress on performance is not as well established as one would like, despite the strong claims that are often made in the applied literature. The reason for this is clear: it is difficult to get experimental situations which reproduce realistic stress or arousal in such a way that controlled, systematic measurement of performance can be made. As a result, our conclusions come from analyses of simple laboratory exercises in which the arousal that is invoked does not correspond to real stresses in actual working conditions, or from anecdotes and unverifiable accounts of performance under actual stressful situations.[c]

The common assumption about performance under different levels of arousal is summarized in the *Yerkes-Dodson curve.* Basically, performance actually appears to improve with an initial increasing level of arousal, but then, as arousal levels continue to increase, performance deteriorates (Fig. 3). Along with this change in performance are several qualitative assessments:

Narrowing. The traditional view is that under high arousal or stress, a person's analyses focus upon a limited set of alternatives. This is thought to apply equally to both perceptual and mental analyses of the situation. The result can be enhanced performance if the issues that the narrowed focus of attention concentrates upon are the most relevant to the situation, but it can lead to deterioration if the issues are not relevant.

Repetition. The same response or analysis is apt to be repeated over and over again. For example, if the door won't open, try again, only harder. The literature has a large number of examples of situations where this behavior has led to disaster (such as cases where death has resulted after continued and unsuccessful attempts were made to escape fire by pushing against a door when it could have simply been pulled open). On the decision-making side this is thought to lead to a rigidity of decision making, so that once an interpretation of the situation has been determined, it will be maintained, even in the face of contrary information (negative information is either not considered or is discounted). I have come to call this *cognitive hysteresis.*

Stereotypical Behavior (and Reversion to Earlier Behavior). It is believed that stress causes reversion to the most well-learned responses to a situation, even when these are not necessarily the appropriate ones. Moreover, if one response pattern has been superseded by a newer one, it is thought that the chance of erroneously returning to the earlier pattern is increased when under stress. (Justification for this belief is intuitively strong, but scientifically weak. Thus, the justification includes stories such as: if you carefully learn a new pattern of behavior for a task — such as a new grip for tennis — when in stress, the old pattern — the old way of holding the racquet — is very apt to return.)

Decreased Resources. It is thought that there are decreased resources available to the stressed person. There are two possible explanations. One is that the physiological aspects of stress change the resources available. The other is that the mental resources are unaffected, but because a person starts devoting more and more attention to the stress itself,

---

[c]Subjective accounts of behavior, especially those offered long after the event, are subject to considerable distortion. Moreover, in the absence of detailed study of the actual events and subject responses, it is usually unclear exactly what alternatives were available and just exactly how performance was affected.

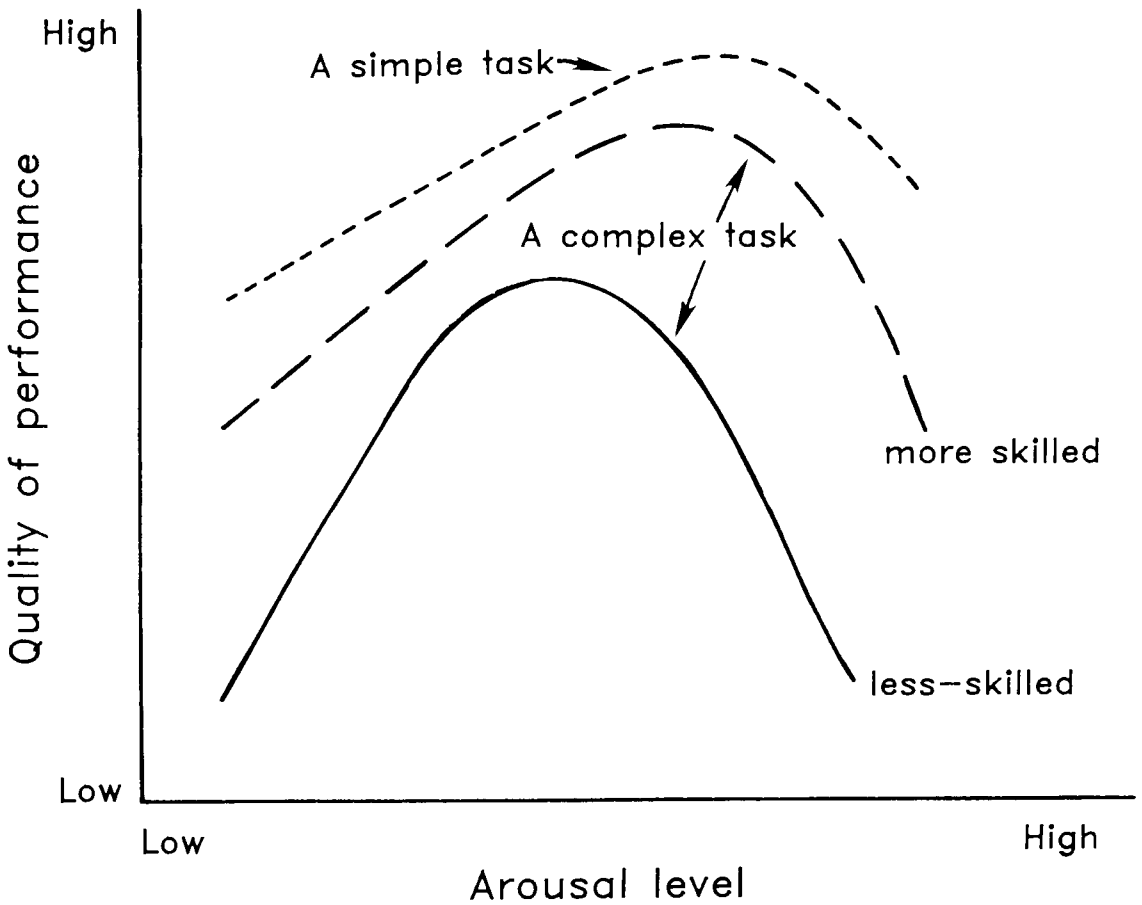


Fig. 3. The Yerkes-Dodson Curve. Maximum performance occurs with some nonzero amount of stress, but higher amounts of stress cause deterioration of performance. The greater the skill of the person, the greater amounts of stress can be tolerated before deterioration occurs (and the greater amount required to reach the maximum performance). These curves should be interpreted with caution, for the data that support them are weak, and the relationship may not always hold. Adapted from Kahneman, ref. 10.

the result is a decreased availability of resources. In either case, the result is functionally equivalent. In similar fashion, the initial peak of the Yerkes-Dodson curve may be a result of increased resources, either as a result of physiological mechanisms or because the early attentional focusing provides a more efficient use of resources.

The Role of Training. Systematic training and overtraining can minimize the effects of stress on performance. Thus, if the person has been repeatedly exposed to the stressful situation (for example, in simulators), the appropriate performance can be taught and practiced. This is the standard procedure in commercial and military aviation. As a result, when a real emergency occurs, performance is apt to be only minimally affected by stress. There are three reasons for this: (1) If the appropriate response has been sufficiently well practiced, it can be performed with a minimum of resources (the result of

"automating" responses). (2) If the responses are sufficiently well practiced, then it is the proper response that is apt to be triggered by the situation. (3) If people have been exposed to the situation often enough in simulation, then the familiarity with the situation and the experience of handling it successfully tends to minimize the buildup of stress. Stress results from the psychological interpretation of the situation, so extreme confidence and familiarity can greatly diminish or abolish stress.

2.3 The Memory System

Approximate models for memory do exist, especially for sensory memory and working (or short-term) memory (STM). Models of long-term memory (LTM) are more complex and require knowledge of representational issues.

2.3.1 Short-Term or Working Memory. Short-term memory (also called working memory) is that part of the memory system that allows us to maintain relatively active and accessible information. STM has a limited capacity, estimated to be approximately 5 to 7 "items." It is easily disrupted by other activity or items "similar" to the information being stored. STM seems to be modality specific, so that information about visual, spatial, auditory (and speech), and motor information all seem relatively independent of one another; either they occupy independent STMs or they do not interfere with one another in a common STM. The properties are easily summarized:

- The capacity of STM is about 5 to 7 "items."
- Information from one modality does not interfere with information from another.
- Meaningful information is much more easily retained than nonmeaningful or unfamiliar material.

Although the existence and properties of STM memory are extensively documented and described, there recently has been renewed interest in attempts to describe the workings of human memory without postulating any separate short-term mechanism. Whatever the status of the theoretical debate, the behavioral phenomena seem clear. Humans have limited working memories; we are able to get easy access to a limited amount of information active at any one time.

Note that in the description of STM, capacity has been described in terms of "items" and interference effects have been described in terms of "similarity." These are critical terms. Do *not* think of STM as a mechanism with a certain fixed amount of space. Rather, it is best to think of it as that part of the entire memory system that is temporarily "activated." Moreover, the item should be thought of as a "pointer" to concepts and relations that already exist within memory. As a result, the amount of information that any given item refers to can vary considerably. This is why experts at any task seem to be able to remember so much more than beginners at the task; the STM capacity is the same, but because each item for the expert is in actuality a constellation of well-learned concepts, the effective STM capacity is increased.

2.3.2 Models of Long-Term Memory. Models of long-term memory (LTM) are more complex than those for STM. LTM models require knowledge of three areas: (1) learning (acquisition), (2) retrieval, and (3) knowledge representation. While all these areas are undergoing considerable study, they are still incompletely developed.

We simply do not know very much about *learning* and *memory retrieval.* From a practical point of view, it is perhaps best to consider the particular problems of memory storage and retrieval more than the formal theoretical models.

The capacity of LTM is huge — probably measured in billions of items. However, there is a major problem in LTM in getting access to information already there. Norman and Bobrow[8] suggest that retrieval is best described as an iterative process in which one starts with a "description" of the information sought, retrieves something, evaluates it, uses this information to refine the initial description, and then repeats the retrieval cycle. (You can see that under stress, this iterated, resource-demanding retrieval process is apt to be seriously disrupted.) One critical thing that determines the success of the retrieval is the manner in which the information was originally encoded or "described." Various mnemonic strategies can be thought of as providing well-established descriptions to aid in the original encoding and subsequent retrieval.

A current way to view the *representation* of information in LTM is in terms of organized units — schemas. Schemas provide structure to memory. Current psychological views of schemas are closely related to the work on knowledge representation in artificial intelligence (and to the concepts of "frames," knowledge representation languages, and production systems).

The important point about schemas for the current purposes can be summarized in this way:

- A schema provides an organized method of representing information, including means for representing the interrelationships among related concepts and beliefs.
- All new information is encoded in terms of previously known information by building new schemas based upon those that already exist for the already known information.
- The structure of schemas provides default values and "inheritance" properties, thus allowing partially specified information to have "typical" values filled in automatically.

The automatic filling in property of schemas (by default values and inheritance) means that people may actually get only partial information, but may believe that they have received full information. This can lead to error if the default values are inappropriate.

2.4 Decision Making

2.4.1 Short-Term Memory Limits. In making judgments, people are heavily influenced by memory limitations. Short-term memory limitations prevent detailed logical reasoning in depth without external aids. To follow a chain of reasoning requires holding in memory considerable information. To evaluate several alternatives means to hold in memory several such chains of reasoning. The game of tic-tac-toe (naughts and crosses) can exceed the limit of short-term memory. By this statement I mean that if you attempted to work out all possible paths mentally, you would not be able to do so: the number and the depth

of the possibilities exceeds STM memory abilities. We do well at things such as tic-tac-toe and chess by learning relevant patterns, by restructuring the situation so that reformulations and heuristics serve, and so on. In tic-tac-toe, expert players know the proper moves and responses for the opening moves though experience and study. It has been estimated that the chess master has acquired a "vocabulary" of 50,000 configurations that can be retrieved from long-term memory to guide in the selection of a move and analysis of the situation.

2.4.2 Long-Term Memory Limitations. Finally, LTM plays several roles in decision making. As we have seen, the expert (as in chess) makes considerable use of experience with a large repertoire of situations. Meaningful events are more easily handled than abstract events, even when the logic of the two is identical. More and more evidence shows that people make decisions based upon experiences, not upon logical deductions. Worse, they make judgments based upon what they remember of their experiences, and this therefore means basing decisions upon those events most available from memory, often a biased sample of their actual experiences.

2.4.3 Use of Negative Information. People do not make good use of negative information. As a result, negative information is not as useful in human decision making as it ought to be. Instead, people will hold on to an erroneous hypothesis for considerable time, continuing to seek positive confirmation, even when negative information that could disconfirm the hypothesis is readily available. (Positive evidence is often not logically as powerful as potentially disconfirming evidence, but nonetheless, people tend to seek the positive reassurance.) (The shortcomings in human inference are treated in refs. 11-13.)

2.5 Social Demands Upon Performance

One major group of factors that influences behavior is that of "social demands": peer pressure, the person's perception of economic demands, the perception of the desires of management, etc. These can lead to conflicting goals, for the behavior that has been instructed or that appears in the manuals may conflict with what is good for peer prestige, for the company, for the customer, or for society.

There are many pressures on people, including their understanding of what others expect of them. These social (and societal) pressures are called demand characteristics and they are especially important in complex systems. There are four major issues that can be considered:

People may tailor their behavior for approval of their colleagues. Example: Piloting a military aircraft into a dangerous situation and escaping successfully wins the admiration of fellow pilots, even though it is a serious error and dangerous to enter the situation. Steering the aircraft so as to go off the runway while taxiing is considered humiliating and possibly detrimental to one's career, even though its consequences are minimal.[14]

People may avoid actions that do not appear to be approved by management. Example: The collision of two 747 aircraft at Tenerife was greatly affected by pressures on the KLM pilot to take off in order to save KLM management hotel and crew costs, coupled with the decision of the Pan American pilots to obey the ground controller against their better judgments (after a few thwarted attempts to argue the point).[15]

People may tailor their actions to protect their own reputation or self-image. Example: Operators may behave conservatively to avoid being blamed, even if the conservative action is not the most appropriate. Military pilots, after crashing and while in danger of fire, have been known to stay in the cockpit, resetting switches and memorizing panel configurations in order to protect themselves for the eventual crash review.

People may protect one aspect of system design at the expense of another. Example: Commercial aviation pilots are so concerned for the comfort of their passengers that on occasion they will not brake sufficiently hard, thus potentially leading to loss of life (a factor in the American Airlines crash on the Virgin Islands[16]).

2.6 Human Error

Humans err, sometimes from lack of knowledge, sometimes because they have received erroneous information, sometimes for extraneous reasons, and sometimes for internal-processing reasons that are not easy to determine. In general, any given error usually has multiple causes, and it is not easy to pin down a single reason. Much of modern technological equipment seems designed to aid in the generation of error and oftentimes is relatively intolerant of error. It is my belief that many of the accidents that result from human error should be blamed more on the equipment design, the social demands, and the task requirements than on the operator. Elsewhere I have analyzed errors and attempted a classification based upon information-processing causes.[17–19] From these analyses, we can learn certain design lessons. To simplify the story, major classes of errors are *mode errors, description errors, capture errors, and activation errors.* The lessons from analyses of these classes of errors can be summarized this way:

- Mode errors suggest the need for better feedback.
- Description errors suggest the need for better system configuration.
- Capture errors imply the need for better feedback.
- Activation issues suggest the importance of displaying the options and of providing feedback.
- People will make errors, so make the system insensitive to them.

3. IMPLICATIONS FOR DESIGN

The various aspects of human performance discussed above contain a number of lessons for design. It seems clear that one should design for the worst case, which means the tired, stressed operator who is receiving a large amount of information about a situation, thereby heavily taxing short-term memory capabilities. It is important to avoid having the operator revert to stereotypical behavior, or to form premature hypotheses (which then will receive considerable positive confirmation — even in the face of disconfirming, negative information), and to avoid tasks that require heavy load on memory or logical reasoning. These considerations mean that one should design for an operator with considerably reduced capabilities.

Note that humans are remarkably capable, so that they can learn to perform well even in the most poorly designed situations. The fact that the skilled operator normally performs well under many circumstances should not disguise the fact that the human operator is still human. Under severe, worst-case conditions, the operator cannot perform as expertly as usual, especially if the conditions are either novel or infrequently encountered. But it is precisely these conditions that must be guarded against. Design should emphasize the 1 out of a 1000, not the usual situation.

Social factors can determine performance as much as cognitive capacities and the physical and operating system demands. Do not neglect the social and organizational structure of the people who use the equipment. Make it possible for an operator to make decisions or ask for help without appearing to be stupid or to be flouting authority, or to be sacrificing a social requirement for an operational requirement.

3.1 Conclusions Based on the Analyses of Stress

If the effects of stress are considered (and the mental capabilities are assumed to be diminished), then often appropriate design can be incorporated to assist the operators in times of stress. The point is, do not assume a fully alert, well-functioning operator, but rather assume one who does not examine alternatives, does not seek out information, and has diminished mental abilities. The situation can be designed so as to minimize the effects of stress. If information about the situation is presented in a direct, meaningful fashion, then decision making will be enhanced. Example: Because the immediate response of people who are trying to exit a room while highly stressed is to push against doors, fire doors are provided with a long, horizontal bar designed so that the door will open, even when locked, by force against any part of the bar.

3.2 Lessons from the Analyses of Errors

An analysis of errors can provide help in avoiding some major problems. These analyses must be supplemented with other methods, including an understanding of the nature of a person's mental image of the system that is being used and an understanding of the human information-processing capability of the user. The analyses presented in this paper make several points:

- *Feedback*: The state of the system should be clearly available to the user, ideally in a form that is unambiguous and that makes the set of options readily available so as to avoid mode errors.
- *Similarity of response sequences*: Different classes of actions should have quite dissimilar command sequences (or menu patterns) so as to avoid capture and description errors.
- *Reversible actions*: Actions should be reversible (as much as possible) and where both irreversible and of relatively high consequence, they should be difficult to do, thereby preventing unintentional performance.
- *Consistency of the system*: The system should be consistent in its structure and design of command so as to minimize memory problems in retrieving the operations.

These considerations, coupled with similar analyses of the properties of the users' mental models of the system, lead to other sets of rules for performance, including the notion of a

system image. The system image should be the first thing set up by the designer, and all commands, feedback, and instruction should be designed to be consistent with that system image. However, this is another story — one that is still in the early stages of development and is not fully ready to be discussed here.

3.3 Design Philosophy

Start with a functional specification of the system that is to be designed. Specify this from the user's point of view. Construct a model of the user, taking into account:

- *The limited short-term memory of the user.* That is, the likelihood that the procedure will be retrieved depends upon the frequency with which it is normally used.

  Implications: Provide STM memory aids such as on-line "help" services, reasonable default values, menus, mnemonics, functional layout, and so on.

- *Mechanical, perceptual, and motor limitations.*

  Implications: Follow established human factor guidelines.

- *The limited attentional capability, both in capacity and in duration.*

  Implications: Minimize resource demands. Don't overload the user with alarms. (The 747 aircraft has 455 warnings and caution alerts.)

- *The tendency to reason through prior examples, or through representative cases rather than by a formal, unbiased analysis procedure.* That is, decisions are affected by the point of view given by the statement of the problem, which includes the way the displays are configured (see ref. 11).

  Implications: If decisions must be made by some explicit weightings of factors (and history), provide the relevant information to the user along with decision-making tools. Provide information with a focus chosen according to the desired decision bias.

- Norman's law: *If there is potential for error, no matter how remote, someone will make it.*

  Implications: If you believe Norman's law, then you will make the design tolerant of error, with sufficient state information available to let the operator identify what has been done and to allow reversible operations.

ACKNOWLEDGMENT

The portions of the research described in this report that were performed at the Institute for Cognitive Science were conducted under Contract N00014-79-C-0323, NR 667-437 with the Personnel and Training Research Programs of the Office of Naval Research, and were sponsored by the Office of Naval Research and the Air Force Office of Scientific Research.

REFERENCES

1. D. A. Norman, *Perspectives on Cognitive Science*, published jointly by Erlbaum, Hillsdale, NJ, and Ablex, Norwood, NJ (1981); Chinese and Japanese translations in progress.
2. A. D. Swain and H. E. Guttmann, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications,* NUREG/CR-1278, draft report for interim use and comment, October 1980, U.S. Nuclear Regulatory Commission, Washington, DC.
3. D. A. Norman and D. G. Bobrow, "On Data-Limited and Resource-Limited Processes," *Cognitive Psychology* **7**, 44-64 (1975).
4. D. Navon and D. Gopher, "On the Economy of the Human Processing System: A Model of Multiple Capacity," *Psychological Review* **86**, 214-55 (1979).
5. W. Schneider and R. Shiffrin, "Controlled and Automated Human Information Processing: I. Detection, Search, and Attention," *Psychological Review* **84**, 1-66 (1977).
6. R. M. Shiffrin and W. Schneider, "Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory," *Psychology Review* **84**, 127-90 (1977).
7. D. A. Norman and T. Shallice, *Attention to Action: Willed and Automatic Control of Behavior*, (Tech. Rep.), Center for Human Information Processing, University of California, San Diego, La Jolla, CA (1980).
8. D. A. Norman and D. G. Bobrow, "Descriptions: An Intermediate State in Memory Retrieval," *Cognitive Psychology* **11**, 107-23 (1979).
9. M. I. Posner, "Orienting of Attention," *Quarterly Journal of Experimental Psychology* **32**, 3-25 (1980).
10. D. Kahneman, *Attention and Effort*, Prentice-Hall, Englewood Cliffs, NJ (1973).
11. *Judgment Under Uncertainty: Heuristics and Biases*, D. Kahneman, P. Slovic, and A. Tversky (Eds.), Cambridge University Press, New York (1982).
12. R. Nisbett and L. Ross, *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice-Hall, Englewood Cliffs, NJ (1980).
13. P. Wason and P. N. Johnson-Laird, *Psychology of Reasoning: Structure and Content*, Harvard University Press, Cambridge, MA (1972).
14. T. V. Golder, *P-3 Pilot Errors: A Conceptual Approach*, Center for Advanced Research, Naval War College (June 1978).
15. P. A. Roitsch, G. L. Babcock, and W. W. Edmunds, *Human Factors Report of the Tenerife Accident. Aircraft Accident Report: Pan American World Airways, Boeing 747, N737 AP; KLM Royal Dutch Airlines, Boeing 747, PH-BUF; Tenerife, Canary Islands, March 27, 1977*, Air Line Pilots Association, Washington (1979) (actual report not dated).
16. *Aircraft Accident Report No. NTSB-AAR-77-1*, National Transportation Safety Board (NTSB), Washington, DC (April 27, 1976).
17. D. A. Norman, "Categorization of Action Slips," *Psychological Review* **88**, 1-15 (1981).
18. D. A. Norman, "A Psychologist Views Human Processing: Human Errors and Other Phenomena Suggest Processing Mechanisms," Invited Address, *Proc. of the International Joint Conference on Artificial Intelligence*, Vancouver (1981).
19. D. A. Norman, "Steps toward a Cognitive Engineering: Design Rules Based on Analyses of Human Error," *Proc. of the Conference on Human Factors in Computer Systems*, Gaithersburg, MD, March 15-17, 1982, Institute for Computer Sciences and Technology, National Bureau of Standards, U.S. Department of Commerce.

Paper 7

THE ROLE OF CONCEPTUAL MODELS IN NUCLEAR POWER PLANT OPERATION

Michael D. Williams, Thomas P. Moran, and John Seely Brown
Xerox Palo Alto Research Center

1. INTRODUCTION

In a high-risk system, such as a nuclear power plant, the most critical property of its control system is that it function *correctly* to keep the plant in safe operating states. This is why, for example, so much effort is expended to validate the correct performance of critical control software. But the human operator also has critical "mental software" for interpreting the state of the plant and for controlling it. The question we raise is: why do we not make an equally serious effort to validate this mental software?

The most obvious component of the operator's software is the standard plant operating procedures; but another, less obvious, component is a *conceptual model* of how the plant works. A conceptual model is necessary for the intelligent use of the operating procedures, such as recognizing the appropriate conditions for their application and making small alterations to them to adapt to specific situations. But the conceptual model is most crucial for dealing with unanticipated situations, such as multiple faults, for which no standard procedures apply. Given the critical role of the conceptual model in plant control, its most important property is its correctness, especially in unanticipated situations—a property we call *robustness.*

A crucial objective in plant operation (and perhaps licensing) ought to be to explicitly train operators to develop, perhaps with computer aids, robust conceptual models of the plants they control. The question is whether we are actually able to develop robust conceptual models and *validate* their robustness. Cognitive science is just beginning to come to grips with this problem. This paper describes some of the evolving technology for building conceptual models of physical mechanisms and some of the implications of such models in the context of nuclear power plant operation.

2. THE EXISTENCE OF QUALITATIVE CONCEPTUAL MODELS

Let us distinguish between two kinds of cognitive models. The first kind consists of models of the behavior or performance of the power plant operator (perhaps embedded in a larger model of the whole power plant system). This is the kind of model that most other modelers at this conference are referring to. The second kind of cognitive model is a model *of the power plant* that the operator is postulated to have in his head. This is the kind of model we are referring to when we speak of "conceptual models."

Consider how an operator might think about a simple mechanism, for example, the pressure-regulating valve diagrammed in Fig. 1. How might the operator explain its behavior, that is, describe *how* it is able to regulate pressure? A likely explanation might run as follows:

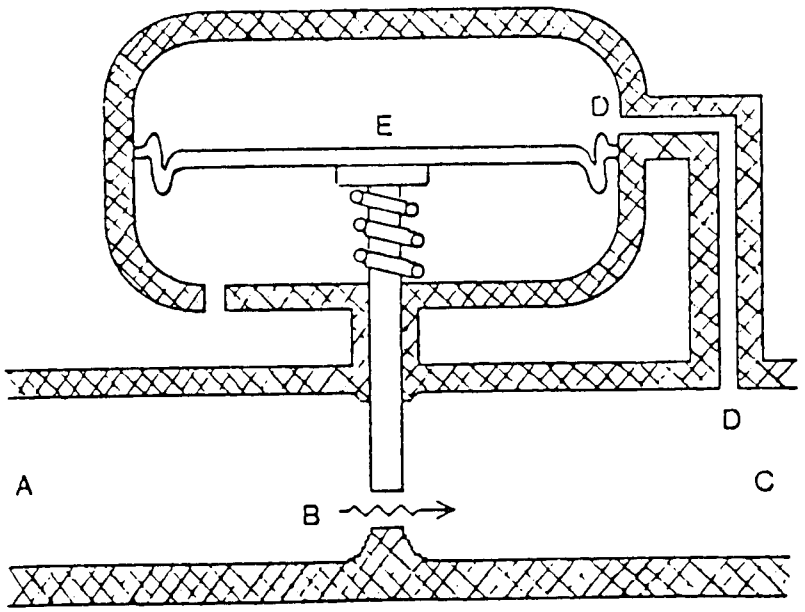


Fig. 1. Diagram of a Pressure-Regulating Valve. (From de Kleer and Brown, ref. 1).

> An increase in source (*A*) pressure increases the pressure drop across the valve (*B*). Since the flow through the valve is proportional to the pressure across it, the flow through the valve also increases. This increased flow will increase the pressure at the load (*C*). However, this increased pressure is sensed (*D*), causing the diaphragm (*E*) to move downward against the spring pressure. The diaphragm is mechanically connected to the valve, so the downward movement of the diaphragm will tend to close the valve thereby pinching off the valve. Because the flow is now restricted the output pressure will rise much less than it otherwise would have and thus remains approximately constant.[1]

The characteristics of this explanation that we want to emphasize are its *qualitative* and *causal* nature. It is not given in terms of complex constraint equations, nor in even vaguely numerical terms. The conceptual model behind this explanation is expressed in terms of increasing and decreasing values and in terms of component parts and explicit local, casual relations between those parts. It is our view that operators understand the behavior of the reactor as a mechanism in just these terms and, further, that this kind of understanding provides operators with the basis for learning, recalling, using, repairing, and modifying the plant operating procedures.

Learning is, by its nature, an incremental process. None of us expects a person to learn a complex operating procedure in a glance. Rather, it must be acquired in bits and pieces. The conceptual models come into the problem by providing some of the semantic glue to connect the bits and pieces of the procedure together. If nothing else, we expect conceptual models to serve as a useful learning tool.

Even if learning were a holistic process, recall is not.[2] As a practical matter, if an operator tries to recall the steps in a procedure, a great deal of patching must be done. Inevitably, the operator forgets a step and must fall back on recalling some other fragment of procedure to patch over what he has forgotten, or he must use various semantic structures to restrict the set of possible next actions, or he must invent that piece of procedure anew from a basic understanding of the mechanism or the constraints on the procedure itself. Among the main sources of information for these recall and reconstruction processes are the same conceptual models we expect to be useful in learning.

Even if one could recall a procedure in its entirety, its execution will almost inevitably require some kind of adaptation to the particulars of the current context. Some step in the procedure is bound not to be applicable or to be simply wrong. For example, in a reactor casualty control procedure, how could it be known in advance that a particular detector has been reading 10% high for the last 2 months or that it is in the midst of calibration or that one of countless other plant conditions has been altered. In such cases we expect people to use common sense to adapt to these minor variations of a basic task. But, even if one had an exact operating procedure described for every anticipated plant condition, operators would inevitably run into novel situations. And we want them to be able to react at least plausibly (if not intelligently) in these unforeseen circumstances. This requires the ability to construct new procedures on the fly. One obvious way to make this manageable is to take a nearly right procedure and modify it or to compose a new procedure out of known procedures that serve as fragments of the solution to the problem at hand.

3. CRITERIA FOR QUALITATIVE CONCEPTUAL MODELS

There are two kinds of criteria any modeler must be sensitive to: (1) criteria that specify *what phenomena* are to be modeled and (2) *performance* criteria that specify the dimensions along which a model is to be evaluated.

Our goal is not to model the operator's behavior directly, but to model the operator's view of the plant, that is, to build models of the plant that capture the same kind of understanding of the plant that operators seem to have with their mental models. Thus, we want our models not only to *predict* the behavior of the plant but also to be able to produce *explanations* of how that behavior is achieved. Further, we want our models to support human-life epistemologies—that is, reasoning in *qualitative, causal,* and *functional* terms. Note in the example of the pressure regulator that the qualitative terminology of the explanation is there deliberately; it is not simply an approximation from some underlying quantitative reasoning process. Ask an operator a quantitative question (for example, *how far* the valve opens), and he will shift to a different form of reasoning, bringing in a whole different set of factors. It is clear that such quantitative inferences are not the precursors to the qualitative explanation.

Performance criteria for evaluating models fall into two categories: *psychological* and *logical.* The psychological criteria include *learnability* (how readily can we teach the

model to an operator?) and *usability* (given that an operator knows a model, how easily can he use it to predict the plant's behavior?). The logical criteria include *precision* (how accurately does the model predict the plant's behavior?), *scope* (what range of plant behaviors does the model predict?), and *robustness* (how well does the model predict when the model itself is perturbed to reflect an unanticipated state of the plant?).

It is our conjecture that the logical criteria ought to be the principal initial criteria for constructing conceptual models of nuclear power plants and, especially, that robustness (which we shall further discuss below) ought to be the primary focus.

4. THE EMERGING TECHNOLOGY FOR ANALYZING QUALITATIVE CONCEPTUAL MODELS

There is an emerging technology for constructing models that attempt to satisfy these criteria. Burton, Brown, and de Kleer,[3] de Kleer and Brown,[1,4,5] Forbus,[6] Forbus and Stevens,[7] Williams, Hollan, and Stevens,[8] and Reiger and Grinberg,[9] among others, discuss a variety of techniques for constructing qualitative models of physical devices (such as electronic circuits, pressure regulating values, thermostats, heat exchangers).

One of the simplest examples is the door buzzer of de Kleer and Brown,[4] shown in diagram form in Fig. 2. The operation of the buzzer is quite simple: If the switch is open, the spring holds the clapper down, and there is no current in the circuit. Once the switch is closed, a current flows. The current through the coil causes a magnetic field that pulls the clapper open. The opening of the clapper breaks the circuit, causing the current to stop. The current stopping makes the field go away, permitting the clapper to close (due to the force of the spring). As this causal chain cycles, the clapper repeatedly strikes its contact, making the buzzer sound.

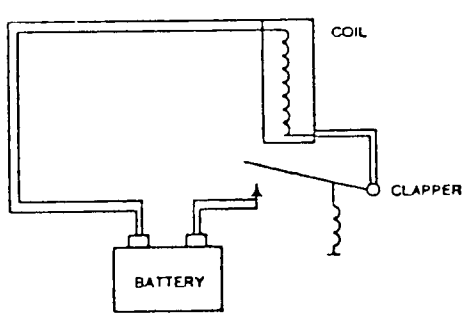
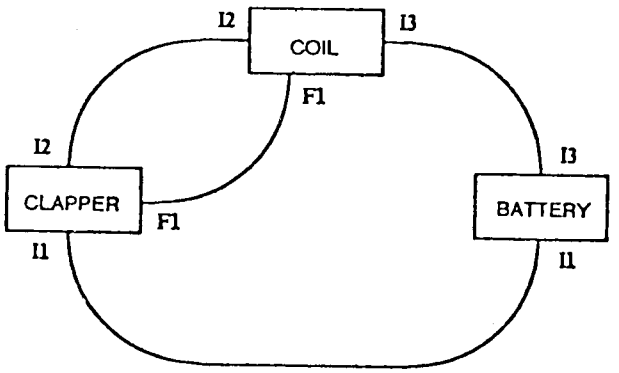


Fig. 2. Diagram of a Simple Door Buzzer. (From de Kleer and Brown, ref. 5).

Figure 3 is a diagram of the simplified *device topology* of the buzzer, that is, the components out of which the buzzer is constructed and the connections between these components. Below the diagram are rules that specify the qualitative behavior of each of the components, called *component models.* This model is precise, and it can be *run* to predict how the components interact to produce the overall device behavior. Not only can behavior be predicted, but also reasons can be assigned for the behavior, in terms of which events in the running cause which other events, that is, causal explanations.




```
CLAPPER : OPEN:
            I1 ⇐ 0, I2 ⇐ 0
            IF F1 = 0 CAUSES: clapper will become CLOSED.
          CLOSED:
            I1 ⇐ 1, I2 ⇐ 1
            IF F1 = 1 CAUSES: clapper will become OPEN.

COIL : ON:
         F1 ⇐ 1
         IF I2 = 0 CAUSES: coil will become OFF
         IF I3 = 0 CAUSES: coil will become OFF.
       OFF:
         F1 ⇐ 0
         IF I2 = 1 CAUSES: coil will become ON
         IF I3 = 1 CAUSES: coil will become ON.

BATTERY : I1 ⇔ I3.
```

Fig. 3. Device Topology and Component Models of the Buzzer. (From de Kleer and Brown, ref. 5).

Figure 4 is a diagram that sorts all of the possible causal relations derived from running the model into a coherent casual structure. This capability of automatically deriving the *behavior* of a system from the *structure* is one of the essential characteristics of this modeling technology.

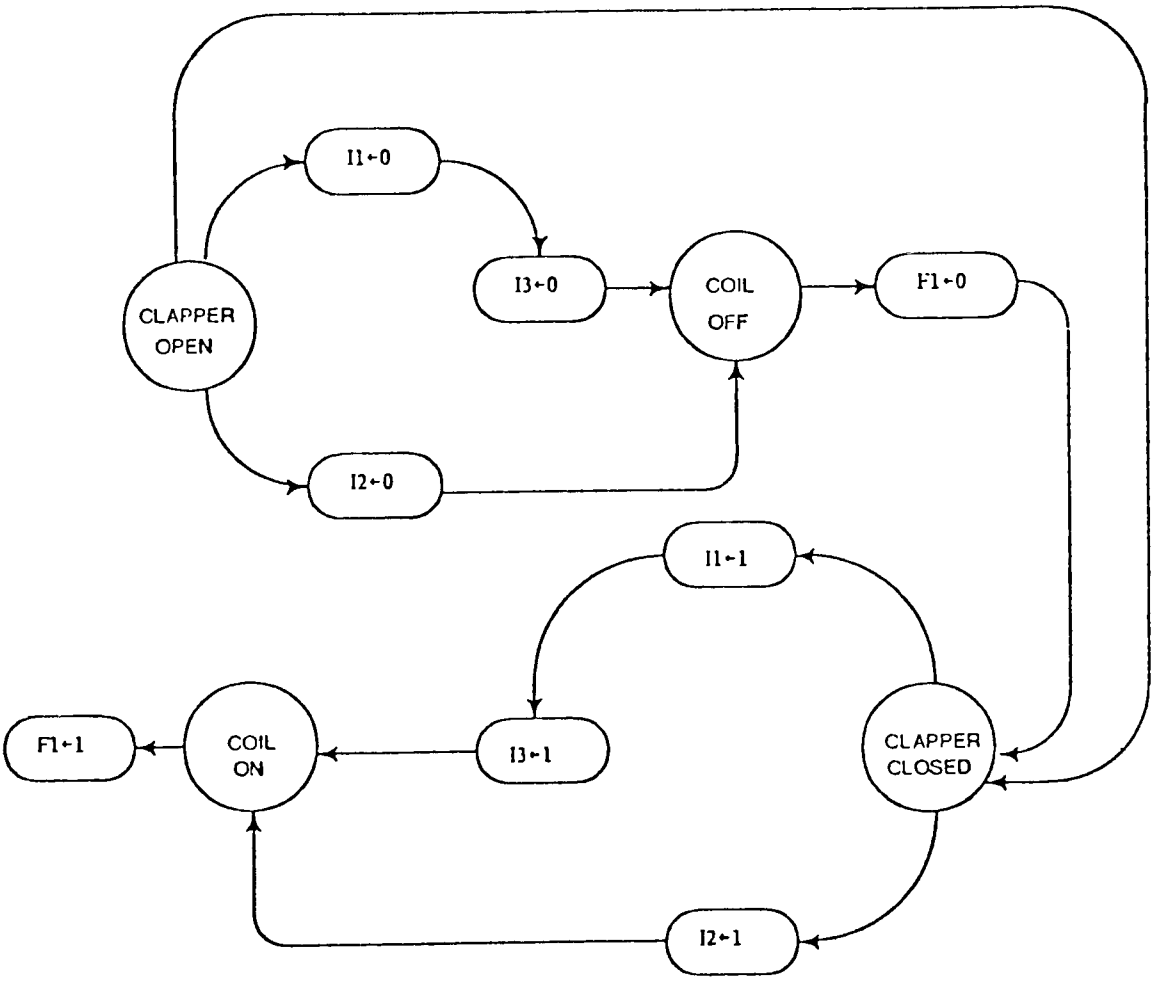


Fig. 4. Causal Structure of the Buzzer. (From de Kleer and Brown, ref. 5).

An important characteristic of these models is that their derivations can be readily converted into plausible causal explanations. The explanation above of how the buzzer works can simply be read off of Fig. 4. Such explanations can be useful in computer-based instructional systems and operational aids.

In nuclear power plant operation, the operator's conceptual model is most crucial for dealing with *unanticipated* situations, such as multiple faults, for which no standard procedures apply. This, the conceptual model's most important property is its *robustness.* Robustness refers to the ability of a model to predict correctly under perturbation (for example, the replacement or removal of one of its component models) or in new situations (as for an unanticipated set of input conditions).

There are a variety of techniques to test the robustness of a conceptual model. One technique is to modify the topology. For example, removing the battery in the model of the buzzer in Fig. 2 reveals that the model is not robust, since the model predicts that the

buzzer still works. (We will discuss the nature of this problem below.) A second technique is to fault one of the components by replacing it with a new component model of the faulted component, say a ruptured diaphram in the pressure regulator, and test the behavior of the modified model. A third technique is to test the model on new problems beyond those it was initially designed for. For example, given a conceptual model of a pressurizer, we might test its behavior under the new situation in which the reactor pressure vessel had drawn a bubble and the pressurizer was not dealing with a two-state fluid on the other side of the surge line. This collection of techniques provides a partial mechanism to validate the robustness of a conceptual model.

The construction of robust models is difficult, but a number of principles for constructing them are emerging. The essence of the difficulty is that *hidden assumptions* are often buried in the model. For example, in the buzzer in Fig. 2, note that the model for the clapper switch assumes the existence of a battery. This is why the model predicts that the device works even with the battery taken out. The principle that is being violated here is the *no-function-in-structure* principle,[4] which states that a robust model is one in which there are no assumptions about overall behavior (function) of the device in the component models and that the behavior must be derived solely from the interactions of the parts of the structure. (This is why perturbing the topology is a good test of robustness.) One of the significant features of qualitative models is that they can deal with such assumptions *within* the formal language of the model; for example, the process of running qualitative models involves careful management of assumptions in sorting out the possible behaviors.[5] This feature gives qualitative models a considerable advantage over quantitative models, such as a set of differential equations, where the assumptions are *outside* the formalism.

5. THE USABILITY OF QUALITATIVE CONCEPTUAL MODELS

Given a robust conceptual model of an entire power plant, there are a number of potential uses for it. We could simply train, test, and evaluate operators on this model with the assurance that if they use it they could understand the plant and operate it safely. However, this presupposes that the model has the appropriate psychological properties, that is, that it is learnable and usable.

Unfortunately, complex conceptual models of the type we have been discussing require a great deal of effort to derive predictions of behavior. People tend to prefer simpler models when they are adequate. For example, even though a nuclear power plant operator may have a detailed understanding of the effect of the control rods on the neutron flux in the core, he probably prefers a simple model of the qualitative reaction between the rods, the neutron flux, and reactor temperature. This is adequate for many purposes. The trick, of course, is to know when to abandon the simple model because its assumptions do not hold, and this is just the kind of knowledge that an operator needs.

We are proposing that a single, complex conceptual model of the entire power plant is not enough. We envision, in addition, a lattice of simplified conceptual models and their associated assumptions, with a robust conceptual model of the entire power plant at the root of the lattice (see Fig. 5), that is, all the simpler models being derived from the robust model by adding assumptions. The robust conceptual model is probably too complex for anyone to learn and too complex to reason with, even if one could learn it. Simplified models could be generated by systematically degrading the robust model to produce models

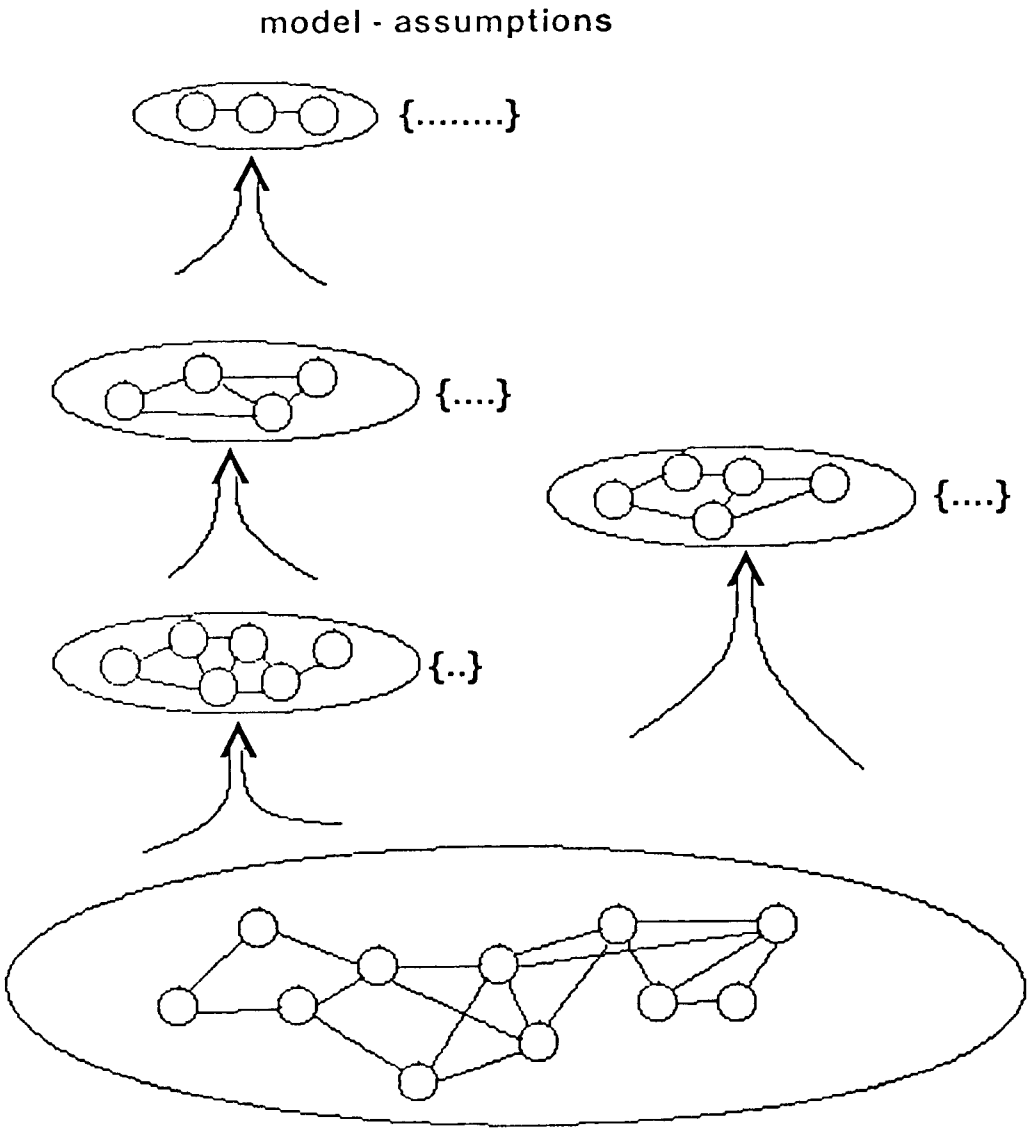


Fig. 5. Lattice of Simplified Conceptual Models and Their Associated Assumptions.

that are both learnable and useable in real-time reasoning processes. For example, if we had a simplified model of the pressurizer, we would have associated with it an explicit assumption that the reactor vessel is solid (i.e., the only steam bubble in the system is in the pressurizer). The advantage of having the entire structure is that we would explicitly know the assumptions associated with each simplified model, and the lattice shows which (more complex) models we could resort to when the assumptions of the current model are violated.

Such a lattice of models provides us with another kind of training tool. For any given simplified model that we teach to the operators, we know in advance exactly where they will fail to hold (that is, where their assumptions are violated), and we can directly train the operators to be aware of these crucial conditions. Thus, if we know that the conceptual model of the pressurizer assumes that the rest of the reactor is solid we can: (1) explicitly inform the operators of this, (2) train them for conditions where the plant is not solid, and (3) provide an alternative (but more complex) conceptual model that works under the assumption that the reactor is not solid.

The lattice of models can also be used to help us specify an effective training program. The key notion is that most physical devices are based on a simple idea, represented by a *kernel model,* that is elaborated to take account of various real-world complexities. The idea is to relate this simple kernel model to the complex, robust model (representing the real plant) by a series of patches on the kernel model that eventually transform it into the robust model, the collection of patched models forming our lattice of models. Consider trying to understand an electronic circuit. Start off by choosing a very simple kernel circuit that captures the core idea but is missing all the frills that really make it work in the context and under the conditions for which it is intended. Then identify a limitation in the kernel circuit (an assumption that restricts its domain of applicability) which can be fixed by patching the circuit. In the ideal case this means taking one component of the given circuit and replacing it with a set of components. Note that if all the required patches are of this simple form, then we have an ideal learning sequence of circuits, in which each component in the target circuit can be understood as either stemming from an underlying kernel or as a patch for a limitation in the kernel. Also, note that this kind of understanding is useful for troubleshooting or reacting to casualties, since each component is understood, in part, in terms of how the overall system might behave if that piece weren't there.

6. CONCLUSION

The two principal goals of this paper are (1) to give a brief tutorial on the status and objectives of an emerging technology for constructing human-like conceptual models of complex physical mechanisms, and (2) to present a conjecture that for the application of this technology to nuclear power plant operation the property of robustness (correctness in unanticipated situations) should be our principal initial goal.

Conceptual models are used to predict the behavior of a mechanism from the structure of that mechanism. They use a qualitative, causal, and functional terminology and can be used to generate plausible, human-like explanations of why a device works as it does. The construction of conceptual models is more difficult than it appears on the surface, though a set of principles for testing robstness is emerging.

We believe that the primary objective of the design of conceptual models for the operation of nuclear plants should be the robustness of the models. Given that a robust model can be constructed, there are a variety of techniques that can be used to enhance operator training, support, and evaluation such that we can improve our confidence in the operator's ability to cope with complex novel situations.

REFERENCES

1. J. de Kleer and J. S. Brown, "Foundations of Envisioning," *Proceedings of the AAAI National Conference on Artificial Itelligence,* 1982.
2. M. D. Williams and J. Hollan, "The Process of Retrieval from Very Long Term Memory," *Cognitive Psychology* **5,** 87-227 (1981).
3. R. R. Burton, J. S. Brown and J. de Kleer, "Pedagogical, Natural Language and Knowledge Engineering Techniques in SOPHIE I, II, and III," in *Intelligent Tutoring Systems,* D. Sleeman and J. S. Brown (Eds.), Academic Press, London (1982).
4. J. de Kleer and J. S. Brown, "Mental Models of Physical Mechanisms and Their Acquisition," in *Cognitive Skills and Their Acquisition,* J. A. Anderson (Ed.), Lawrence Erlbaum Associates, Hillsdale, N.J. (1981).
5. J. de Kleer and J. S. Brown, "Assumptions and Ambiguities in Mechanistic Mental Models," in *Mental Models,* D. Gentner and A. S. Stevens (Eds.), Lawrence Erlbaum Associates, Hillsdale, N.J. (1982).
6. K. D. Forbus, *Qualitative Process Theory,* Massachusetts Institute of Technology A.I. Memo No. 664 (1982).
7. K. D. Forbus and A. Stevens, *Using Qualitative Simulation to Generate Explanations,* Technical Note, Navy Personnel Research and Development Center (1981).
8. M. D. Williams, J. Hollan and A. Stevens, "Human Reasoning About a Simple Physical System," in *Mental Models,* D. Gentner and A. S. Stevens (Eds.), Lawrence Erlbaum Associates, Hillsdale, N.J. (1982).
9. C. Reiger and M. Grinberg, "The Declarative Representation and Procedural Simulation of Causality in Physical Mechanisms," *Proceedings of the Fifth International Joint Conference on Artificial Intelligence,* 250-255 (1977).

Paper 8

OUTLINE OF A MODEL OF HUMAN PROBLEM SOLVING IN FAILURE SITUATIONS*

William B. Rouse
Search Technology, Inc.
Norcross, Georgia

1. INTRODUCTION

Man-machine interaction has been a topic of formal study for well over 50 years. The earliest investigations focused on the environment as it affected the human operator's safety and ability to perform his job. Later investigations began to consider also the design of equipment in terms of identifying possible limitations and devising potential enhancements of operator performance. Much progress has been made in these areas, although a great deal remains to be accomplished.

More recently, the impact of automation has come to be of increasing importance. In aircraft, ships, process plants, transportation networks, and other large-scale systems, more and more control loops that were once closed manually are now automatically controlled. As a result, the human operator is becoming more of a monitor and supervisor of automation.[1]

The possibility of failures is the primary reason for having humans monitor automatically controlled processes. If hardware and software failures could not occur and if the automation were capable of handling all contingencies, then human operators would be unnecessary. However, failures and design limitations are quite possible and, therefore, a primary task of the human operator is to detect these events and deal with them appropriately. If the current trends continue, this task will come to dominate the human's responsibilities.[2]

It seems reasonable to make the general claim that the manual activities of the human operator will increasingly be supplanted by problem-solving activities. The objective of this paper is to outline an overall model of human problem solving in detecting and diagnosing system failures and compensating for them.

2. OUTLINE OF AN OVERALL MODEL

From a recent review of the literature,[3] it is obvious that considerable effort has been invested in the study of human problem solving in general, and in the study of human detection, diagnosis, and compensation for system failures in particular. However, most of the models focus on a single aspect of problem solving. Only a few of the models consider the full breadth and robustness of human problem solving behavior.[4–7] What is needed is a model that, at least conceptually, captures the whole of problem solving and, at the same

*Much of this paper is excerpted from W. B. Rouse, "Models of Human Problem Solving: Detection, Diagnosis, and Compensation for System Failures," *Automatica*, in press.

time, can be operationalized within specific task domains. This paper proposes such a model.

2.1. Pattern Recognition Orientation

A conclusion that surfaced repeatedly in the review of the literature[3] was that humans, if given a choice, would prefer to act as context-specific pattern recognizers rather than to attempt to calculate or optimize. Obviously, life would be difficult indeed if one had to constantly recalculate various things in order to make choices. Thus, human preference for pattern recognition is justifiable both scientifically and practically.

However, if the human does not recognize a pattern, a mode of problem solving other than context-specific pattern recognition must be employed. The alternative modes may be called heuristic, analytical, topographic, etc. A common characteristic of these modes is that the humans must go beyond the surface features of the problem. Since the focus of this paper is on system failures, this notion can be more precisely stated as the human must go beyond the system state and consider the system structure.

Figure 1 illustrates this fundamental concept. As can be seen, the human is assumed to have a clear preference for proceeding on the basis of state information. The use of structural information is definitely a less-preferred alternative. The mechanism depicted in Fig. 1, which is elaborated upon throughout the remainder of this paper, is proposed as the central and only mechanism necessary for an overall model.

2.2. Levels of Problem Solving

Considering the literature reviewed by the author,[3] as well as a variety of studies of human problem solving in aircraft, ships, and process plants carried out by the author and his colleagues,[8–15] it seems reasonable to conclude that problem solving occurs on several levels. Perhaps the most obvious example of multilevel problem solving is the coordination of compensation and diagnosis where the human must try to sustain system operations while also attempting to locate the source of the problem. The concept of multiple levels is, however, much more general than the idea of coordinating tasks.

It appears that three general levels of problem solving are needed to model human behavior: (1) recognition and classification, (2) planning, and (3) execution and monitoring. Recognition and classification involves detecting that a problem-solving situation exists and assigning it to a category. Planning is the process whereby the approach to solving a problem is determined, and execution and monitoring is the actual process of solving the problem.

Table 1 summarizes how the basic mechanism in Fig. 1 applies to the three levels of problem solving. At the highest level (i.e., recognition and classification), the human is assumed to identify the context and category of a problem. If he finds the state information to match a familiar frame,[5] problem solving proceeds on that basis. If an appropriate frame is not in the human's repertoire, structural information might provide clues to an analogy or be used to employ basic principles of, for example, the scientific method.

At the next level (planning), the human must decide how the problem will be attacked. Based on the state information, he may conclude that the problem-solving situation is familiar and the appropriate script[6] or standard procedure can be employed. If no script is

Table 1. Decisions and Responses for Three Levels of Problem Solving

Level	Process		
	Decision	Response	
		State-Oriented	Structure-Oriented
1. Recognition and classification	Frame available?	Invoke frame	Use analogy and/or basic principles
2. Planning	Script available?	Invoke script	Formulate plan
3. Execution and monitoring	Pattern familiar?	Apply appropriate S-rule	Apply appropriate T-rule

available, the human must use structural information to plan in terms of generating alternatives, imagining consequences, valuing consequences, and so on.[16]

Actual problem solving occurs at the lowest level (i.e., execution and monitoring) where scripts or plans are executed and monitored for success. Familiar patterns of state information may allow for the use of context-specific symptomatic rules (S-rules) that map directly from observation to hypothesis or action (e.g., if the engine will not crank, check the battery). If the pattern is not familiar, structural information may allow use of topographic rules (T-rules) for searching the structure of the problem (e.g., if a component's inputs are good and its outputs bad, the component has failed).

All of the responses noted in Table 1 (invoke frame, use analogy, etc.) invoke the same mechanism as shown in Fig. 1. This mechanism is recursively invoked until actions are produced and the problem is solved. Thus, in contrast to many of the models found in the literature,[3] the model outlined here is very simple, involving a single mechanism that is recursively employed for all aspects of problem solving.

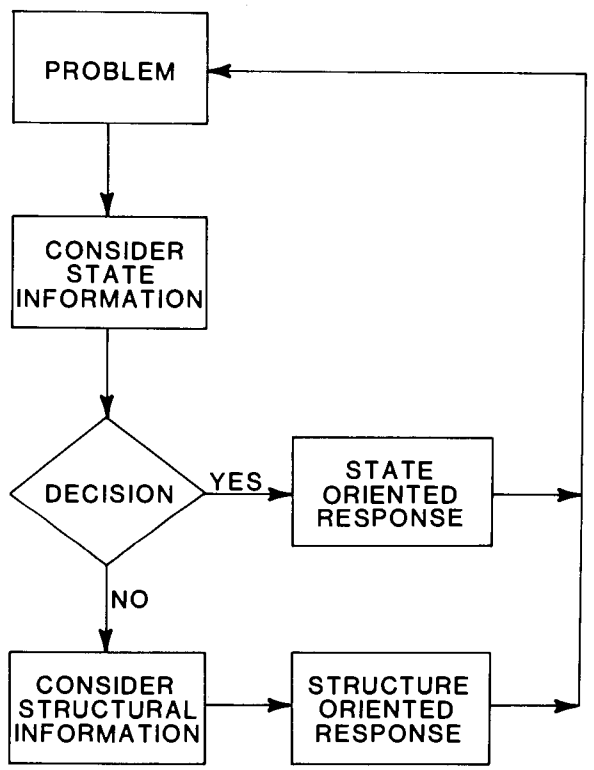


Fig. 1. Basic Mechanism of Proposed Model.

2.3. Hierarchical vs Heterarchical Approach

If one views problem solving from an operations research or management science perspective, one should hierarchically consider goals, objectives, attributes, alternatives, etc. This hierarchical approach has often been adopted by computer scientists when designing knowledge-based expert problem-solving systems.[17] The model proposed in this paper certainly could perform hierarchically by first invoking a frame or an analogy, then invoking a script or planning, and finally acting via S-rules and T-rules.

It has been argued, however, that human problem solving is heterarchical or opportunistic rather than hierarchical.[18] In other words, the human does not solve problems in a purely top-down or bottom-up manner. Instead, he appears to operate on all levels almost simultaneously.

The proposed model can produce this type of behavior if one assumes that the three decisions (i.e., frame available, script available, and pattern familiar) are constantly, but not necessarily consciously, being re-evaluated. Thus, for example, the model might be using T-rules to plan on the basis of an analogy and suddenly realize the applicability of a script. This could result in the preemption of planning and the rapid application of a sequence of S-rules that provide new information and result in a familiar frame being recognized, which, in turn, leads to a new script and so on.

If all three decisions are constantly being reevaluated, it is possible that conflicts will arise in terms of which decision should take precedence. Such conflicts might be resolved by giving more credence to closer matches (e.g., a very familiar pattern is more captivating than a somewhat familiar script). Another method of resolving conflict is to place more weight on alternatives that maintain the current direction of the problem solving. In other words, the model could incorporate the assumption that the human would like to avoid changes in frame, script, etc.

2.4. Behavior of the Model

Since the model outlined here represents a synthesis of the many concepts and models discussed in the literature,[3] it should not be surprising that it will produce the types of behavior noted in this literature. The strength of the model is that a very simple mechanism and method of organization can produce such an impressive range of behaviors. Of course, this possibility, from a slightly different perspective, has been investigated by others.[4]

A particularly interesting aspect of the model's behavior, as well as that of humans, is its potential for making errors. The model has two inherent possibilities for causing errors. The first relates to the model's recursive use of the basic mechanism in Fig. 1. As the model recursively invokes this mechanism, it needs a "stack" or some short-term memory for keeping track of where it is and how it got there. If short-term memory is limited, as it is in humans, the model may recurse its way into getting lost or pursuing tangents from which it never returns. To constrain this phenomenon, it is probably reasonable to assume that lower priority items in the stack are more likely to be lost first. For example, one is more likely to forget one's umbrella than to forget to go to work.

The second possibility for causing errors is the matching of irrelevant or inappropriate patterns. For example, the model, or a human, may be captured by an inappropriate but similar script or S-rule. As a result, the model may pursue an inappropriate path until it suddenly realizes, perhaps much too late to be able to recoup, that it has wandered far afield from where it thought it was headed.

The fact that the proposed model has inherent possibilities for making errors, particularly somewhat subtle errors, provides an interesting avenue for evaluating the model. Most models are evaluated in terms of their abilities to achieve the same levels of desired task performance as humans. A much stronger test would involve determining if the model deviates from desired performance in the same way and for the same reasons as humans. The proposed model can potentially be evaluated in this manner.

3. DISCUSSION AND CONCLUSIONS

Considering the relationships of the overall model outlined here to the models found in the review of the literature,[3] this conceptual model is, to a great extent, an outgrowth of earlier work by the author and his colleagues[19–24] and is also closely related to work by Rasmussen.[7,25–27] The frame, script, and S-rule aspects of the model are also fairly similar to concepts proposed by Newell and Simon.[4] The manner in which the model utilizes both state and structural information and the recursive use of the same basic mechanism on all levels of problem solving are perhaps the model's most unique characteristics. In this way, it is potentially capable of dealing with unfamiliar problem contexts via analogies, T-rules, etc.

The model's ability to operate almost simultaneously on several levels provides the potential for representing the coordination of compensation and diagnosis, an aspect of problem solving that few if any other models explicitly describe. Because of short-term memory limitations, the model also allows for the possibility of errors in coordination (e.g., focusing on one task to the exclusion of the other). What is not clear at this point, and is the topic of several current investigations, is the nature of S-rules, T-rules, scripts, and plans relative to coordinating compensation and diagnosis. This important topic deserves considerable study.

In this paper, an outline of an overall model of human problem solving has been proposed. This outline is, in a sense, a set of model specifications based on a synthesis of a wide range of concepts and models as well as a variety of experimental results. Currently, the software necessary to support this model is being designed. Initial evaluations of the model will focus on coordinating of compensation and diagnosis of failures in process control systems.

REFERENCES

1. T. B. Sheridan and G. Johannsen (Eds.), *Monitoring Behavior and Supervisory Control,* Plenum Press, New York (1976).
2. J. Rasmussen and W. B. Rouse (Eds.), *Human Detection and Diagnosis of System Failures,* Plenum Press, New York (1981).
3. W. B. Rouse, "Models of Human Problem Solving: Detection, Diagnosis, and Compensation for System Failures." *Automatica,* in press.
4. A. Newell and H. A. Simon, *Human Problem Solving,* Prentice-Hall, Englewood Cliffs, New Jersey (1972).
5. M. Minsky, "A Framework for Representing Knowledge," in *The Psychology of Computer Vision,* P. H. Winston (Ed.), McGraw-Hill, New York (1975).
6. R. C. Schank and R. P. Abelson, *Scripts, Plans, Goals, and Understanding,* Lawrence Erlbaum Associates, Hillsdale, New Jersey (1977).
7. J. Rasmussen, *On the Structure of Knowledge — A Morphology of Mental Modes in a Man-Machine Context,* Report No. M-2192, Riso National Laboratory, Roskilde, Denmark (November 1979).
8. G. Johannsen and W. B. Rouse, "A Study of the Planning Process of Aircraft Pilots in Emergency and Abnormal Situations," pp. 738-745 in *Proc. 1980 IEEE Intl. Conf. on Cybernetics and Society,* Boston (October 1980).
9. G. Johnannsen and W. B. Rouse, "Problem Solving Behavior of Pilots in Abnormal and Emergency Situations," pp. 142-150 in *Proc. First Ann. European Conf. on Human Decision Making and Manual Control,* Delft, The Netherlands (May 1981).
10. S. H. Rouse, W. B. Rouse and J. M. Hammer, "Design and Evaluation of an Onboard Computer-Based Information System for Aircraft," *IEEE Trans. Syst. Man. & Cybern.,* SMC-12 (1982).
11. W. B. Johnson and W. B. Rouse, "Analysis and Classification of Human Errors in Troubleshooting Live Aircraft Powerplants," *IEEE Trans. Syst. Man. & Cybern.,* SMC-12, 389-393 (1982).
12. W. B. Johnson and W. B. Rouse, "Training Maintenance Technicians for Troubleshooting: Two Experiments with Computer Simulations," *Hum. Fac.* **24** (1982).
13. J. M. van Eekhout and W. B. Rouse, "Human Errors in Detection, Diagnosis, and Compensation for Failures in the Engine Control Room of a Supertanker," *IEEE Trans. Syst. Man. & Cybern.,* SMC-11, 813-816 (1981).
14. W. B. Rouse and N. M. Morris, "Human Problem Solving in a Process Control Tank," pp. 218-226 in *Proc. First Ann. European Conf. of Human Decision Making and Manual Control,* Delft, The Netherlands (May 1981).
15. W. B. Rouse and N. M. Morris, "Studies of Human Problem Solving in a Process Control Task," pp. 479-487 in *Proc. 1981 IEEE Intl. Conf. on Cybernetics and Society,* Atlanta (October 1981).
16. G. Johannsen and W. B. Rouse, "Mathematical Concepts for Modeling Human Behavior in Complex and Man-Machine Systems," *Hum. Fac.* **21,** 733-747 (1979).
17. E. D. Sacerdoti, "Planning in a Hierarchy of Abstraction Spaces," *Art. Intell.* **5,** 115-135 (1974).
18. B. Hayes-Roth and F. Hayes-Roth, "A Cognitive Model of Planning," *Cog. Sci.* **3,** 275-310 (1979).
19. W. B. Rouse, "A Model of Human Decision Making in a Fault Diagnosis Task," *IEEE Trans. Syst. Man & Cybern.,* SMC-8, 357-361 (1978).
20. W. B. Rouse, "A Model of Human Decision Making in Fault Diagnosis Tasks That Include Feedback and Redundancy," *IEEE Trans. Syst. Man & Cybern.,* SMC-9, 237-241 (1979).
21. W. B. Rouse and S. H. Rouse, "Measures of Complexity of Fault Diagnosis Tasks," *IEEE Trans. Syst. Man & Cybern.,* SMC-9, 720-727 (1979).

22. W. B. Rouse, S. H. Rouse and S. J. Pellegrino, "A Rule-Based Model of Human Problem Solving Performance in Fault Diagnosis Tasks," *IEEE Trans. Syst. Man & Cybern.*, SMC-10, 366-376 (1980).

23. W. B. Rouse and R. M. Hunt, "A Fuzzy Rule-Based Model of Human Problem Solving in Fault Diagnosis Tasks," pp. XV 95-101 in *Proc. Eighth World Congress of IFAC,* Kyoto (August 1981).

24. R. M. Hunt and W. B. Rouse, "A Fuzzy Rule-Based Model of Human Problem Solving," *Proc. Am. Control Conf.,* Washington (June 1982).

25. J. Rasmussen and A. Jensen, "Mental Procedures in Real Life Tasks: A Case Study of Electronic Troubleshooting," *Ergonomics* **17,** 293-307 (1974).

26. J. Rasmussen, *Notes on Diagnostic Strategies in Process Plant Environment,* Report M-1983, Riso National Laboratory, Roskilde, Denmark (January 1978).

27. J. Rasmussen, "Models of Mental Strategies in Process Plant Diagnosis," pp. 241-258 in *Human Detection and Diagnosis of System Failures,* J. Rasmussen and W. B. Rouse (Eds.), Plenum Press, New York (1981).

Paper 9

MODELING HUMAN-COMPUTER INTERACTION

Thomas P. Moran
Xerox Palo Alto Research Center

(NOTE: The talk given by the author at the Workshop was based on a book that he has recently completed. Because of the demands on his time in writing the book, he agreed to participate in the Workshop with the proviso that he would not be required to prepare a paper for the proceedings. As a result, only an abstract of his presentation is included here.)

The main products of an applied cognitive psychology are cognitive performance models. For such models to be both applied and scientific suggests that they be of a certain form. In Chapter 1 of our book, *The Psychology of Human-Computer Interaction*,[1] S. Card, A. Newell, and I have laid out a view of "calculational" applied psychology models, of which one model, the Keystroke-Level Model, is a good example. The Keystroke-Level Model, described in detail elsewhere,[2] is a simple model for predicting one aspect of human performance: the time it takes an expert user to perform a given task on a given computer system. It is based on counting keystrokes and other low-level operations, including the user's mental preparations and the system's responses.

The Keystroke-Level Model does not exist in isolation, but rather is embedded in a theoretical model of cognitive skill (actually, a series of embedding models: the GOMS Model, the Problem Space Model, and the Model Human Processor). A psychological picture of cognitive skill can be sketched in which mentally compiled *methods* (or procedures) are the central units. Modeling the execution of methods, then, is the key to predicting expert cognitive skill performance. It is also useful for *calibrating* skilled performance (for example, to distinguish skill from problem solving). On the other hand, modeling the *complexity* of methods seems to be important for predicting *novice learning* of cognitive skills. A simple model of learning explains our observed correlation (between different systems) between expert performance and novice learning.

We view the Keystroke-Level Model as a *system design tool*. A more complicated and refined model could be formulated, thus increasing its accuracy, by relaxing some of its serious restrictions (for example, a model that predicts methods or that predicts errors). One of the great virtues of the Keystroke-Level Model is that it puts a lower bound on the effectiveness of new models, any one of which must do better to merit serious consideration.

REFERENCES

1. S. Card, T. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, New Jersey (1983).
2. S. K. Card, T. P. Moran, and A. Newell, "The Keystroke-Level Model for User Performance Time with Interactive Systems," *Communications of the ACM* 23, 396-410 (1980).

Paper 10

COGNITIVE MODELS EMBEDDED IN SYSTEM SIMULATION MODELS

Arthur I. Siegel and J. Jay Wolf
Applied Psychological Services, Inc.

1. INTRODUCTION

If we are to discuss and consider "cognitive models," we must first come to grips with two questions: (1) What is cognition? (2) What is a model? Presumably, the answers to these questions can provide a basis for defining a cognitive model. Accordingly, this paper first places these two questions into perspective. Then, cognitive models are set within the context of computer simulation models and a number of computer simulations of cognitive processes are described. Finally, pervasive issues are discussed vis-a-vis cognitive modeling in the computer simulation context.

2. DEFINITIONS AND CONSIDERATIONS

2.1. Cognition

Cognition has been traditionally defined in a somewhat diverse and broad manner. Guilford, who described the "structure of the intellect" as the result of extensive factor analytic studies, provided the following definition of cognition:

> Immediate discovery, awareness, rediscovery, or recognition of information in its various forms; comprehensive or understanding (Guilford and Hoepfner[1]).

The *Comprehensive Dictionary of Psychological and Psychoanalytic Terms* presents the following primary definition of cognition:

> 1. a generic term for any process whereby an organism becomes aware or obtains knowledge about an object ... It includes perceiving, recognizing, conceiving, judging, reasoning (English and English[2]).

Taft summarized the proceedings of a set of meetings on "cognition — theory, research, promise" held at the University of Kansas as follows:

> No one in this symposium on cognition has offered us a definition of this concept. We could do little better than refer to Martin Scheerer's definition in Lindzey's *Handbook of Social Psychology*. He says that cognition is the process of gaining information and understanding the world (Taft[3]).

From these definitions it seems that the term *cognition* can be applied to mental activity ranging from mere recognition to reasoning and understanding. In spite of this diversity, there is a common core throughout the three definitions. Cognition appears to be a generic term that describes the interaction of the human with that which is sensed, the processing of these sensory data (called "information"), and the transformation of the data into some meaningful form. It seems to include sensation, information processing, understanding, and reasoning. Accordingly, cognitive models can include almost everything and anything that would be included under the terms *perception* and *higher mental functions*, as well as the processes which link the two.

2.2. Models

This paper is concerned with computer simulation models. Computer simulation models have also been variously defined. Examples of current definitions are:

> A computer simulation is a logical-mathematical representation of a concept, system, or operation programmed for solution on a high-speed electronic computer (Martin[4]).
>
> A simulation model can be defined as an abstraction of some real system that can be used for purposes of prediction and control (Naylor et al.[5]).
>
> ...form a picture of (the system's) behavior by sampling from all the ways it might behave (Kendell[6]).
>
> The (simulation model) produces outputs resembling those observed in the real world, and inspires confidence that the real causal process has been accurately represented (Dutton and Starbuck[7]).
>
> A simulation model is an operator that generates a set of variables X, given a set of variables Y, such that: (*a*) the sets X and Y represent characteristics of the referent situation, (*b*) the set of X is indistinguishable with respect to an explicit criterion from the corresponding characteristics of the referent situation, and (*c*) the operator itself represents the causal process of the referent situation (Starbuck[8]).

The essence of these definitions appears to be that computer simulation models attempt to mimic or to represent some aspect of real life and that the simulation process is embodied in the form of a high-speed digital computer program.

Mere representation is not a sufficient goal in itself for a behavioral model. To be useful for more than descriptive purposes, the representation should be of such a nature that it allows extrinsic (situational/environmental/equipment) or intrinsic (individual and group) variables to be manipulated so as to determine the effects of the manipulation on some output measure or on some other variable which exerts an affect on the output.

Hence, we define a behavioral computer simulation model as a representation of behavior and behavioral influences implemented on a digital computer so as to allow control or prediction of an event, or event sets. This definition emphasizes both the structural and the utility aspects of a behavioral simulation model.

2.3. Models vs. Theories

While a model will often be based on theory, a model of human behavior or of some aspect of behavior is not, per se, a theory of behavior. The two should not be confused. Table 1 compares and contrasts models and theories. A model possesses a different purpose than a theory:

- Feigl[9] proposed that the purpose of a theory is to state functional relationships. While a model may include functional relationships within its structure, the goal of simulation modeling, as stated above, is to allow for prediction or control of the effects of manipulating extrinsic or intrinsic variables on an output measure.
- Chapanis[10] held that models are to be judged on the basis of utility while the test of a theory rests on validity. The "utility" of a model means the extent to which the output assists the decision maker in reaching a decision.
- Because they are fundamentally representational in nature, models are descriptive in content. Because the goal of theories is to state functional relationships, as a minimum, they must rely on some type of correlational content.
- Theories may postulate unknown or intervening variables of unknown or partially known dimensions. In contrast, each variable in a simulation model must be defined and be scalable, and its relationship to other variables must be clearly specified.
- Whereas models are highly quantified so that specific representations can be formulated with precision, in theories quantification is not essential.

Table 1. Comparison of Models and Theories

	Behavioral Model	Theory
Purpose	Predict or control behavioral implications — "What happens if"	Describe functional relationships — "How it works"
Criteria of acceptance	Utility or assistance in decision making	Validity
Content	Descriptive	Correlational
Variables	Measurable/scalable	May be hypothetical
Quantification	High	Optional
Generality	From full to limited	Full

2.4. Why Computer Simulation?

If any real-life situation is to be modeled, a rather complex situation is involved. This,in turn, dictates that the internal structure of a model which represents the situation contain numerous interactive internal variables. Implementation of such models has been made practical because of the advent of the high-speed digital computer. The computer program allows for a complex, interactive representation of the process which is modeled, and the machine allows the process to be exercised expeditiously and under a wide variety of conditions.

When cognitive models of real-life situations are considered, it is hard to conceive that the computer simulation approach can be avoided. For example, we have been analyzing maintenance activities in nuclear power plants. Some of the variables which can impact the cognitive processes of the maintenance personnel are:

Environmental Variables	**Personnel Variables**
Temperature	Ability
Acoustic noise	Motivation
Radiation	Supervision
Human factors situation	Stress tolerance
Time limits	Fatigue
Protective clothing	Training
Procedures	Number of persons
Spare parts availability	Cross training

If such a number and variety of variables is to be represented and exercised interactively, there seems to be little option other than computer simulation. The option is entirely practical because of:

- the exponential growth in the availability of computational power and reduced hardware costs;
- the wide acceptance of and significant growth of educational processes leading to today's computer scientist, which is related to
  - (1) the current acceptance of the computer field as a science rather than an art,
  - (2) the development of many theoretical underpinnings of the field (e.g., structured programs, correctness proofs),
  - (3) the explosive growth of the number of trained personnel in analysis and programming;
- the impact of on-line program development and maintenance, as well as on-line conversational program runs;
- the orders-of-magnitude reduction in the cost per bit of storage, the cost per bit per second of long-distance transmission, and the basic cost of computation per unit arithmetic/logical operation;
- the feasibility of transmitting large volumes of data over long distances;

- the improved computer system reliability via redundancy, error correction, fault tolerant operating systems, and component count reductions;
- the adoption of federal standards for programming languages, storage media, transmission interface protocols and the like;
- the drastic reduction in the volume of space required by computing equipment coupled with corresponding reductions in power (operating) costs.

Given that computer simulation is possible and practical, how good is it? Jenkins[11] has compared various cognitive models, including models based on the computer simulation approach (Siegel-Wolf simulation). His results, which support the computer simulation approach, are presented in Table 2. Similarly, both Meister[12] and Adams[13] have supported the approach in the noncognitive domain.

Table 2. Rating of Four Cognitive Models[a]

Criteria	Rasmussen S-O-R	Siegel-Wolf Simulation[b]	Sheridan Supervisory Control	Artificial Intelligence
1. Acceptance by the user	3	4	Not observed	3
2. Verisimilitude	5	4	4	3
3. Validity of input data	3	5	5	4
4. Documentation to support use	Not observed	5	Not observed	5
5. Model validation	1	4	1	3
6. Output at the level consistent with the objectives	5	4	5	5
7. Easy to use interface	4	4	Not observed	3
Total	21	30	16	26
Mean Rating	3.5	4.3	4.0	3.6

[a]From Jenkins, ref. 11; the higher the number, the better the rating.
[b]Computer simulation.

3. PRIOR COGNITIVE MODELS THROUGH COMPUTER SIMULATION

Over the years, Applied Psychological Services has been active in the modeling of human performance in complex, real-life equipment systems. Within a number of these computer simulation models, some type of cognitive modeling has been embedded. Our cognitive models are possibly best characterized as input-output models. They attempt to describe the probable output, given a variable input set. They do not purport to present the underlying physiological, biological, or chemical processes. Because of our interest in simulating real-life human performance and because of the effects of extrinsic and intrinsic variables on human performance, selected cognitive model elements have been embedded within larger models of total individual and group performance. As such, they have formed their function at a subroutine level within a total computer simulation model (i.e., program).

The following sections present examples of various recent cognitive simulation elements. Each of these is drawn from and constitutes a part of a total or larger simulation.

3.1. Decision Simulation

A decision module was developed[14] to simulate the operator's decision processes involved in an advanced electronic imagery system. The simulation logic follows from Simon and Newell's problem-solving theory[15,16] but expands on that conceptualization by considering task complexity, decision utility, operator ability, stress, and Bayesian concepts. The Simon-Newell theory describes the problem solver as stepping, node-by-node, through the problem space until a solution (i.e., final decision) is reached. The node stepping represents the decision process where each step involves data collection, analysis, and evaluation and the problem space represents the size and structure of the decision task.

The decision subroutine is based on decisions involving up to five decision alternatives (one and only one of which is correct) and six nodes (for a total of 11 states). The representation, as shown in Fig. 1, may be viewed as a hexagonal structure which may be initially entered at any node. From any node, 11 choices are possible; stepping to one of the other five nodes, remaining at the node, or stepping to a solution state. Stepping to a solution state completes the stepping process.

Decision time, within the simulation, is calculated by considering three variables: task complexity, stress on the simulated operator, and the number of steps to reach a solution. Decision correctness is determined by matching the solution reached (via the stepping process) to a given (input) correct solution. Within the step process, the step probability values, initially supplied as input, are affected by: (1) operator ability, (2) utility, and (3) a Bayesian process.

A top-level flow chart of the simulation process is presented as Fig. 2. After selecting the most proficient operator on the basis of his speed and proficiency, the step probabilities are first adjusted for operator ability (proficiency) and decision utility. Next (following resets), the decision implement step is taken. If a decision state is reached, the correctness of the decision is determined and the simulation proceeds to the decision time calculation sequence (boxes 9 through 12). If no decision state is entered, the input probability matrix is altered by a Bayesian process, and a probability ratio, required in the later complexity calculation, is obtained. Then, performance of the next step is simulated, and the decision—no decision check is again made. This process continues until a decision state is entered, or until the number of steps has reached the prespecified limit.

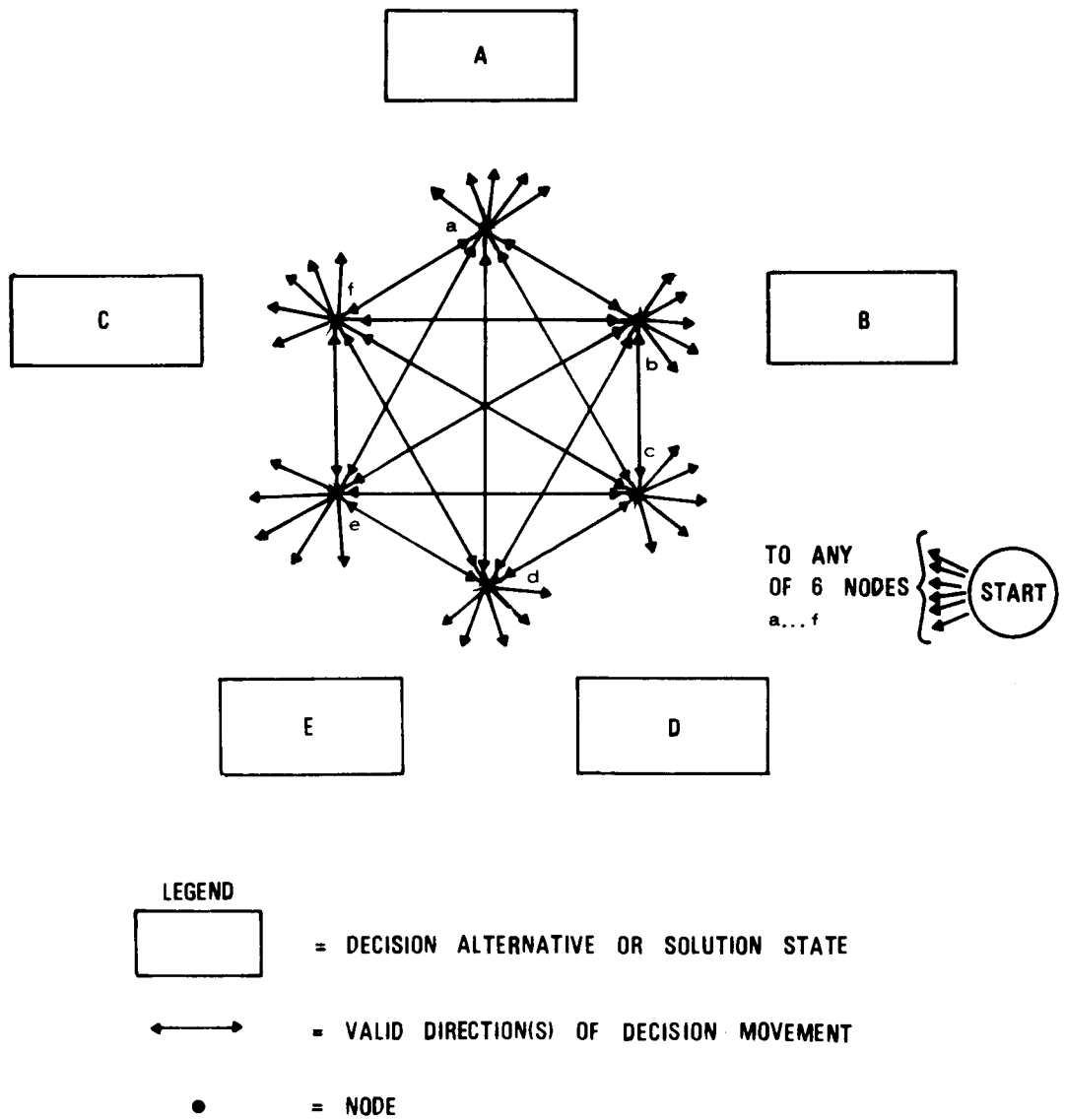


Fig. 1. Partial Representation of a Five-Solution, Six-Node Problem Representation.

The operator ability adjustment causes an increase (higher proficiency operator) or decrease (lower proficiency operator) in the probability of a correct choice and in the time to reach a solution. The utility function is included because the best solution is not necessarily the solution with the highest probability. Utility is calculated on the basis of: (1) the importance of the decision on up to any three preselected mission goals, and (2) the effects of each course of action on the goals. The resultant utility values are used to adjust the input probability of moving toward each goal.

Monte Carlo processes are employed to complete the stepping process and the final decision is compared with the correct decision to determine decision correctness. If the

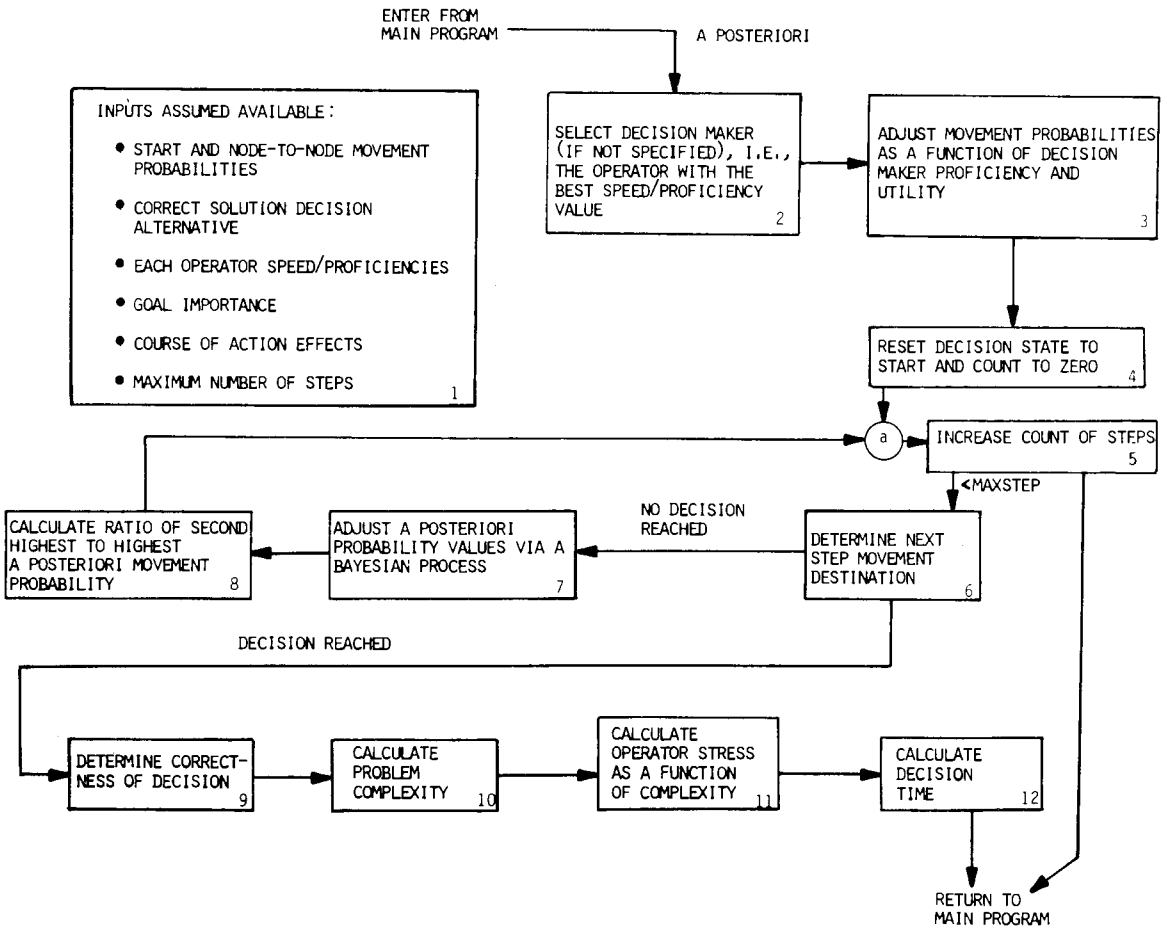


Fig. 2. Overall Flow Logic of Decision Subroutine.

completion of a step does not cause the simulated decision maker to enter a decision state, at the conclusion of the step the input probabilities are further altered on the bases of the Bayesian logic. This represents changes in the simulated decision maker's state on the basis of the information gained during the step.

3.2. Recognition Model

Recognition of objects represents a cognitive activity which is important to almost every real-life situation. The present recognition model[14] was developed as a part of a total simulation of the actions of operators of an advanced radar system. The subroutine does not include the search, detection, or tracking activities since these are simulated in other subroutines. Only the object identification (classification) function is considered here.

The goal here was determining the capability of an operator to recognize one of several predefined "primitive" objects. The model was based on the thinking of Reed,[17] who indicated the relative ascendancy of "features analysis" over "template matching" because features analysis rests on structuring the stimulus pattern in terms of primitive-characterized strokes to correspond to environmental invariants.

The recognition simulation first calls for determining an initial classification probability on the basis of the number of strokes in the object under consideration. This first approximation is based on the logic of Deese[18] that "stimulus complexity and ease of identification appear to be related as a U-shaped function." That is to say, features can neither be too few nor too many and neither too similar nor too dissimilar to expedite the recognition of pictorial images. The function employed in the simulation (Fig. 3) was developed on the basis of the data presented by Levine and Eldredge[19] who asked experienced photo interpreters to classify a set of already detected targets.

This initial classification probability is then successively moderated (degraded) by four functions: (1) deviation from primitive in number of strokes, (2) deviation from primitive in angle of main stroke, (3) deviation from primitive in length of main stroke, and (4) deviation from primitive in curvature of main curved line.

The first of these functions, the deviation of the subject figure from the primitive in terms of number of strokes (Fig. 4), is based on the conjecture that classification accuracy will degrade as the detail in a representation deviates from learned or anticipated detail. In a sense, a figure which is the same as a learned figure can be considered to be a structured stimulus. The influence of structure in form perception was acknowledged by Aiken and Brown,[20] who asserted that "Thus it appears that the O, in dealing with relatively weakly structured stimulus configurations, will impose a consistent structure of his own generation, to organize the stimulus configuration" (p. 282). Arnoult[21] concluded similarly. Empirical image evaluative studies[22–23] in which the amount of structure (learned detail) was varied also support this conjecture.

The function for deviation from primitive in angle of main stroke rests on the argument that a shaped rotated out of the orientation in which it was learned is less readily recognized. As stated by Hake[24] (p. 151):

> Actually, the fact that recognition of forms and comparative judgments of forms deteriorates when forms are rotated with respect to the observer has been known for a long time.

Hake reported a number of studies which support the contention of modified association value for slanted (rotated) figures. In the simulation, the results of Arnoult[21] were relied on to provide the basic function. Arnoult varied 10 nonsense shapes over eight angular positions and asked subjects for "same" or "different" judgments when the rotated shape was compared to a standard. Arnoult's data were smoothed and rescaled. However, the shape of the Arnoult curve was closely approximated. The resultant function is shown in Fig. 5.

The function for deviation from primitive in length of main stroke (Fig. 6) presupposes that subjective response alternatives increase as the target size decreases so that the match between stimulus and its mental template in memory is more difficult. That is, guessing must be instituted until a correct "match" between current stimulus-input and memory-pattern has been achieved when very small targets are involved. The constraints that preexist in the real environment need to be discovered through successive trial-and-error episodes until stimulus features correspond to predefined pattern features sufficiently that the intended recognition response is obtained. Bruns, Bittner, and Stevenson[25] in a dynamic television identification task concluded that "Target effects were primarily related

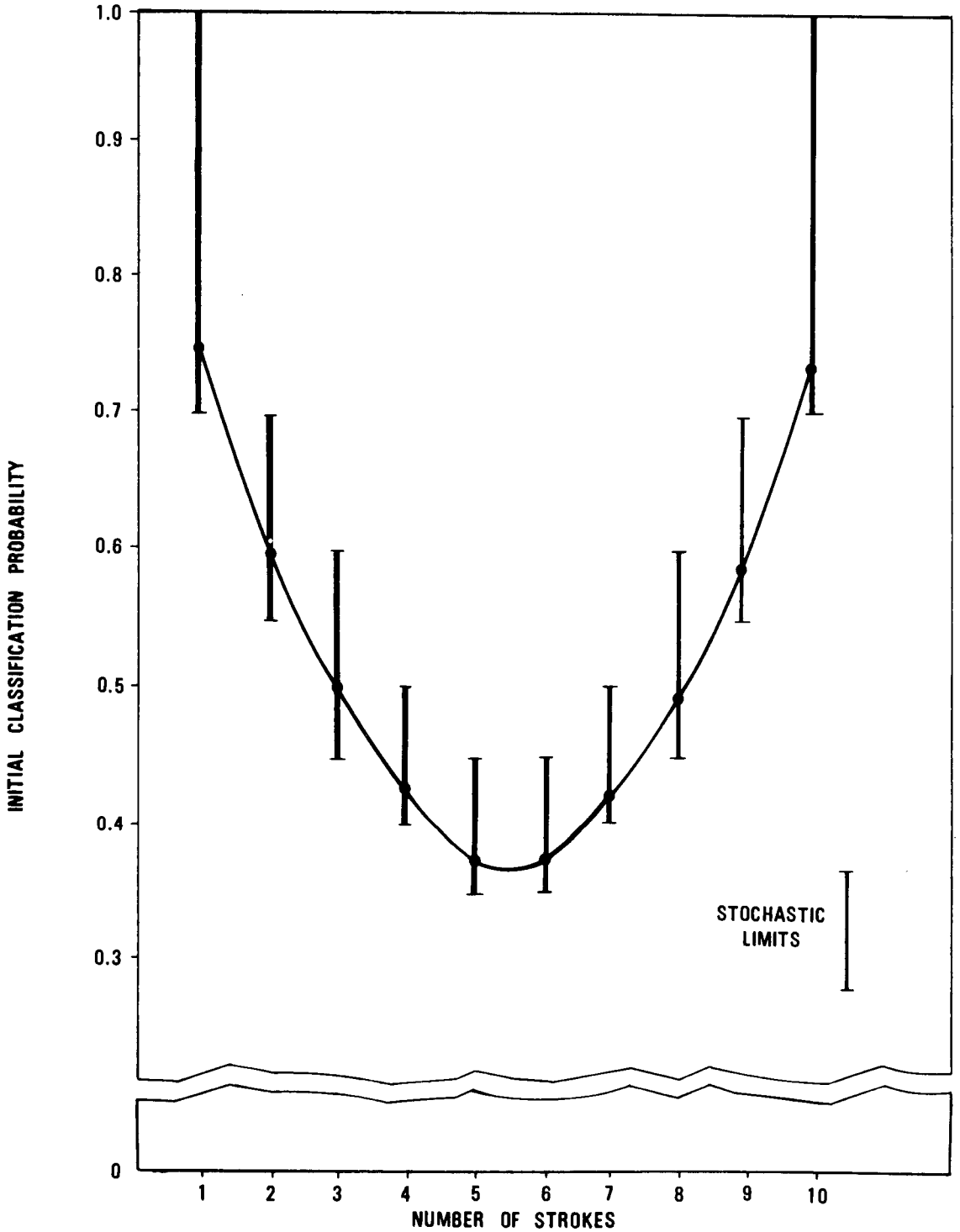


Fig. 3. Initial Classification Probability as a Function of the Number of Strokes in Generated Target.

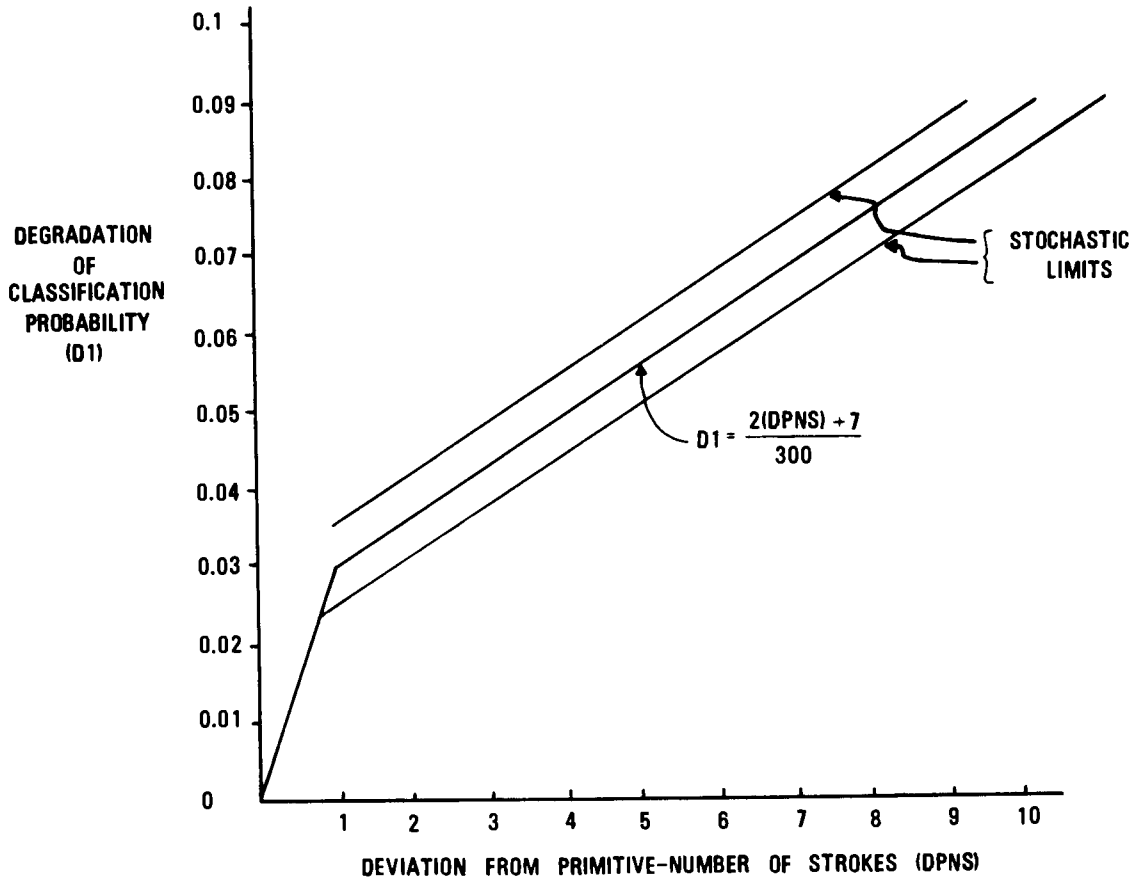


Fig. 4. Degradation of Classification Probability as a Function of Deviation in Number of Strokes.

to target size expressed either as target area or target diagonal." Their data, employed to derive the function shown in Fig. 6, indicated an almost linear relationship ($r = 0.98$) between slant-range identification and target diagonal, with a slant-range degradation ratio of about 6:1 for "small" as compared with "larger" targets.

The fourth function, for deviation from primitive in curvature (Fig. 7), was similarly based on the contention that deviation from an anticipated image will decrease classification accuracy. No studies were identified that reported data relative to the effect of curvature on recognition accuracy. Accordingly, the function was derived on the basis of the best professional judgment of the developers of the simulation.

The recognition module also considers context, proficiency of the operator, and stress. The context effect was based on the work of Miller, Heise, and Lichten,[26] who reported a range of context effects of about 30 percent. This effect is simulated by drawing and applying a random number that is equiprobable in the range 0 to 0.30. Operator proficiency and the stress effect are simulated in much the same manner as for other models we have developed.[27]

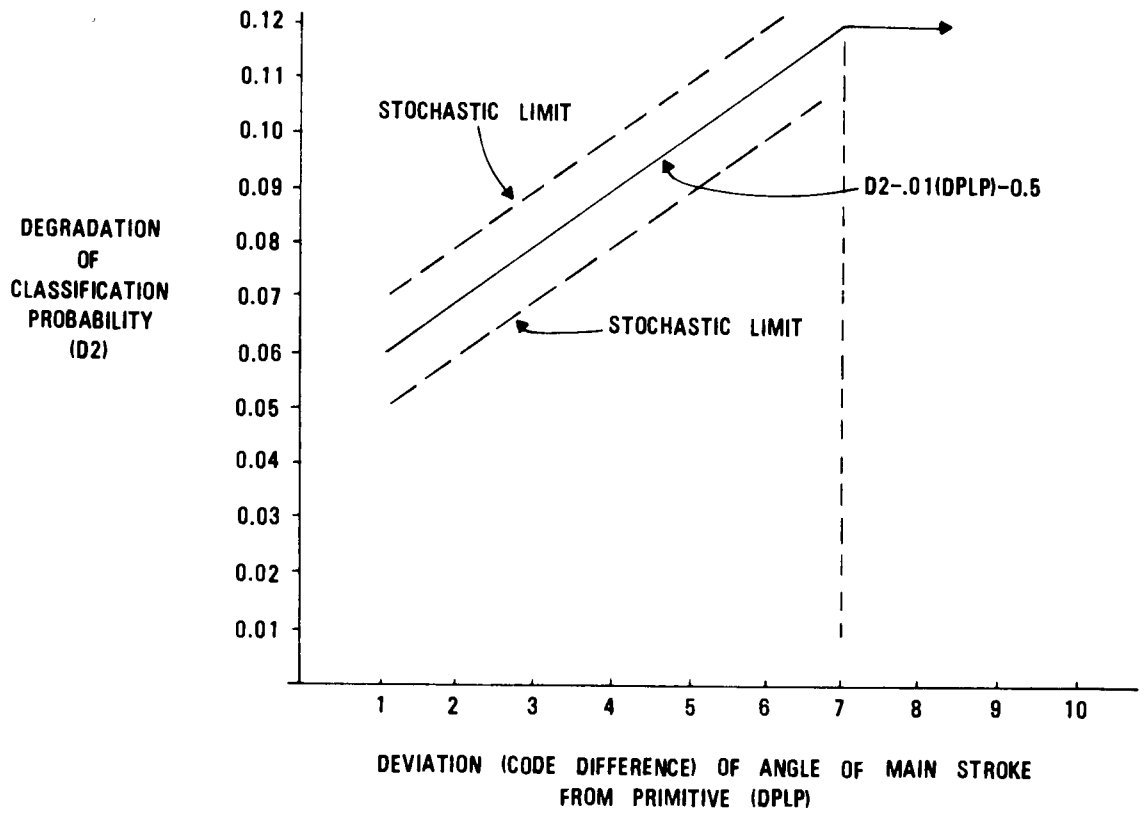


Fig. 5. Degradation of Classification Probability as a Function of Deviation from Angle of Main Stroke.

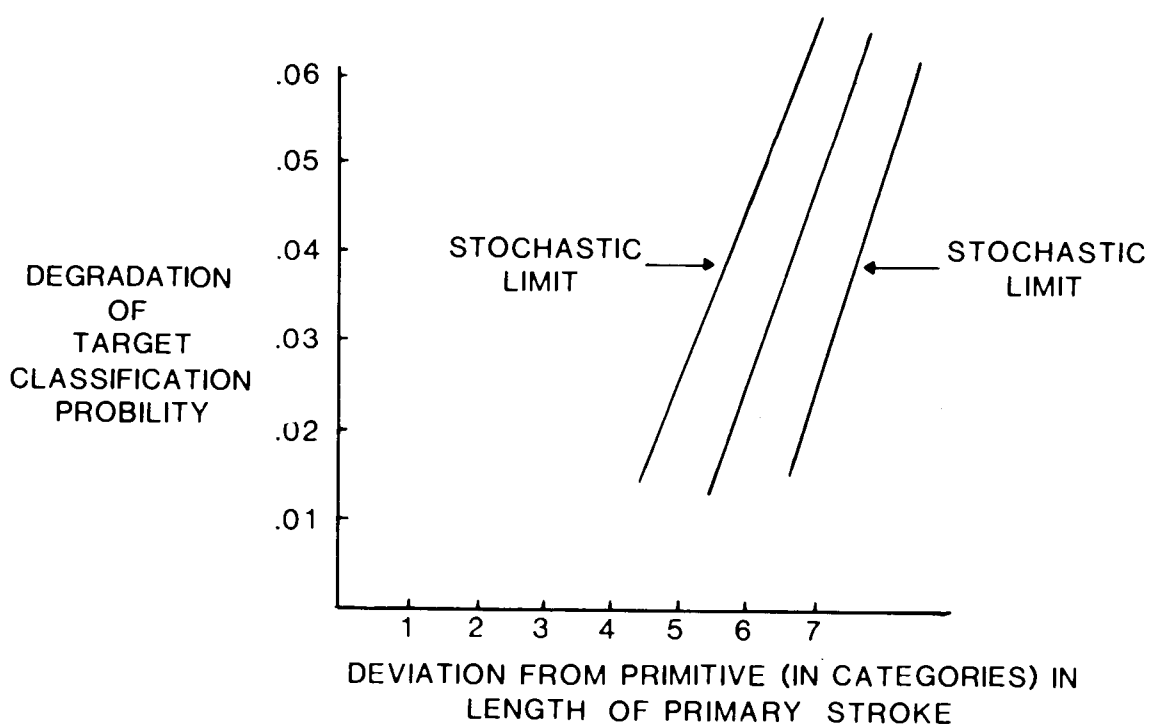


Fig. 6. Degradation of Classification Probability as a Function of Deviation from Length of Main Stroke.

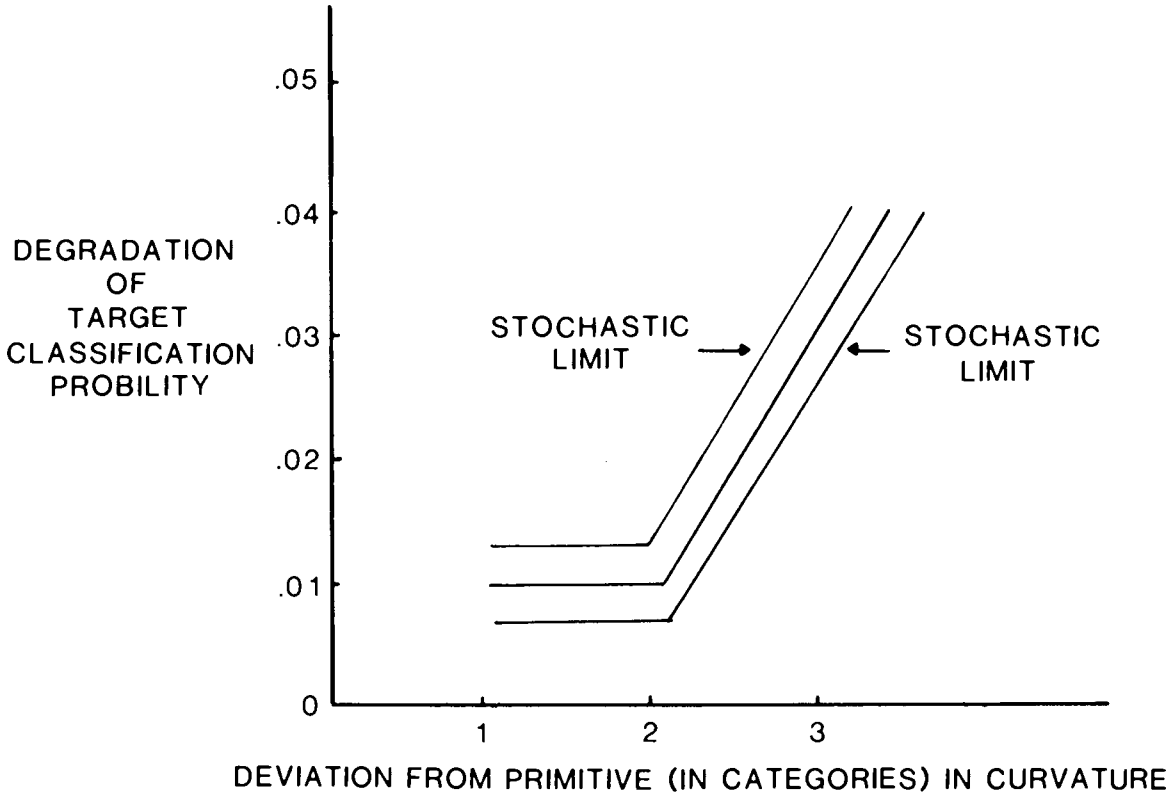


Fig. 7. Degradation of Classification Probability as a Function of Deviation in Curvature of Main Curved Lines.

The computer logic flow for the total recognition model is presented as Fig. 8. Here RY represents a random number in the range of 0 to 1, STR(M) represents the stress threshold, M specifies an operator, and F(M) represents operator proficiency.

3.3. General Nuclear Power Plant Decision Model

Finally, it is possible that the decision making for maintenance tasks in nuclear power plants might be simulated through a trait approach. As the result of a set of job analyses,[28,29] it is known that work of maintenance personnel in nuclear power plants is dependent on at least three primary mental abilities: (1) cognition, (2) problem solving, and (3) memory. The level of each of these abilities required for performing various instrument and control technician functions is shown in Fig. 9. In Fig. 9, the required levels are scaled from 1 to 7, where a value of 1 indicates "no" requirement for the ability when performing the function and a value of 7 indicates "very much" of a requirement for the ability.

If one knows the required ability level for a function, it seems that the cognitive processes can be modeled quite directly. Again, such a model would be best embedded within a total performance model if a real-life situation is involved.

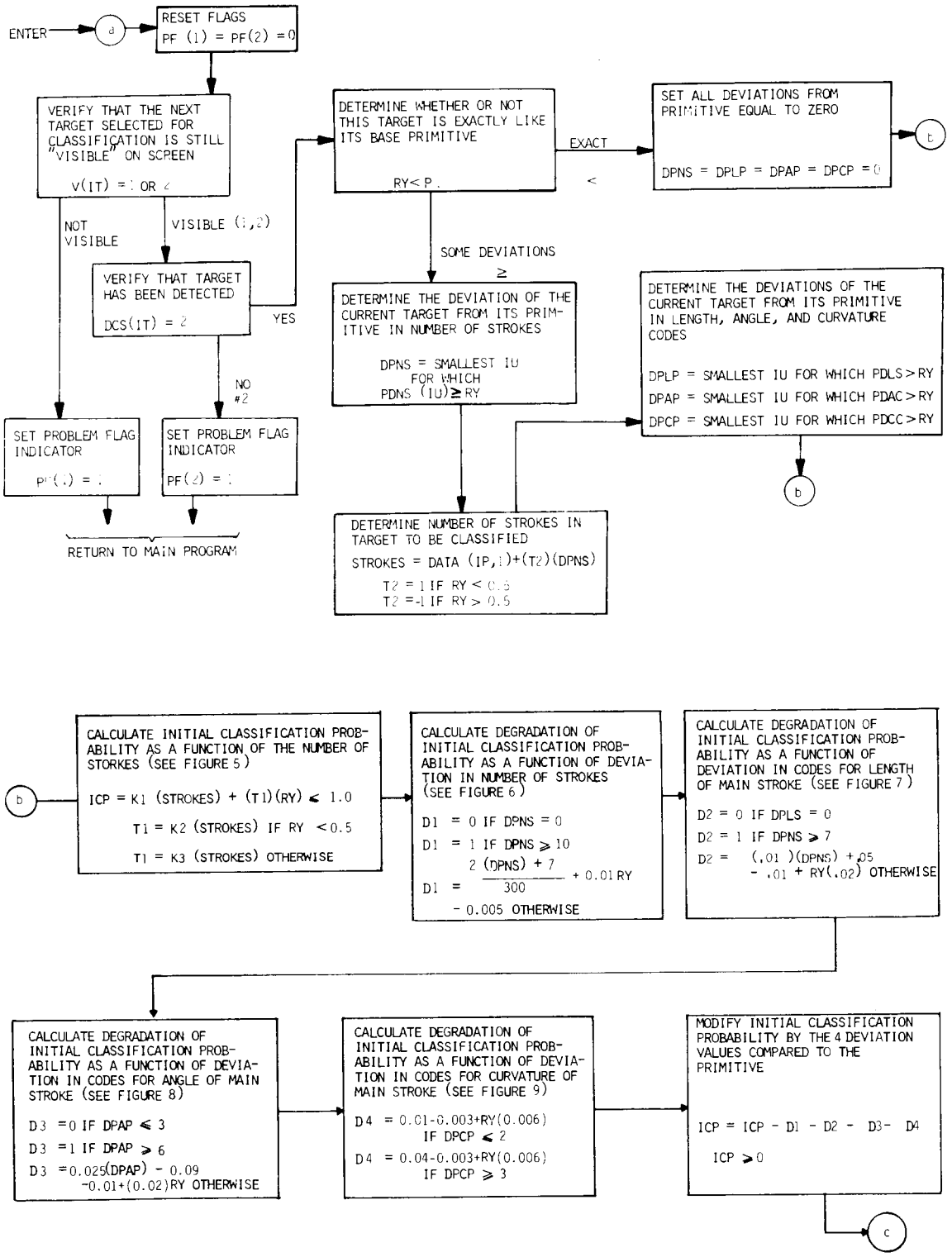


Fig. 8. Detailed Subroutine Flow Logic.

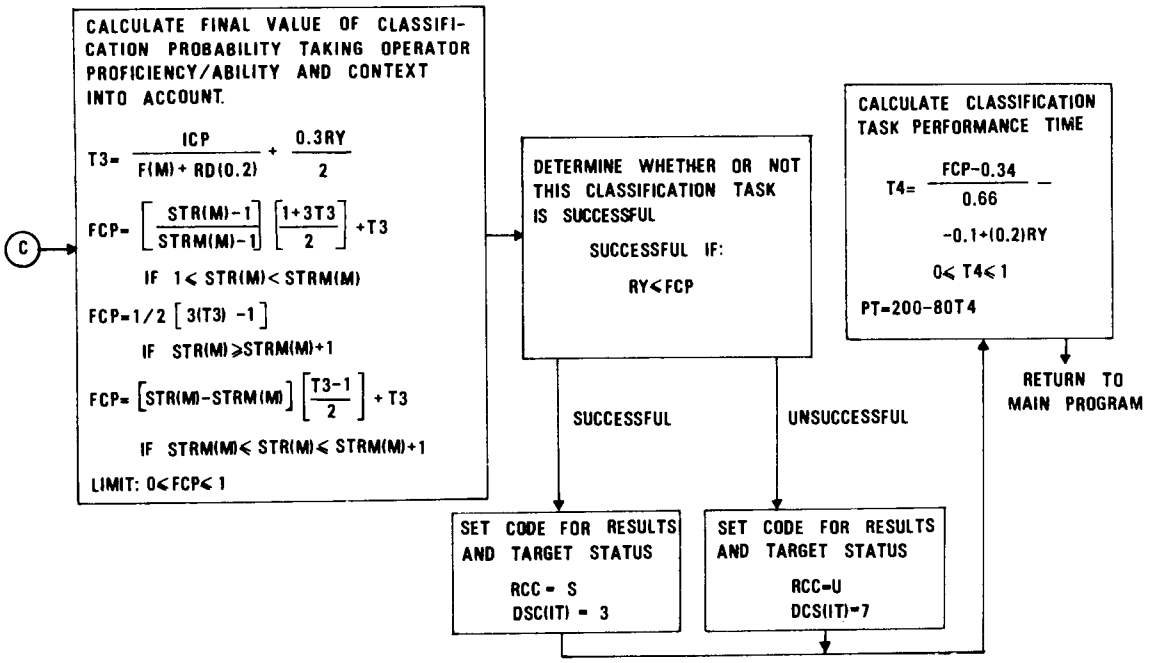


Fig. 8. Detailed Subroutine Flow Logic (continued).

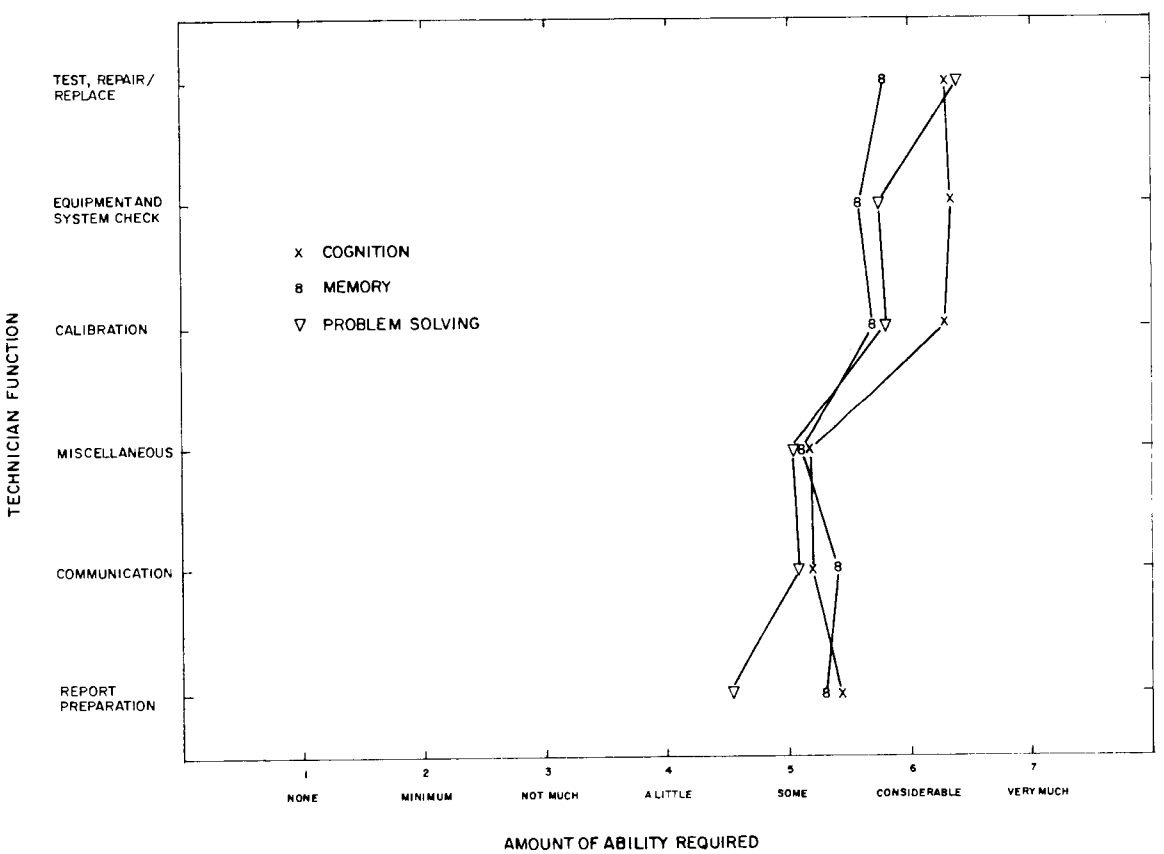


Fig. 9. Mental Ability Requirements for Instrument and Control Technician Functions.

The model for deriving the probability of a correct decision is drawn from prior psychometric work in test construction and item analysis. The concept of the selected logic is based on the model presented by Rasch[30] for item analysis. Rasch pointed out that the probability of a correct response to a test item depends on both the examinee's ability and the difficulty of the item. The relationship is best described by the normal ogive function:

$$P = \frac{e^x}{1+e^x} = \text{probability of a correct response},$$

where x is the difference between the test question difficulty and the examinee's ability. The equation holds that the difference between required ability and actual ability is the key to success probability determination. The Rasch approach has been demonstrated to be robust despite violations of its assumptions.[31]

To model the decision function using this concept requires estimates of decision difficulty and of the operator's ability. Data relative to decision difficulty can be extrapolated from Fig. 9, where the "ability required" data can be constructed to represent "decision difficulty level" for each function. Assume that a computer simulation model "knows" (from parameter inputs) the level on each of the three mental abilities of the operator(s) who must perform the function. The probability that the operator will successfully make the decision is determined by comparing the "actual" (simulated) ability levels of the operator with the levels required for function performance. The comparison is weighted by the logarithm of the operator's ability plus the load (i.e., required ability):

$$D = (\text{ability - load}) \log (\text{ability} + \text{load}) .$$

The weighting is based on a logarithmic logic because of: (1) the logarithmic nature of many psychological functions,[32] and (2) recent information about the exponential nature of the distribution of errors in nuclear power plants.[33]

When a decision maker's ability exactly matches the function load, the resulting difference is zero. Negative values indicate function requirements in excess of the operator's ability. Positive values indicate that the decision maker's ability is greater than that required by the task. These difference values are employed in the logistic model to provide the success probability for the function.

If one assumes that the probability of a correct performance of a function or a decision is 0.90 when requirements and ability match, the curve for the probability of a correct decision which emerges is that shown in Fig. 10. This is accomplished by an assignment for $x = x + 2.2$ in the ogive function for P above. Using this curve, the probability for all combinations of required abilities and actual operator abilities can be found. These are shown in Table 3. The computer logic is shown in Fig. 11.

4. DISCUSSION

The prior discussion attempted to set into perspective cognitive modeling within the context of computer simulation. To this end, three examples of cognitive models that have been or could be implemented within full computer simulation performance models were

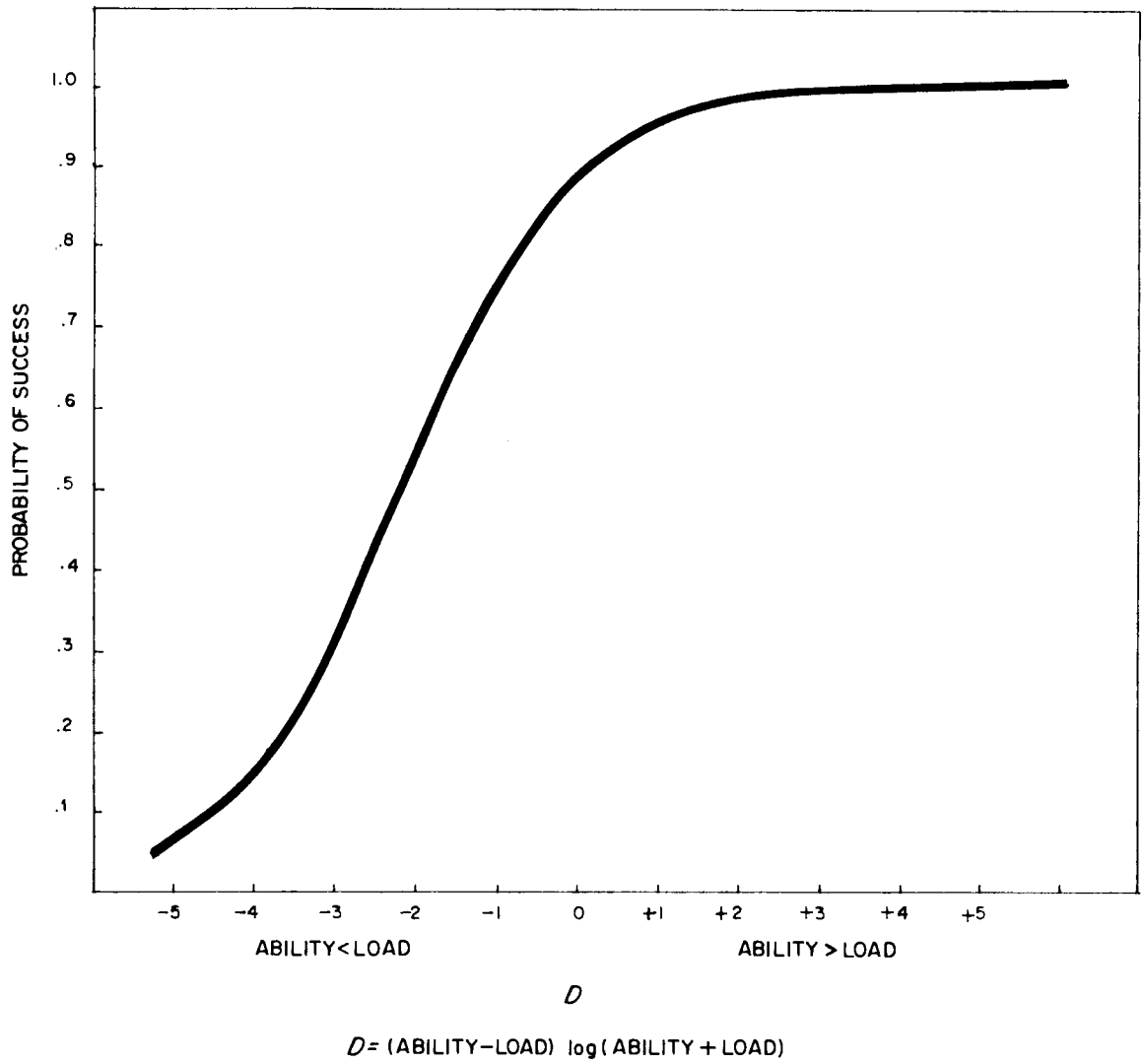


Fig. 10. Probability of Success as a Function of the Difference Between Subtask Load and Operator Ability.

presented. Cognition is viewed as one aspect of human behavioral simulation. Certainly, any total model must consider not only the representation of cognition but also the influence of other variables on the cognitive process. Examples of these other variables for nuclear power plant operation and maintenance were presented earlier. Because of the interactive complexity of these variables among themselves and with cognition, it seems that the computer simulation approach is the most promising for performance simulation. Certainly, cognitive subroutines should be included within such total performance models. Total performance models are within the state of the art, as is the incorporation of cognitive subroutines within such models. Cognition does not occur in isolation. It occurs in a context. For this reason, any model which considers cognition in isolation from a total situation or which contains no cognition elements is apt to be barren or insensitive.

Table 3. Success Probability for Different Combinations of Subtask Loading and Operator Ability (OA)

Subtask Loading	Success Probability						
	OA = 1	OA = 2	OA = 3	OA = 4	OA = 5	OA = 6	OA = 7
1	0.900	0.936	0.968	0.987	0.995	0.998	1.000
2	0.849	0.900	0.948	0.977	0.991	0.997	0.999
3	0.730	0.818	0.900	0.955	0.982	0.994	0.998
4	0.526	0.656	0.795	0.900	0.959	0.985	0.995
5	0.286	0.417	0.597	0.777	0.900	0.962	0.987
6	0.117	0.196	0.340	0.550	0.761	0.900	0.965
7	0.038	0.071	0.142	0.284	0.510	0.748	0.900

Computer simulation and behavioral modeling are dependent on the availability of acceptable theories for the various psychological and psychosocial elements. Behavioral models, including cognitive models, will prosper or suffer in accordance with the adequacy of the theories on which they are based. Given adequate theories, models can be designed to represent the theories. And, given adequate input data, the models can be validated and implemented. The gaps in the available behavioral theory are probably greater than the gaps in the computer modeling technology. This should not be construed to suggest that such modeling should not take place. We have shown by example that, at least with respect to cognition, adequate theories exist upon which to base simulation. It seems most prudent, therefore, to make the best use of those theories and data we have, while at the same time attempting to improve our theories and our data bases.

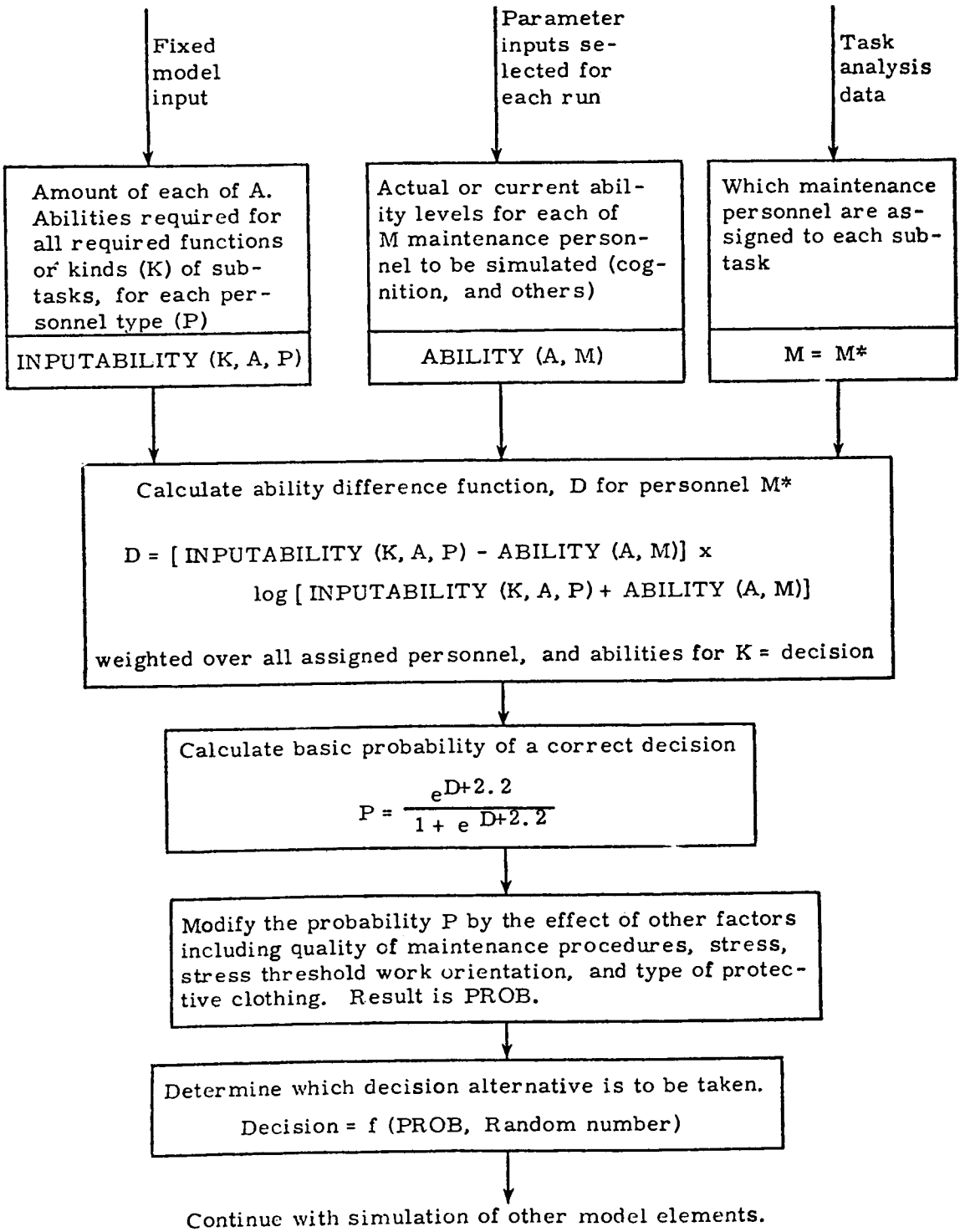


Fig. 11. Computer Logic for Simulating Decisions.

REFERENCES

1. J. P. Guilford and R. Hoepfner, *The Analysis of Intelligence*, McGraw-Hill, New York (1971).
2. H. B. English and A. C. English, *A Comprehensive Dictionary of Psychological and Psychoanalytic Terms*, Longmans Green, New York (1958).
3. R. Taft, "Summation," in *Cognition: Theory, Research, Promise*, C. Scheerer (Ed.), Harper Row, New York (1962).
4. F. F. Martin, *Computer Modeling and Simulation*, Wiley, New York (1968).
5. T. H. Naylor, J. L. Balintfy, D. S. Burdick, and K. Chu, *Computer Techniques*, Wiley, New York (1966).
6. M. G. Kendall, "Introduction to Model Building and Its Problems," in *Mathematical Model Building in Economics and Industry*, Hofner, New York (1968).
7. J. M. Dutton and W. H. Starbuck, "Finding Charlie's Run-Time Estimator," in *Computer Simulation of Human Behavior*, J. M. Dutton and W. H. Starbuck (Eds.), Wiley, New York (1971).
8. W. H. Starbuck, "Testing Case-Descriptive Models," in *Computer Simulation of Human Behavior*, J. M. Dutton and W. H. Starbuck (Eds.), Wiley, New York (1971).
9. H. Feigl, *Readings in Philosophical Analysis*, Appleton-Century-Crofts, New York (1949).
10. A. Chapanis, "Men, Machines, and Models," *American Psychologist*, **16**, 113-31 (1961).
11. J. P. Jenkins, *Cognitive Models for Complex Man-Machine Systems Application* (draft) (1982).
12. D. Meister, "Letter to the Editor," *Human Factors Society Bulletin*, Santa Monica (June 1982).
13. J. A. Adams, "Letter to the Editor," *Human Factors Society Bulletin*, Santa Monica (June 1982).
14. A. I. Siegel, J. J. Wolf, and A. R. Williams, *Computer Simulation of Human Performance in Electronic Processed Imagery Systems*, Applied Psychological Services, Wayne, PA (1978).
15. A. Newell and H. A. Simon, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, NJ (1972).
16. H. A. Simon and A. Newell, "Human Problem Solving: The State of the Theory in 1970," *American Psychologist* **26**, 145-59 (1971).
17. S. K. Reed, *Psychological Processes in Pattern Recognition*, Academic Press, New York (1973).
18. L. Zusne, *Visual Perception of Form*, Academic Press, New York (1970).
19. J. M. Levine and D. Eldredge, *Effects of Ancillary Information Upon Photointerpreter Performance*, Army Research Institute for the Behavioral and Social Science, Arlington, VA (1974).
20. L. S. Aiken and D. R. Brown, "A Feature Utilization Analysis of the Perception of Pattern Class Structure," *Perception and Psychophysics* **9**, 279-83 (1971).
21. M. D. Arnoult, "Prediction of Perceptual Responses from Structured Characteristics of the Stimulus," *Perceptual and Motor Skills* **11**, 261-68 (1960).
22. T. Coluccio and J. Wasielewski, *Electro-Optical Image Enhancement*, RADC-TR-70-35, Griffin Air Force Base, Rome, NY (1970).
23. S. McLeod, *Photointerpreter Performance Studies*, RADC-TDR-64-326, Griffin Air Force Base, Rome, NY (1964).
24. H. H. Hake, "Form Discrimination and the Invariance of Form," in *Pattern Recognition*, L. Uhr (Ed.), Wiley, New York (1966).
25. R. A. Bruns, A. C. Bittner, andnR. C. Stevenson, *Effects of Target Size, Target Contrast, Viewing Distance on Dynamic Televisual Target Detection and Identification*, TP-72-24, Naval Missile Center, Point Magu, CA (1972).
26. G. M. Miller, G. A. Heise, and W. Lichten, "The Intelligibility of Speech as a Function of the Context of Test Materials," *Journal of Experimental Psychology* **41**, 329-45 (1951).

27. A. I. Siegel and J. J. Wolf, *Man-Machine Simulation Models: Psychosocial and Performance Interaction*, Wiley, New York (1969).

28. A. I. Siegel, W. D. Bartter, and F. F. Kopstein, *Job Analysis for Maintenance Mechanics in Nuclear Power Plants*, Applied Psychological Services, Wayne, PA (1981).

29. W. D. Bartter, A. I. Siegel, and P. J. Federman, *Job Analysis for Instrument and Control Technicians in Nuclear Power Plants*, Applied Psychological Services, Wayne, PA (1982).

30. F. M. Lord and M. R. Novick, *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, MA (1968).

31. R. Forsyth, U. Saisangjam, and J. Gilmer, "Some Empirical Results Related to the Robustness of the Rasch Model," *APM* **5**, 175-86 (1981).

32. R. S. Woodworth and H. Schlosberg, *Experimental Psychology*, Holt, Rinehart, and Winston, New York (1954).

33. D. W. Joos, Z. A. Sabri, and A. A. Husseiny, "Analysis of Gross Error Rates in Operation of Commercial Nuclear Power Stations," *Nuclear Engineering and Design* **52**, 265-300 (1979).

Paper 11

A RULE-BASED APPROACH TO COGNITIVE MODELING OF REAL-TIME DECISION MAKING

Perry W. Thorndyke
Perceptronics, Inc.
Menlo Park, California 94025

1. INTRODUCTION

Recent developments in the fields of cognitive science and artificial intelligence have made possible the creation of a new class of models of complex human behavior. These models, referred to as either "expert" or "knowledge-based" systems, describe the high-level cognitive processing undertaken by a skilled human to perform a complex, largely mental, task. Expert systems have been developed to provide simulations of skilled performance of a variety of tasks. These include problems of data interpretation, system monitoring and fault isolation, prediction, planning, diagnosis, and design.[1] In general, such systems strive to produce prescriptive (error-free) behavior, rather than model descriptively the typical human's errorful behavior. However, some research has sought to develop descriptive models of human behavior using the same theoretical frameworks adopted by expert-systems builders (see, for example, articles in ref. 2).

This paper presents an overview of this theoretical framework and modeling approach, and indicates the applicability of such models to the development of a model of control room operators in a nuclear power plant. Such a model could serve several beneficial functions in plant design, licensing, and operation. First, the model could support the development of a task analysis of the operator's functions. Such an analysis would be useful in the design and development of SPDS* displays or other procedures aids. Second, the model could be used in a variety of training functions — to assist in the development of performance criteria, to model typical operator deficiencies and errors, or to embed in an automated instructional system with a computer-based surrogate instructor. Third, the model could be used as a surrogate operator in system experiments. Such experiments would utilize a computer-based simulation of the operator in conjunction with a plant simulation to evaluate the impact of new systems on plant operations. In such similar experiments, the system configuration would be treated as an independent (i.e., manipulated) variable, the operator would be a controlled variable, and the overall performance of the system would be the dependent (that is, measured) variable. Fourth, the model could be used to evaluate proposed operator procedures. In this role, the operator model would become the independent variable in system experiments, with operator behavior governed by the alternative sets of proposed procedures. Fifth, the model could be used to conduct sensitivity analyses on the impact of operator performance on plant operations. For example, the analyst might systematically vary the behavior of the simulated operator to examine the effects of different skill levels, error types, or attentional failures on plant performance. Finally, the model might be used to design task allocations and organizations for multiperson teams of control room personnel.

*Safety Parameter Display System.

These intended uses of a cognitive model of the control room operator place constraints on the design of the model. As in other models and simulations of cognitive problem-solving behavior, modeling the control room operator requires a realistic simulation of the skills, actions, hunches, procedures, and heuristics (that is, rules of thumb) of the human problem solver. Achieving realism in the simulation entails producing a functionally accurate and "reasonable" model of the operator's behavior. In order for the analyst to understand the reasons underlying the production of a particular behavior by the simulation, it should model the intermediate states of the operator's memory and attentional field prior to the generation of an action. This requires the model to represent and the simulation to provide continuously available output of the operator's working memory at each point during the simulation. Such output of the simulation should be intelligible — that is, the system should be capable of providing rationales for produced behaviors. The capability of tracing simulation behaviors to their sources allows the analyst to both understand the dynamics of the operator's performance and to identify inadequacies or deficiencies in the model. Finally, the simulation embodying the cognitive model should be flexible and modifiable to allow easy incremental revision and experimentation with the governing rules of behavior.

In previous work on expert systems, two techniques have been successfully applied to the development of realistic, intelligible, and flexible simulations. First, researchers have modeled cognitive behavior as a collection of discrete, interacting processes operating at multiple levels of abstraction. The individual processes have unique triggering conditions, operations they perform, and outputs they produce. Second, English-like programming languages have provided the software tools for implementing complex cognitive models in comprehensible, modifiable codes. These languages feature the use of IF-THEN production rules to represent the detailed operation in each component cognitive process. This multilevel, rule-based approach has been applied to modeling such activities as planning, tactical decision making, and air traffic control. Like power plant monitoring and control, these activities will require time-stressed situation assessment, plan formulation, plan execution and monitoring, and dynamic replanning.

The remainder of this paper will address four topics. First, the tasks of the control room operator will be viewed as an example *model-based situation assessment and planning* (SA&P).[3] Second, a general framework for modeling cognitive performance on SA&P tasks will be presented. This framework assumes that multiple cooperating "expert" processes share common resources to achieve the task objectives. Third, we will describe the advantages of rule-based systems for implementing the cognitive model. Fourth, we will briefly review previous and current research projects in the area of rule-based modeling of SA&P.

2. MODEL-BASED SITUATION ASSESSMENT AND PLANNING

Like many other cognitive tasks that require attention and response to conditions in the external environment, the operator's functions comprise largely monitoring systems and indicators, diagnosing problems when they occur, and developing and executing plans to maintain desired states. Taken together, these tasks constitute examples of *model-based situation assessment and planning.* A general form of this model is shown in Fig. 1. The "model" comprises a set of concepts, facts, rules, procedures, beliefs, and heuristics in the

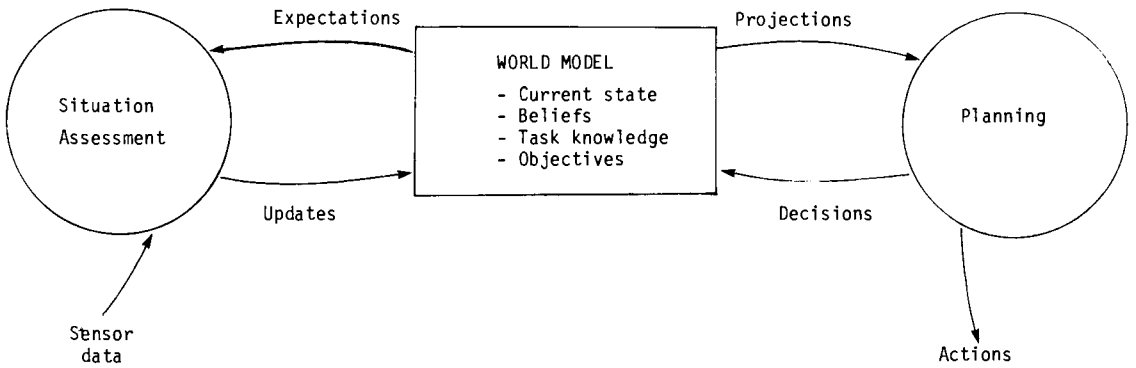


Fig. 1. Model-Based Situation Assessment and Planning.

head of the operator. This model constitutes the knowledge base the operator uses to carry out the assigned functions.

During situation assessment, the operator receives sensor data or information communicated from other personnel and evaluates it against his current model of the external world. That model supports the development of expectations about what observations to anticipate — expectations that may or may not be met by the actual data. As the external data change, the operator must update his mental model to reflect the new state of the environment. As the environment changes, the operator must determine which of the new conditions invite or necessitate responsive actions. In addition, his assessment will involve the generation of hypotheses concerning the causes of the new data he is receiving.

During planning, the operator must select one or more goals to be achieved, based on the current situation assessment, the operator's model of the state of the environment, and projections about how the environment will change over time. The operator must then synthesize one (or more) plans for achieving the objectives and refine them through the selection of appropriate actions. If more than one plan exists, the operator must then evaluate the probable outcomes of plan execution, evaluate the costs and benefits, and select the best option.

During plan execution, the operator performs the actions specified by the selected plan and monitors the resulting state of the environment, replanning as necessary to compensate for surprising events or failed plans. Any or all of these functions may require communication with other individuals or systems — either to provide status reports or to coordinate plans and actions with other personnel.

Numerous applications of this model to real-world tasks suggest the robustness of this general framework. In the military arena, real-time situation assessment of battle conditions requires analysis of sensor data, a model of force characteristics and tactics, and the identification of threats and opportunities. Planning entails directing or redirecting forces to meet the identified mission objectives. In air traffic control, the controller monitors radar displays and voice communications to detect conflicts and clearance requests.[4] Planning requires the rerouting of aircraft to satisfy safety constraints and minimize delay. Another application, motivated by chemists from Oak Ridge National Laboratory, involves

the identification and location of hazardous chemical spills. The situation assessment problem requires hypothesizing possible spill sources and identities, while the planning problem concerns what to measure and where to perform the measurements.[5]

3. THE ORGANIZATION OF THE COGNITIVE MODEL

To model the organization of the various activities required for time-stressed SA&P, we have found it useful to adopt the "cooperating experts" paradigm. This paradigm, first proposed as a model of speech understanding,[6] had been used to model planning,[7] air fleet control,[4] and tactical decision making.[8] It assumes that a set of independent processing modules cooperate and communicate via a common world model, or data base. Figure 2 illustrates the particular version of the cooperating experts paradigm used to simulate the real-time control of aircraft.[4] Each ellipse represents a functional specialist with its own scope of concern, goals, and procedures for accomplishing those goals. The arrows between the ellipses indicate data or results transferred to or from the world model, around which the specialists are organized. Each specialist uses the world model to support its activities, and each alters the contents of the model as part of its actions.

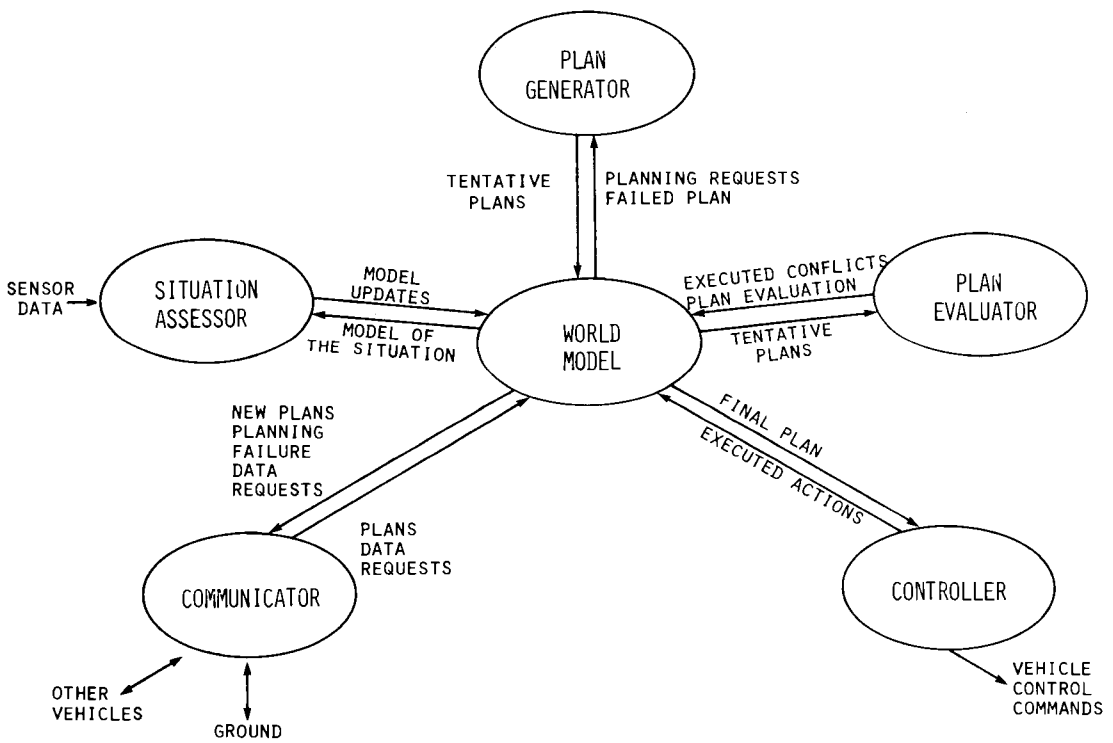


Fig. 2. Architecture of the Model for Air Fleet Control.

The specialists adopted for the cognitive models of air fleet control and tactical decision making include situation assessment, plan generation, plan evaluation, control, and communication. The situation assessment specialist interprets sensor data, recognizes altered states of the world, and determines conditions that require the human to take action. The plan generation specialist sets goals and designs a configuration of actions to achieve the goal. The plan evaluation specialist determines the uncertainties and projected outcomes associated with each plan provided by the plan generator, and selects the best of the available plans. The control specialist performs the actions specified by the selected plan and monitors plan execution. The communication specialist reports status to and coordinates actions with other individuals participating in the problem-solving process.

Since only one process may be active at a time, a scheduling process must control the allocation of processing control to the various specialists. Thus, three features characterize models organized according to the cooperating experts paradigm: the number and type of specialists, the structure and content of the world model, and the control regime for scheduling execution of the specialists. The remainder of this section considers these three features in more detail.

3.1 Processing Specialists

The expertise contained in each processing module is represented by a set of rules with the form IF (condition) THEN (action). Each rule contains a number of predicates that must be satisfied by the data objects in the world model and their relationships. If the tests are satisfied, then the action associated with the rule may be performed. The predicates include both actions performed by the operator, new facts or beliefs, or constraints on his future behavior. For example, the following rule might be part of the situation assessment specialist for a power plant operator:

```
IF PZR pressure is falling
      AND PZR level is falling
      AND PRT temperature is normal
      AND PRT pressure is normal
      AND CNTMT/AUX alarm is on
THEN create an emergency
      AND assert that the type of the emergency is
        loss-of-coolant-accident with probability .8
      AND let the source of the emergency be unknown.
```

Other rules in the situation assessment module would recognize other types of plant malfunctions or reason about the sources of the problems.

Newly asserted information is recorded in the world model for inspection and use by all of the processing specialists. For example, the plan generation module would contain rules to set and achieve goals, given the current assessment. One such rule might be:

```
IF there is an emergency whose type is loss-of-coolant-accident
THEN assert that stop-coolant-loss is a goal whose status is unattained
      AND whose posting-time is (the current time)
      AND let the criticality of the goal be high.
```

3.2 Structure of the World Model

The world model corresponds to the working and long-term memory of the modeled human. It comprises a relational data base of concepts, attributes of the concepts, and relations among the concepts. For example, the following description of one particular concept might reside in the world model:

Emergency

ID	= Event28
Type	= Loss-of-coolant-accident
Time	= 0947
Date	= 072679
Source	= Unknown

Depending on the particular application, the knowledge contained in the world model may be organized in a variety of ways. For example, in B. Hayes-Roth's model of planning,[7] knowledge is partitioned among five conceptual planes, or knowledge bases: the world knowledge plane, the plan plane, the plan abstraction plane, the meta-plan plane, and the execution plane. Facts in the world knowledge plane record data and observations about the world, as in the example above. Facts in the plan abstraction plane contain knowledge of the desired attributes of the plan, without specifying actual planned actions (for example, attend to the most critical emergency first). Knowledge in the meta-plan plane describes how to approach the problem, what methods to apply, and criteria for generating prospective plans. These types of knowledge all support the development of actual plans to be executed. The executive plan contains knowledge useful in deciding how to schedule the execution of the various specialists. This knowledge may include requests for activation from the various specialists, information about priorities, or knowledge about the expected results to be obtained from each specialist. Knowledge on the plan plane records elements of the developing plan and actions the planner expects to perform. In this and other planning models, these plan elements include decisions at multiple levels of abstraction. At the highest level, the planner may specify the goals to be achieved by the plan. At the next level of abstraction, the planner may specify the tactics, design, or approach by which the objectives will be achieved. At the next lower level of abstraction, the planner may indicate the sequence of procedures to be used to achieve the goal. At lower levels of abstraction, the planner may specify the individual actions and operations required to carry out the selected procedures.

3.3 Control Regime

Specialists monitor the state of the world model and, when appropriate, post requests for activation and processing time. This knowledge is used by an executive specialist that schedules the activation of other specialists based on these requests and a set of scheduling heuristics. These heuristics take into account a procedural model of the operator's task (that is, what tasks are prerequisites to what other tasks), priorities, emergencies, the current focus (that is, activity), and the expected utility of scheduling various activities.

This design leads to two features of the model that correspond to characteristics of human processors in time-stressed decision-making situations. First, the human has limited attentional and processing resources to allocate to the multiple tasks and the glut of data he must assimilate. Thus, the executive specialist must determine the most cost-effective use of their limited resources. Second, the decision maker does not engage in a strictly sequential performance of the various activities he must perform. That is, he does not merely assess the situation, develop and select a plan, and then execute the plan in strict succession. Rather, the assessment of the situation is an ongoing process, whose partial results contribute the development and evaluation of the plan at hand. As new actions are performed and new data arrive, plans may be modified, discarded, or re-evaluated. Thus, time-stressed decision-making performance is best viewed as a set of asynchronous processes whose performance is mediated by a common memory (the world model) and whose execution is coordinated by an executive scheduling process.

4. RULE-BASED SIMULATIONS OF COGNITIVE BEHAVIOR

Traditional simulations of complex systems are designed to be used in "batch" mode. After the user specifies the initial values of the simulation parameters, the program runs to completion, producing the simulation results (often an event history). The user then must analyze this output to recover the results of interest.

Such simulations have a variety of shortcomings. First, they are often unintelligible, comprising thousands of lines of, for example, FORTRAN code. In such systems, critical assumptions and designs governing the performance of the model are buried in difficult to decipher program code. Therefore, these simulations often lack credibility, because they are unable to justify or display the underlying reasons a particular behavior was produced. Furthermore, such systems are typically monolithic, inflexible, and difficult to modify.

The use of rule-based programming languages to engineer simulations offers the opportunity to eliminate these shortcomings. By representing the system's behavior in rules such as those illustrated above, the modeler provides a comprehensible, English-like encoding of the expertise used by the simulation to produce its behavior. These rules may be readily generated and examined by both researchers, subject matter experts, and individuals interested in the knowledge used by the simulation. Since the rules are modular, it is relatively simple to modify portions of the model and incrementally refine portions of the simulation without disrupting the integrity of the remaining rules. Such capabilities make it possible for the subject matter expert to play a more prominent role in developing and experimenting with the system than is possible with traditional simulators. Recently developed languages for building rule-based systems include ROSS,[9] ROSIE,[10] EMYCIN,[11] OPS[12] and AGE.[13]

The rule-based formalism provides the flexibility to represent many types of knowledge. The rules can readily represent heuristics and nondeterministic events and outcomes, which are essential to any realistic model of complex behavior. In addition, rules provide a convenient language for representing routine and standard procedures, many of which are expressed in IF-THEN form.

Several of the rule-based programming languages provide facilities for examining interactively and in detail the behavior of the simulation. Users may inquire why a particular behavior was generated, and receive a trace of the rules that were executed in the simulation that led to the action of interest. By altering specific rules or heuristics, the user can conduct experiments or perform sensitivity analyses to determine the consequences of adding or deleting rules, conditions, or values.

Computer simulations embodying rule-based models can serve both training and test/evaluation functions. In a training function, the simulation can serve as a surrogate instructor, generating responses to problem scenarios and contrasting them to the responses selected by the trainee. Alternatively, the system might simulate the behavior of another individual with whom the trainee interacts during task performance. For example, Perceptronics is currently developing a rule-based simulation of tactical decision making for use as an automated opponent in naval training games. The simulation will develop offensive and defensive strategies and tactics for use in simulated naval battles, and will direct forces against a human player. A key feature of the system is the use of English-like rules to encode the knowledge used by the situation assessment, tactics selection, and planning modules of the system. By using modular rules to encode the rules of behavior, the instructor will be able to manipulate the skill and behavior exhibited by the opponent by selectively editing or deleting particular rules.

In the test/evaluation mode, rule-based systems can simulate operator or system performance in a variety of experiments. For example, the ROSS (Rule-Oriented Simulation System) developed at The Rand Corporation simulates strategic air battles.[9] The system users can postulate alternative sensor or weapons systems, force ratios, or tactical options by editing the rules and knowledge base the system uses. These manipulations can then be used to perform strategic assessments by running and examining the results of the battle simulation.

A similar methodology could be applied to the evaluation of power plant system performance, in which operator behavior was simulated by a rule-based system. An interface between this system and an existing or another rule-based system of plant equipment operation would allow the evaluation of the combined human-machine system. Experiments could simulate plant performance and assess operator responses under a variety of assumptions about operator skill, workload, error rates and types, and response speed. These assumptions could be varied by modification of the rules used by the simulation. Alternatively, the experiments could vary simulated plant conditions to determine likely operator behaviors and overall performance of the combined system.

5. SUMMARY

Cognitive models of situation assessment and planning appear to offer promising modeling tools for the development of detailed models of power plant operators. When combined with the existing technology offered by artificial intelligence and rule-based systems, such models can result in flexible simulations for modeling, analysis, and evaluation. The research approach for achieving this capability has five major steps:

1. Design and build the cognitive model.
2. Conduct experimental investigations to refine and validate the model.
3. Apply the model to the analysis of individual differences.
4. Apply the model to the analysis of errors and skill deficiencies.
5. Infer principles for systems design and operation (for example, selection, training, plant design and procedures, automated aids, etc.).

ACKNOWLEDGEMENTS

Major contributions to the ideas presented in this paper have been made over the past five years by Barbara Hayes-Roth, Frederick Hayes-Roth, and Phil Klahr.

REFERENCES

1. M. Stefik, J. Aikens, R. Balzer, J. Benoit, L. Birnbaum, F. Hayes-Roth and E. Sacerdoti, "The Organization of Expert Systems," *Artificial Intelligence* **18**, 135-173, 1982.
2. D. Waterman and F. Hayes-Roth (Eds.), *Pattern-Directed Inference Systems,* New York: Academic Press (1979).
3. F. Hayes-Roth, *Artificial Intelligence for Systems Management,* P-6573, The Rand Corporation, Santa Monica, CA (1980).
4. P. Thorndyke, D. McArthur and S. Cammarata, "AUTOPILOT: A Distributed Planner for Air Fleet Control," pp. 171-177 in *Proceedings of the Seventh International Joint Conference in Artificial Intelligence* (1981).
5. F. Hayes-Roth, D. Waterman and D. Lenat (Eds.), *Building Expert Systems,* Boston: Addison-Wesley (1983).
6. L. Erman, F. Hayes-Roth, V. Lesser and R. Reddy, "The Hearsay-II Speech Understanding System: Integrating Knowledge to Reduce Uncertainty," *Computing Surveys,* 213-252 (1980).
7. B. Hayes-Roth, *Human Planning Processes,* R-2607-ONR, The Rand Corporation, Santa Monica, CA (1980).
8. P. Thorndyke, *Intelligent Opponent Simulation for Tactical Decisionmaking Training,* P-82-127, Perceptronics, Woodland Hills, CA (1982).
9. P. Klahr and W. Faught, "Knowledge-Based Simulation," *Proceedings of the First Annual Conference of the American Association of Artificial Intelligence* (1980).
10. F. Hayes-Roth, D. Gorlin, S. Rosenschein, H. Sowizral and D. Waterman, *Rationale and Motivation for ROSIE,* N-1648-ARPA, The Rand Corporation, Santa Monica, CA (1981).
11. W. van Melle, *A Domain Independent System that Aids in Constructing Consultation Programs,* Report No. STAN-CS-80-820, Computer Science Department, Stanford University (1980).
12. C. Forgy and J. McDermott, "OPS, a Domain-Independent Production System Language," pp. 933-939 in *Proceedings of the Fifth International Joint Conference in Artificial Intelligence* (1977).
13. P. Nii and N. Aiello, "AGE (Attempt to Generalize): A Knowledge-Based Program for Building Knowledge-Based Programs," pp. 645-655 in *Proceedings of the Sixth International Joint Conference in Artificial Intelligence* (1979).

REPORTS OF WORKING GROUPS

Working Group No. 1

SURVEY OF DEVELOPMENT OF MODELS AS APPLIED TO NUCLEAR PLANT OPERATORS (I)

W. B. Rouse, Chairman

C. E. Feehrer
M. G. Johnson
J. Fragola
T. P. Moran
N. Moray
W. R. Nelson
J. Rasmussen
A. I. Siegel
P. W. Thorndyke

Working Group No. 1 attempted to catalog a variety of existing modeling methodologies with regard to several attributes. First, the degree to which various models could reasonably be used to describe different aspects of human behavior was addressed. The pursuit of this issue was structured in terms of one of Jens Rasmussen's conceptualizations as shown in Fig. 1. Based on this figure, the group decided to concentrate on the behaviors above the wavy line.

Table 1 summarizes the attributes of the models or methodologies discussed. As is evident, there was a strong emphasis on models that have been applied to some engineering systems and/or models that are obviously applicable in this way. As a result, many psychological concepts and theories are not included. This was not meant to slight their importance. Indeed, they are essential ingredients for the successful application of the methods that are tabulated.

From this perspective, it should be stressed that these methods are very broad and can be specialized to a variety of applications, incorporating a variety of psychological and engineering concepts, laws, etc., in the process. This leads to a succinct conclusion: While the cognitive modeling of the nuclear power plant operator is a difficult problem, there are a variety of robust and well-tested modelling methodologies available.

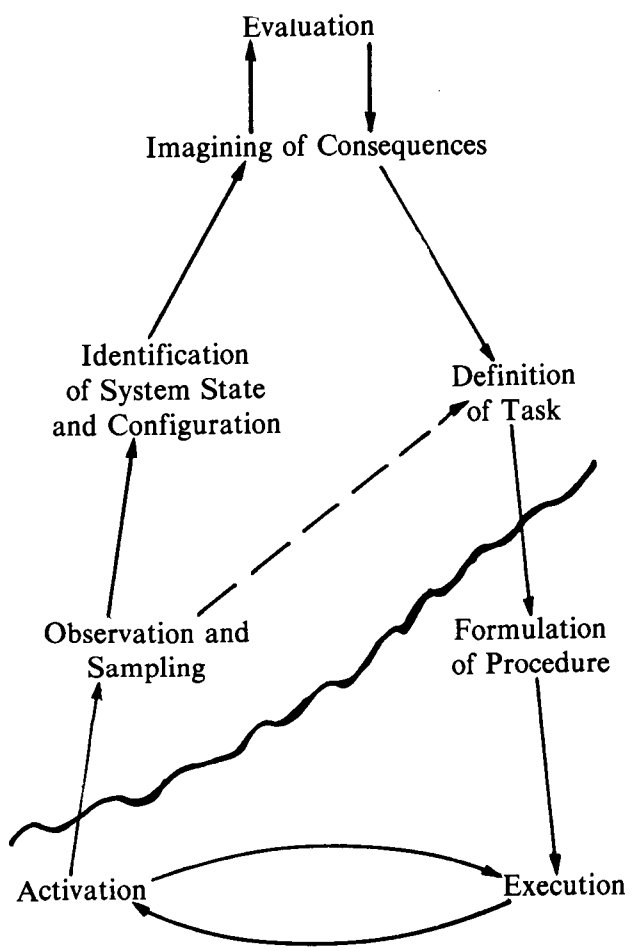


Fig. 1. Simplified Map of Mental Activities Sequence from Jens Rasmussen.

Table 1. Summary of Attributes of Models, Methodologies, or Theories for Various Applications

Model/Theory	Attribute
	For Observation and Sampling
Information theory	- Has been applied to aircraft cockpit design and evaluation - Closed form calculation is possible - Has narrow range of applicability
Queueing theory	- Has been applied to aircraft cockpit design and evaluation and air traffic control - Closed form calculation is often possible - Has wide range of applicability
Supervisory theory (decision theory based sampling)	- Some applications to radar monitoring have been made - Is potentially applicable
	For Observation → Identification (Detection)
Information theory (time to decide)	- Has been applied to other problems - Closed form calculation is possible - Is potentially applicable - Has narrow range of applicability
Signal detection theory (yes/no decision)	- Has been applied to medical decision making, sonar, avionics maintenance, and process control systems - Closed form calculation is possible - Is potentially applicable
Statistical pattern recognition (yes/no decision)	- Has been applied to medical decision making and process control systems - Closed form calculation is possible - Has wide range of applicability
Optimal estimation/filtering (yes/no decision)	- Has been applied to other problems - Has been applied to aircraft instrument failures - Closed form calculation is possible - Has narrow range of applicability
Rule-based models (yes/no decision)	- Have many related applications - Calculations are non-closed form - Have wide range of applicability
	For Identification (Diagnosis)
Statistical pattern recognition	- Has been applied in medical diagnosis - Has been applied in other judgement processes - Closed form calculation is possible - Has wide range of applicability
Rule-based models	- Have *many* applications, especially in artificial intelligence - Calculations are non-closed form - Have wide range of applicability
Fuzzy set theory	- Has been applied to medical diagnosis, fault diagnosis, and process control systems - Is "somewhat" obscure

Table 1 (continued)

Model/Theory	Attribute
	For Imagining of Consequences (Predictions)
Statistical pattern recognition	- Has been used to determine medical treatment - Closed form calculation is possible - Has narrow range of applicability
Rule-based models	- Have been used for air traffic control and other applications, especially in artificial intelligence - Calculations are non-closed form - Have wide range of applicability
Semantic models	- "Steamer" (qualitative simulator) - Tighter than rule-based models - Calculations are non-closed form
State equation models	- Have been used with great success in manual control - Have been used to describe plant - Closed form calculations are possible - Models are descriptive, but not necessarily explanative
	For Evaluation (Models of Choosing)
Utility theory	- (EV, SEV, EU, SEU)* - Has many relevant applications - Is normative
Prospect theory	- Has substantial potential - Is descriptive
Policy capturing	- Tells how decisions *were* made
	Fundamental Tradeoffs
Range of applicability vs. tightness of formulation Range of comprehensibility vs. tightness of formulation Broadness of conceptualization vs. runnability	

*EV, expected value; SEV, subjective expected value; EU, expected utility; SEU, subjective expected utility.

Working Group No. 2

SURVEY OF DEVELOPMENT OF MODELS AS APPLIED TO NUCLEAR PLANT OPERATORS (II)

J. G. Wohl, Chairman

S. Baron
B. J. Bell
A. M. Fullerton
J. P. Jenkins
W. J. Luckas, Jr.
D. A. Norman
R. Pulliam
M. Williams
D. D. Woods

In its survey of the development of cognitive models, Working Group No. 2 first asked the question: Is cognitive modeling/cognitive science a productive area for NRC to support? The consensus of the group was that the answer should be yes — but with provisos. While to date there have been few applications, those few have included some small-scale uses with high payoff. Certainly more basic science/exploration is needed. In the following outline, the group suggests a possible approach to the problem. The format is that of a series of questions and answers, the latter including recommendations for courses of actions.

Question 1: What can/should NRC do?

- NRC should support both basic science and applications.
- It should decide on a level of investment in the profession.
- It should not attempt to develop global models.
- It should back off to projects that have
  - — scientific merit and
  - — can be applied with high payoff.

Question 2: What work would give a high payoff?

- Scientific research projects that
  - — advance the art in formal and descriptive models,
  - — develop part-task models,
  - — develop human subsystem models,
  - — apply modern methods (e.g., artificial intelligence, optimal control and estimation).
- Model applications to
  - — assist and help direct task analyses,
  - — provide cognitive science support to specific NRC problems [such as human reliability/diagnostic error predictions (see Fig. 2 of addendum to this Working Group Report)].

Question 3: What are NRC's specific problems? *[Note: We have been asked for a list of models, that is, for a list of answers. Before we can provide the answers, we need a list of problems specified at the proper level so that we can determine which existing models can be applied and what new model developments and scientific investigations will be necessary.]*

- Suggestion: NRC should form an interdisciplinary task group to identify those problems ripe for immediate application.

Question 4: What kinds of problems can models be applied to? For what purposes are models useful?

- Example (a): Determining adequacy of system; that is,

Are operator resources { instrumentation, displays, procedures, controls, ... }

adequate to ensure "safe" { accurate and timely detection, diagnosis, action, within acceptable workload, ... } system operation?

- Example (b): Determining effect of operator capabilities/limitations on system performance, looking at performance of human/machine system (HMS) both for "best" operators and for "behaviorally deficient" operators, where the latter may denote
  - — inadequate mental model,
  - — failure to know or follow procedures,
  - — use of S-rules only (see addendum to this Working Group Report),
  - — inappropriate goals.

- Examples (c): Possible early applications, such as
  - procedural aids, that is, aids to

    increasing difficulty ↓

    - publication
    - documentation/justification
    - modification
    - critique
    - evaluation
    - generation

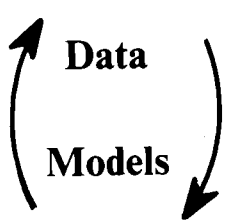
    of operating, maintenance and emergency procedures.

  - conceptual simulations (for training),
  - decision support for existing systems,
  - redesign of HMS interfaces for existing and planned systems.

- Examples (d): Applying models in other areas, such as:
  - methodology development,
  - sensitivity analysis,
  - hardware innovations for
    - display engines,
    - portable/low cost symbolic processors,
  - software innovations for
    - rapid prototyping.

- Example (e): Using task group approach as way to leverage scarce resources for maximum gain, for example,
  - possible analysis of SPDS* as test bed for application of task group, with goal to develop "lessons learned," as well a possible SPDS improvements.

Question 5: What is the data-model problem and its relationship to task analysis?



- We need data to apply models of cognitive activities.†
- We need models to know what data to collect.†
- We need to specify HMS functions, objectives, and decision requirements (see Figs. 1-3 for details of approach).

---

*Safety Parameter Display System.

†See addendum to this working group report.

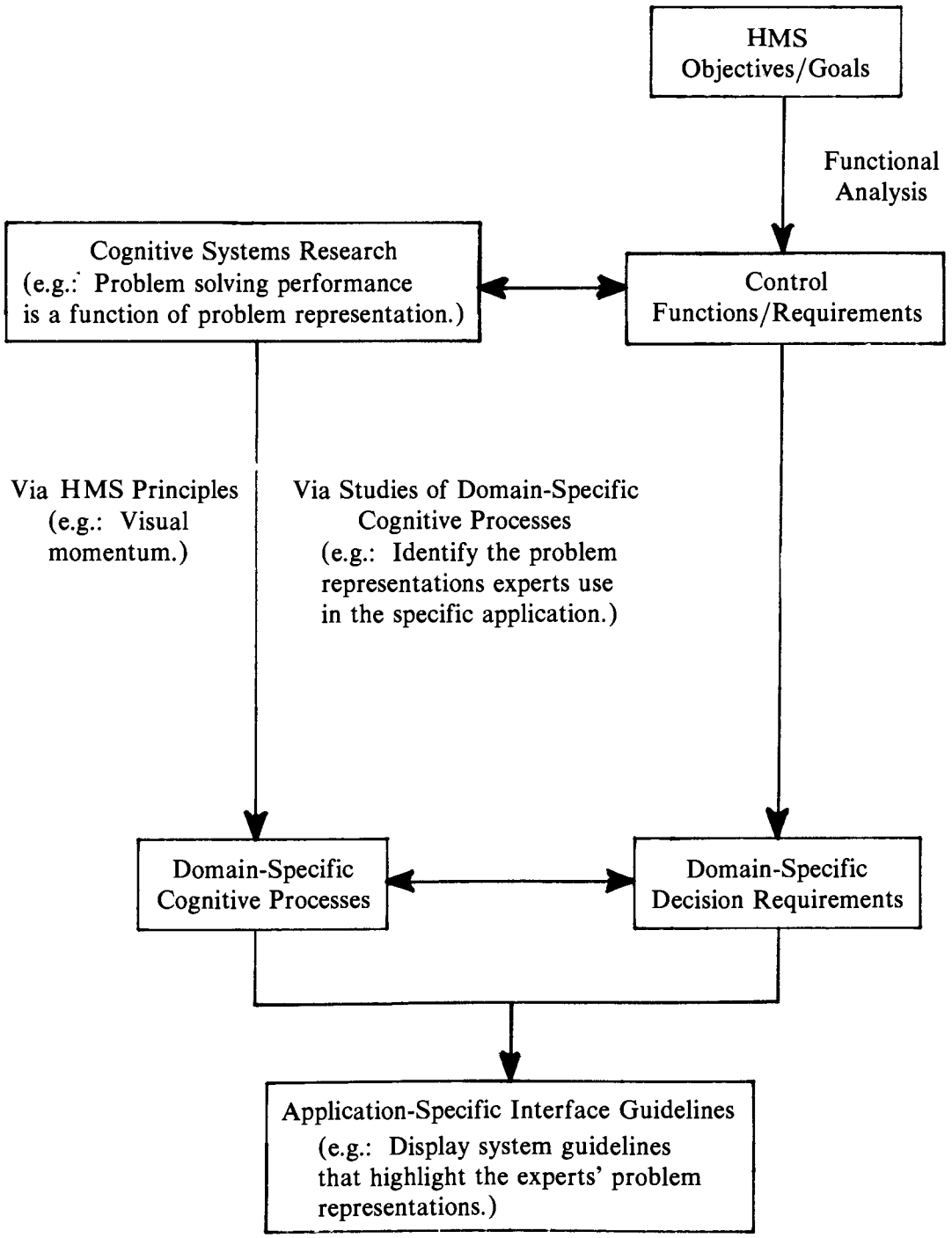


Fig. 1. Approach. (After Woods and Hollnagel.[1])

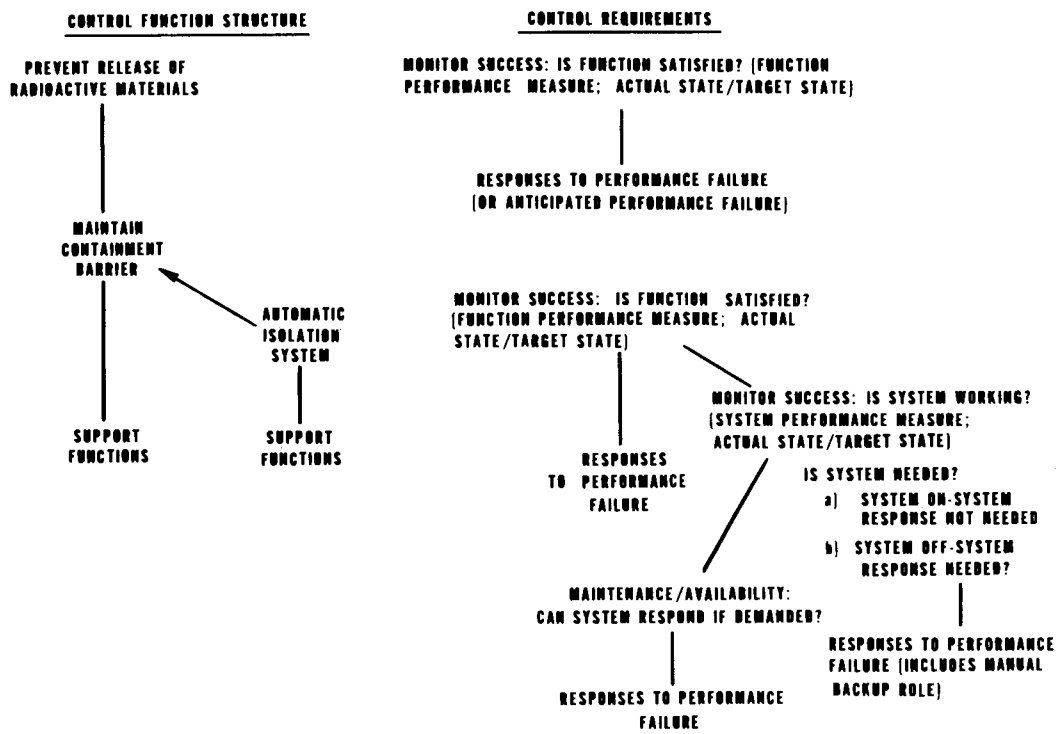


Fig. 2. Resolution of Functions. (From Woods and Hollnagel.[1])

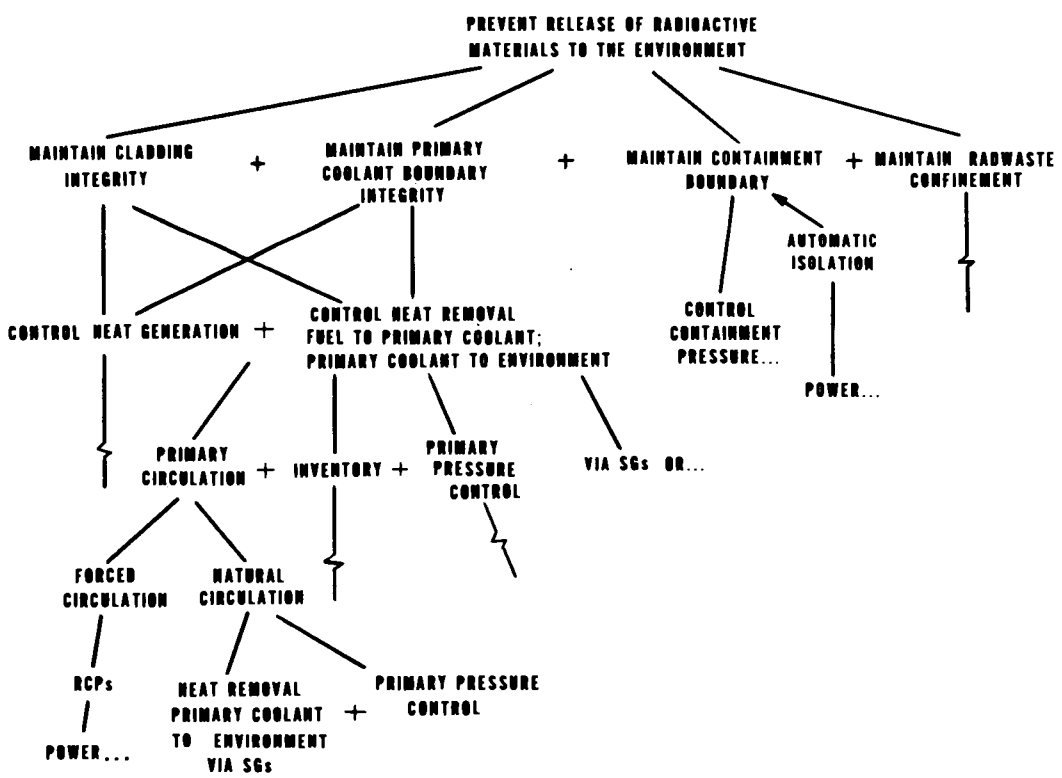


Fig. 3. Resolution of Objectives. (From Woods and Hollnagel.[1])

Question 6: How are we to get the necessary data?

- More than just funding will be required; for example, we will need access to
  - subject matter experts,
  - plants, simulators, and design shops.

Question 7: Where are we and in what directions should we be going?

- Current NRC program includes
  - expert system/AI problem solving, display requirements, experimental evaluation,
  - maintainer performance prediction model, task analysis, validation,
  - evaluation of psychophysical scaling methodology, use of utility theory to predict success/failure performance,
  - THERP[†] upgrade and final documentation,
  - human performance data collection from simulator training.

- We are further along than most people seem to believe, but big holes exist in the knowledge base. Examples of fundamental cognitive research problems we need to be working on are:
  - mental models (operator's mental representation of system),
  - mutual models (operator's mental models of each other),
  - models vis-a-vis task analysis (must expand task analysis technology to include cognitive activities),
  - multiple mental models (M^3) (operators exhibit not one but several overlapping models which affect/determine their behavior),
  - system's (designer's) model of the user (determining operator characteristics implicitly assumed by the system and the conditions under which these characteristics are beyond the capabilities of (1) the "best" operator and (2) the "behaviorally deficient" operator),
  - operator's cognitive activities,
  - structure of the environment (required for model of user).

Question 8: What about our knowledge of the nuclear power plant operator?

- "Go to the operator, learn his ways, and be wise." We need to understand much more fully the operator and "where he is coming from" (see Figs. 4 and 5).

[†]Technique for Human Error Prediction.

Increasing experience and capability	Knowledge-based behavior	Expert expert	Knows system intimately, has rich mental model of relationships, can modify mental model to account for events and can improvise new procedures in unusual circumstances.
		Novice expert	Knows system and rules, has simplified mental model of relationships and detailed understanding of system elements, and recognizes when rules no longer apply and mental model is inadequate.
	Rule-based behavior	Expert novice	Is still learning system elements, knows rules and when to apply them, and has primitive mental model of gross system relationships.
		Novice novice	Is learning system and rules.

NOTE: Many jobs/tasks are sufficiently simple not to require knowledge-based behavior.

Fig. 4. Relationship Between Types of Behavior and Degree of Expertise. (From Wohl.[2])

Question 9: What are our conclusions?

- NRC should look for some immediate applications of cognitive modeling with high payoff.
- NRC should use task group approach to select the applications and also to point out directions for new research.
- A long-range approach to fundamental research issues should be taken (science before engineering).
- Funding should be applied judiciously.
- Powerful results can be expected in "N" years from better understanding of cognitive functions, depending on
  - — a maturing technology,
  - — an expanded profession.
- We must all keep asking the question "What is the best direction?"

References:

1. D. D. Woods and E. Hollnagel, *Cognitive Task Analysis* (in press).
2. J. G. Wohl, *Mathematical Representation of Cognitive Processes in Human-Machine Interaction* (in press).

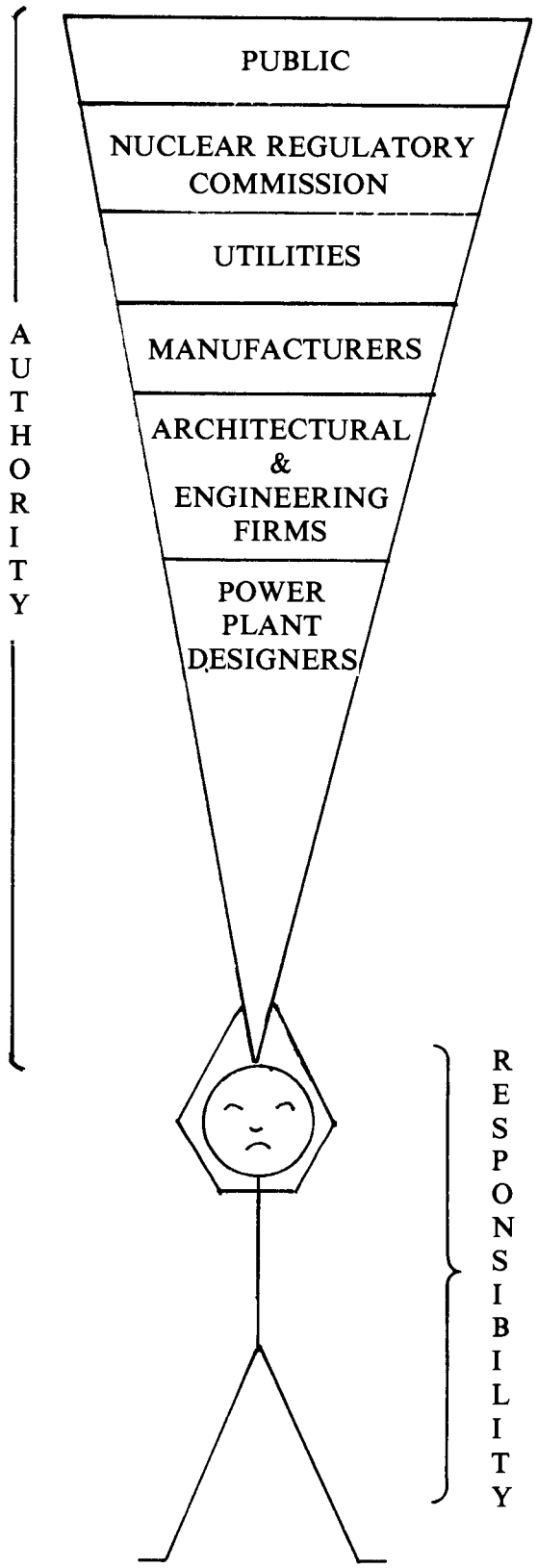


Fig. 5. Overburdening of Operator with Responsibility But with Little Authority Over System.

Addendum to Report of Working Group No. 2:

COGNITIVE CAPABILITY VS. SYSTEM COMPLEXITY IN ELECTRONIC MAINTENANCE

Joseph G. Wohl
Alphatech, Inc.

Abstract

Diagnostic behavior is strongly influenced by the interaction between the diagnostician and the failed equipment. Previous research resulted in a model of diagnostic behavior which predicts repair time distributions and expected values. The present paper shows the relationship between expected value of repair time, equipment complexity and human cognitive limitations.

Background

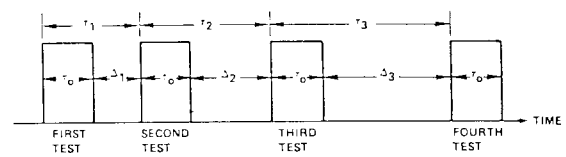
In a previous paper[1] a predictive theory of diagnostic behavior in electronic maintenance was proposed. The theory was based on the following observations and assumptions:

- Repair time distributions generally are highly skewed, such that mean repair time is largely determined by the degree of skewness (i.e., the tail of the distribution of repair times).
- Very long repair times appear to be due primarily to diagnostic difficulty. Since all data were corrected for time spent waiting for spares, and since data acquisition times per test tend to be less than an hour, this assumption appears warranted.
- A failed component generates a first-order suspect component set (namely, all those to which it is directly connected) of size S and mean $\bar{S} = I+1$, where I = equipment complexity index as defined below.
- The time to test a member of the set is a random variable with constant mean τ_0.
- The time between successive tests, i.e., diagnostic interpretation time, is a random variable with geometrically increasing mean Δ_n.
- The order of successive tests is not significant. It can be based on experience, relative failure probability, symptom familiarity, or random choice.

The relationship of τ_0 and Δ_n to the diagnostic process is shown in Fig. 1.

The theory was embodied in an analytical model involving the three variables I, τ_0, and Δ_n:

I = the average number of items (components, modules, circuit boards, etc.) directly connected to any one item.



$$\tau_n = \tau_0 / a^{n-1}$$

$$\Delta_n = \tau_n - \tau_0 = \tau_0 \left[\left(1 / a^{n-1} \right) - 1 \right]$$

WHERE

τ_0 = AVERAGE COMPONENT TEST TIME (INCLUDES REMOVAL, TEST AND REPLACEMENT)

Δ_n = AVERAGE TIME BETWEEN SUCCESSIVE TESTS; I.E., TIME FOR THE n-TH DIAGNOSTIC STEP

τ_n = TOTAL AVERAGE TIME REQUIRED FOR n-TH TEST PLUS n-TH DIAGNOSTIC STEP

Fig. 1. Relationship Between Time Per Test and Time Between Tests (Diagnosis Time). (From ref. 1.)

τ_0 = the average time in fractional hours to acquire and process information for a single diagnostic step or test.

Δ_n = diagnostic time (i.e., hypothesis generation and action selection time) for the nth diagnostic test step.

The first variable, I, is a direct measure of the complexity of the equipment. It is an equipment design characteristic and can be obtained by a simple enumerative process applied to circuit diagrams.

The second variable, τ_0, is a direct measure of the ease of diagnostic information access. It is a human-machine interaction characteristic and can be obtained by field observation.

The third variable, Δ_n, was utilized for curve-fitting as follows: Let n = number of tests performed. Then for the nth test or diagnostic step,

$$\Delta_n = \tau_0[(1/a^{n-1}) - 1], 0 < a < 1, n \geq 1. \quad (1)$$

Δ_n is taken to include the time for cognition, hypothesis development and analysis, and action selection. This formula represents the assumed geometric growth of diagnostic time (see Fig. 1). Simple variations in the parameter a allowed adjustment of the model to data (i.e., curve-fitting).

As reported in ref. 1, the analytical model was subjected to initial testing in both standard and expected value forms with encouraging results (see Fig. 2, for example).

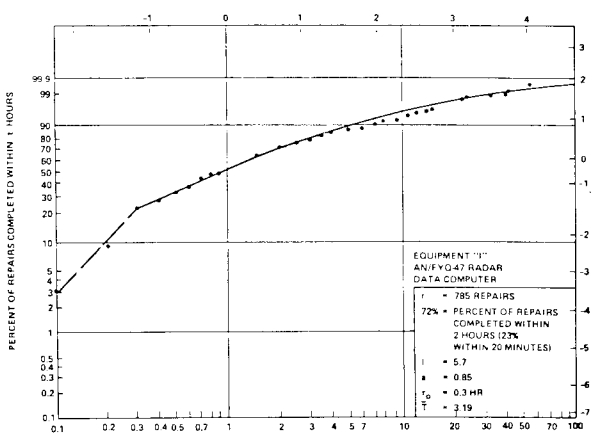


Fig. 2. CFD Data for AN/FYQ-47 Radar Data Computer Plotted Against F(t) As Computed from Model. (From ref. 1.)

The solutions to the expected-value equations for the overall model, as given in ref. 1, fell into three separate domains:

For $a = 1$,

$$\bar{T} = \tau_0 \left[\frac{1}{I\left(1-\frac{1}{I}\right)^2} \right] \left\{ \frac{1}{2} I^2 \left[1 - \left(\frac{1}{I} \right)^2 \right] - \ln I \right\}. \tag{2}$$

For $1 > a > \left(\frac{I-1}{I} \right)$,

$$\bar{T} = \tau_0 \left[\frac{1}{I\left(1-\frac{1}{I}\right)^2} \right] \times \tag{3}$$

$$\left\{ \frac{-a^2(2-a)\ln\left(\frac{1}{I}-1+a\right) - a\ln I - Ia(1-a)\left(I-\frac{1}{I}\right)}{(1-a)^2} \right\}.$$

For $a \leqslant \left(\frac{I-1}{I} \right)$,

$$\bar{T} = \infty. \tag{4}$$

Issues

In retrospect, the analytical results raised several important issues:

(1) Of the three separate domains defined in Eqs. 2-4 above, the first corresponded to the laboratory data cases; i.e., for those instances in which laboratory data were taken, the value of a required to fit the data was always 1.0, resulting in a simple exponential distribution of repair time. Why should the laboratory environment produce this result, as opposed to the field environment?

(2) The second solution domain (Eq. 3) corresponded to the field data cases. For the 10 equipments for which field maintenance data were available, the value of a required to provide a best visual fit was found to be between 0.85 and 0.90. Why such a narrow range of values? What is the significance of this parameter? Does it somehow represent (or at least involve) some basic human characteristic or limitation in the technician-equipment interaction?

(3) Most puzzling of all was the third solution domain; for if the inequality of Eq. 4 holds, then mean repair time is infinite. This means that there can be cases in which equipment complexity is such that human diagnosis will never be successful (or, more precisely, that diagnostic time can grow without bound). Can this be so? And if so, why?

(4) Finally, one of the consistent findings in ref. 1 was that all of the field maintainability data showed a characteristic bimodality (see Fig. 2, for example). Could this be related to fundamental differences in search techniques such as those represented by Rasmussen's symptomatic (S) and topographic (T) search strategies?[2] If so, how?

Results

With regard to the first issue, that of the discrepancy between laboratory and field repair time distributions, a discrete-variable differentiation of Eq. 1 shows that the rate of increase of Δ_n quickly approaches the asymptotic value (1-a) as n increases. Thus, for the laboratory data cases, in which it was found that $a=1$, it would appear that Δ_n, rather than increasing geometrically with n, remains constant. This

effect can be interpreted as follows. If a strictly random search of the suspect set S is made, then the probability of locating the failed component on the nth "try" is represented by a uniform distribution of range S and uniform probability $P_n = 1/S$, $1 \leqslant n \leqslant S$. If the distribution of S itself is Poisson for any given equipment (e.g., Equipment A in ref. 3), then this assumption gives rise to the exponential distribution of repair time typical of laboratory data. This effect is entirely predictable as long as "rare events" such as "intermittents" or unusual component characteristics, which often account for the longer diagnostic times, are not generally introduced in laboratory situations.

To address the second and third issues identified in the preceding section, let us begin by solving Eq. 4 for I:

$$I \geqslant \frac{1}{1-a} \text{ or } I \geqslant k, \text{ where } k = \frac{1}{1-a}. \qquad (5)$$

This states, in essence, that if equipment complexity is at least as great as some number (say, k), then $\overline{T} = \infty$. This result in itself is suggestive. But what is the value of this number k, and more importantly, what is its significance?

The answer lies in the narrow range of values of a among the 10 cases of field maintenance investigated in ref. 1:

$$0.85 \leqslant a \leqslant 0.90 \quad .$$

Substituting the extremes for a in Eq. 5, we obtain the following direct interpretation:

If $I \geqslant k$,

$$\text{then } \overline{T} = \infty \text{ if } \left\{ \begin{array}{ll} k = 6.67 & (a = 0.85) \\ k = 10.0 & (a = 0.90) \end{array} \right\} .$$

The meaning of this inequality now becomes clear: If equipment complexity exceeds the technician's ability to understand and deal with it, then $T = \infty$ indeed. Hence the curve-fitting parameter a, along with its derivative parameter k, must be related to a human cognitive complexity limitation.

We suggest that k is, simply, the number of items that can be simultaneously held in human short-term memory. Thus if equipment complexity as measured by the average number of connections per component exceeds human "buffer storage" capacity, one may expect eventually to experience maintenance situations involving unsuccessful fault isolation over increasingly long periods of time. Technicians often speak about troubles that have defied diagnosis to the point where equipment had to be either completely overhauled or discarded; software "bugs" have been found to resist isolation for years. So there is anecdotal evidence that such cases have indeed occurred, if rarely.

The fourth issue is of interest from a design standpoint. Rouse and Hunt[4] refined Rasmussen's taxonomy of search strategies and defined their boundaries, calling them S-rules and T-rules. S-rules are system-determined and serve to map symptoms directly onto *system* structure, i.e., failure *sources*, while T-rules are context-free and serve as guidelines for searching or analyzing the *problem* structure, i.e., the symptom *reasons*. Since S-rules are efficient, they are generally used in preference to T-rules and are abandoned only after they are found not to work. The relative efficiency of S-rules can be presumed to account for the early or accelerating part (t<1 hour) of the field maintainability data plots in ref. 1 (e.g., as in Fig. 2) and the T-rules to account for the later or decelerating part of the plots. In addition, S-rules can easily be automated, and much built-in test equipment and test routines embody such rules. T-rules, on the other hand, are much more difficult to automate, and it would appear that future progress in designing equipment to reduce diagnostic time will occur as a result of efforts in this area.

Discussion and Conclusions

An attempt has been made to explain some of the puzzling consequences of the theory of diagnostic behavior proposed in ref. 1. Of course, the theory itself requires further experimental validation. But the foregoing results serve to simplify the theory by essentially fixing one of the parameters, viz: $k = 7$ (ref. 5) or, equivalently, $a = 0.857$. This leaves only the two variables I and τ_0 to be measured experimentally and eliminates the arbitrary curve-fitting parameter. It is also clear that the theory proposed in ref. 1 applies only to T-rule behavior. However, this suggests that applications of the theory may extend beyond electronic maintenance and into other human problem-solving and diagnosis areas in which T-rules play a major role.

Finally, one of the potentially important and disturbing implications of the theory should be pointed out.* The equipment for which repair time data were available was representative of 1950's and 1970's technology. Note that this equipment was all designed by engineers with normal cognitive complexity limitations. The electronics technology of the 1980's and 1990's is moving increasingly towards Computer Aided Design (CAD), for which no such limitations exist. What are the implications of CAD for field maintenance in general, and in particular for system partitioning, test equipment design (built-in and otherwise), diagnostic procedures, personnel selection and training, and logistics and spare parts policy? It may be necessary, for example, to require that CAD constrain the physical partitioning of the resultant design such that its complexity index is less than 7. In any event, it will be necessary to match the diagnostic abilities of future test systems with the complexity of future electronic systems.

---

*I am indebted to Dr. Joel Lawson, Chief Scientist, C^3 Division, U.S. Naval Electronic Systems Command, Washington, D.C. for raising this issue.

References

1. J. G. Wohl, "Maintainability Prediction Revisited: Diagnostic Behavior, System Complexity, and Repair Time," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-12, No. 3, May/June 1982.
2. Jens Rasmussen, "Models of Mental Strategies in Process Plant Diagnosis," pp. 241-258 in *Human Detection and Diagnosis of System Failures*, NATO Conference Series, Vol. III: Human Factors, Plenum Press, New York and London.
3. J. G. Wohl, "System Complexity, Diagnostic Behavior, and Repair Time: A Predictive Theory," pp. 217-230 in *Human Detection and Diagnosis of System Failures*, J. Rasmussen and W. B. Rouse, Eds., Plenum Press, New York (1980).
4. W. B. Rouse and R. M. Hunt, "A Fuzzy Rule-Based Model of Human Problem Solving in Fault Diagnosis Tasks," pp. XV 95-101 in *Proc. Eighth World Congress of IFAC*, Kyoto, August 1981.
5. G. A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychological Ref.*, Vol. 63, pp. 81-97 (1956).

Working Group No. 3

USE OF COGNITIVE MODELS OF NUCLEAR PLANT OPERATORS IN DESIGN AND OPERATION

J. Wreathall, Chairman

W. R. Corcoran
D. Farr
J. M. Gallagher, Jr.
T. Hatcher
K. Murphy, Jr.
R. A. Newton
T. Ryan
J. W. Senders

The overall conclusion of this working group was that we could not reach any unanimous agreement whether the development of models of cognitive behavior were worthwhile on the part of the industry. The group found it difficult to understand how those methods of cognitive modeling currently available and discussed both in the position papers and by the other working groups could address the present concerns of the industry. In particular, it does not appear that the concerns facing the industry, such as control room reviews (a la NUREG 0700), the installation of safety parameter display systems, and the development of alternative approaches to emergency procedures can be analyzed by the kinds of models that have been developed for very different kinds of man-machine systems, such as aircraft-tracking tasks.

Because the group could not see any potential payoff in these areas of concern, there was a general reluctance to support the development of models of cognitive behaviors — at least explicit, quantitative models. However, the group did recognize the need for a framework that would relate activities in the different areas of man-machine systems, such as procedures development to training practices, display system design to both of these, and so on — i.e., a framework of integration. This could be provided in part by inputs from cognitive psychology *in conjunction with* systems engineering methods. In other words, understanding from the field of cognitive (and other kinds of) psychology should not be excluded, but neither should it be considered an answer in itself.

Examples of how information from cognitive psychology could be incorporated were described in the context of task analysis, where the addition of cognition (what kinds of decisions need to be made, and when) to the earlier and more mechanistic kinds of task analysis enables a much more comprehensive identification of significant task-related factors. However, until the system designers (in the widest sense) establish a sufficient definition of the roles of operators (and this interfaces with other aspects of an operating system), only a piece-meal approach will result.

The group did recognize and implicitly expressed its concern over the very real difficulties of communications between designers of control rooms, procedures, and training programs on the one hand and cognitive psychologists and mathematical modelers of human behavior on the other. These two groups use entirely different vocabularies and a high degree of confusion exists between them. There are no forums where the two groups interact except at very infrequent specialist workshops (for example, at this workshop and at the two Myrtle Beach workshops).

The time scales of concern to the two groups also tend to be very different. Designers generally must respond to the short-term needs of the industry and the licensing authorities. By contrast, much of cognitive psychology (which was first identified as a discipline only in the mid-1950s) is a science of gradual developments. The work that has produced existing models has largely been a result of concentrated effort and resources on relatively small areas (for example, in the study of aircraft-pilot tasks) which are not necessarily the areas of concern of the nuclear industry.

In view of these differences, there was a general feeling by both those from design backgrounds and those from modeling backgrounds that unless the communication barriers can be reduced, little progress can be made in the near future. The needs of the industry must be stated *sufficiently clearly* to those in the cognitive psychology area — the capabilities of the modelers must be stated *sufficiently clearly* to the industry. If the NRC can provide a possible link between the groups, it will be helping the industry to be safer. The fact that the aerospace industry and defense industries have developed better ways of solving their human factors problems should encourage the nuclear industry to expect a similar return on their investment.

Below is part of a list jointly agreed upon by both the modelers and the potential users, as a list of terms applicable to themselves:

Modelers	Potential Users
• Ignorance	• Arrogance
• Incompetence	• Ignorance
• Insularity	• Shortsighted/tunnel vision
• Lack of understanding of needs of users	• Pride
	• Insularity
	• Stupidity
	• Pessimism
	• Lack of perceived needs

This was our consensus point. It gives rise to an optimistic feeling that each group recognized its own failings — particularly the last item in each list — as well as those of the other groups.

In conclusion, this working group demonstrated that significant gaps in communications exist between the industry groups on one hand and the cognitive modelers on the other. The modelers seem to have a limited understanding of what are the current problem areas in human-system interactions in nuclear power plants, and those concerned with design and operation in the industry often have little understanding of the insights into human performance which could be obtained from the modeling of cognitive processes. The NRC is in a position to bring these two groups together by developing an interface process. Another working group raised the possibility of organizing a small task force to act as an interface between the groups. An alternative, which is already in place, is the IEEE/NPEC Subcommittee 7, Human Factors and Control Facilities, which includes representatives from industry, psychological modelers and the NRC.

The process of communication and mutual education is as much a part of the overall development and implementation of models as any structured modeling. It deserves its share of any resources devoted to this area.

Working Group No. 4

USE OF COGNITIVE MODELS OF NUCLEAR PLANT OPERATORS IN REGULATION AND LICENSING

P. M. Haas, Chairman

L. Beltracchi
J. M. Christiansen
R. Curry
D. E. Embry
H. Halff
L. F. Hanes
F. R. Mynatt
C. M. Overbey
R. W. Pew

Working Group No. 4 focused on the needs of NRC and the potential use of cognitive models. It de-emphasized the state-of-the-art of modeling, though the subject was certainly discussed with regard to what might be available on a "near-term" versus "long-term" basis.

The working group operated by first having "mini-presentations" by several group members on topics which were either of direct concern to the individual or were summaries of information/knowledge from the individual's personal experience which could be shared with the group. Those making presentations were:

(1) Dick Pew, who discussed some concerns about the emphasis on "reliability/PRA" methods versus cognitive modeling and some obstacles to extended use of cognitive modeling by NRC.

(2) Henry Halff, who addressed several questions relevant to regulation in the areas of training and emergency procedures; he suggested that while complete cognitive models of operator behavior might not be available for some time, there are several methodologies in cognitive science that could be brought to bear on the questions both in the near-term and in the mid-term.

(3) Fred Mynatt, who described the "pressurized thermal shock" problem, which is an example of a critical issue now facing NRC that could benefit from cognitive modeling or at least from input from cognitive science.

(4) Leo Beltracchi, who presented a possible iconic which could be incorporated into a display to be used as an operator aid, this being an example of key concerns he has about control room displays which could be addressed by cognitive science.

(5) David Embrey, who summarized the need for cognitive modeling to address critical questions of operator reliability as part of PRA.

These five presentations provided a base for the group's discussions and in fact led to specific examples of how cognitive modeling (or at least results/conclusions from the field of cognitive science) could be applied to problems faced by NRC.

The overall conclusion of the group, based on group discussions and the presentations and discussions of the total workshop, was that there is indeed a need for long-term research. Cognitive science is in a relatively early stage of development, but certainly

there are models and results of work in the field that are applicable to critical issues facing NRC and are *available on a near-term basis.* Perhaps the most significant point made is that NRC is already making use of "models" of cognitive behavior to make critical decisions about operator qualifications and training, control room design, operator aids, etc., but these models are, in general, not written down, not structured, and not validated. They are, in fact, the "mental models" of NRC staff who are in the position of having to make decisions now. It was the group's opinion that the need for a more structured application of existing models and continued work to develop and validate models is obvious.

In the group's report to the workshop as a whole, three speakers discussed specific areas currently being addressed by NRC for which cognitive modeling or at least the body of existing knowledge in the cognitive science area would be beneficial. These areas, in order of presentation, were:

(1) *Review and evaluation of operator aids and, more generally, control displays.* In a brief "simulation," or role-playing, a "typical" display design for a proposed operator aid was presented by a vendor to NRC for an acceptance review. The kinds of evaluation and concerns of NRC and the way in which the body of knowledge available in cognitive science might help NRC in the acceptance review were brought out by the role-playing session.

(2) *Requirements for operator training and for abnormal/emergency operating procedures.* Henry Halff reviewed the Office of Naval Research's experience in training and procedures development/evaluation which appeared to be directly applicable to NRC needs.

(3) *Probabilistic risk assessment.* David Embrey summarized the status of human reliability estimation as part of probabilistic risk assessments currently in progress, the apparent criticality of "cognitive action" in event sequences, and some recommendations for research in this area. As a specific example, he summarized the current evaluations of the "pressurized thermal shock" problem, noting that resolution of this problem is considered one of the most critical research questions facing NRC at this time, if not the most critical, and that apparently some of the most critical questions involved with the sequences are those associated with the operator's reliability in "cognitive action."

The material presented by these three speakers plus the introductory material presented by the working group chairman are given in the addendums that follow.

Addendum A to Report of Working Group No. 4:

INTRODUCTION OF GROUP PRESENTATIONS

P. M. Haas, Group Leader

In thinking about NRC's use of cognitive models of nuclear plant operators with respect to regulation and licensing, we first asked ourselves what we perceived NRC's regulatory and/or licensing activities to be. The answer was that NRC is (1) to issue rules, regulations, and requirements for "things" that must be provided or accomplished by the licensee and (2) to review and accept or reject "things" that are submitted by the licensee. In all cases the criteria used by NRC in carrying out these activities must be defensible.

With regard to the operator in the control room, NRC's activities include many specific items in a number of areas. For example, NRC is concerned with the design and layout of both the control room and the display panels in the control room. It is also involved with the evaluation of the procedures used in the control room. And, of course, it is concerned with determining control room staffing requirements and the criteria for qualification and training of the staff. Finally, NRC has made a major commitment to the use of PRA (probabilistic risk assessment) to identify and quantify risk to the public, and human reliability assessment is an integral part of PRA.

After listing what NRC's activities were to be, those of us in the working group then asked ourselves whether NRC needed information about cognitive behavior and/or whether it needed cognitive models to carry out these activities. The concensus was that there are aspects of cognitive behavior that need to be considered by NRC in each of these areas. Furthermore, we agreed that NRC already has cognitive models and is using them to make specific decisions that will have a major impact on cost and (possibly) safety. Examples are: determining the requirements for SPDSs (safety parameter display systems), arriving at symptoms versus event based procedures, and setting education requirements. In general, however, these models are implicit, ill-structured, and not formalized. That is, they are NRC's "public cognitive models(s)" of the operator and the rest of the system.

We believe that NRC needs a better basis for making decisions like those cited above and that cognitive modeling would provide a better basis. We did not systematically address costs and benefits of modeling; nor did we discuss priorities of research. There was a concensus that long-term research requirements exist; however, some capabilities are available and can be applied to meet near-term needs (within one to two years). Examples of possible uses (both near-term and long-term) for cognitive modeling in addressing specific NRC concerns are given in the following presentations.

Addendum B to Report of Working Group No. 4:

ROLE-PLAYING SESSION: PRESENTATION AND REVIEW OF PROPOSED CONTROL ROOM DISPLAY

Vendor: L. Beltracchi
NRC Decision Maker: L. Hanes
NRC's Consultant: R. Pew

The following role-playing session was conducted by Working Group No. 4 to illustrate to the workshop participants how NRC might use knowledge from cognitive science during an acceptance review of a proposed operator aid. First a rankine cycle iconic display is presented by the vendor to the NRC decision maker, who decides he needs advice from an expert consultant. It is to be pointed out that in the actual presentation, the vendor did not use all the information included below.

Vendor: A temperature-entropy iconic display of the Rankine cycle based upon the measured variables of the secondary loop conveys the operating status of heat transfer to and from the secondary loop of a PWR. By processing and integrating data from each sensor in the secondary loop, an iconic overview of the process in the secondary loop is obtained. With the addition of the primary coolant system parameters and the condenser cooling water parameters, a heat source to heat sink process information iconic is obtained. Finally, by adding the saturation line for water and for steam, additional useful data on plant safety is possible.

A rough sketch of this type of iconic display for a PWR at design power is shown in Fig. 1. For a PWR, the primary coolant is a few degrees subcooled. The line AB represents the primary coolant heat addition in the reactor core. The line BA represents the heat

removal from the primary coolant in the steam generator. The small space between the line AB and the saturated water line indicates subcooling of the primary coolant but not the amount of subcooling. Data from currently measured parameters, such as temperatures and pressures, along with the thermodynamic properties of compressed subcooled water, are needed to generate the display.

At steady-state conditions, the primary coolant loop pressurizer consists of saturated steam and nearly saturated water. From measured water level, temperature, pressure, and the volume of the pressurizer, the density of steam, the density of water, the mass of the steam, and the mass of the water in the pressurizer may be estimated. From these, the computer calculates the quality of the mixture in the pressurizer and this is symbolized as point C in Fig. 1. The vertical distance between point C and point B is the magnitude of subcooling for the average core exit coolant.

The secondary loop is shown in DEFGHIJ, a Rankine cycle with no reheat. Heat addition in phase change from water to steam is represented by DE, heat removal by FGH. Other details are standard thermodynamic fare and may be found in any good thermo book. The tertiary loop, condenser cooling water, is also shown in Fig. 1. This is the ultimate heat sink, the environment.

The iconic display is based on a temperature versus entropy diagram for the process. The issue of display resolution between subcooled water and saturated water for both the primary coolant and the feedwater is poor. This is due to the nature of the temperature versus entropy properties of subcooled compressed water.

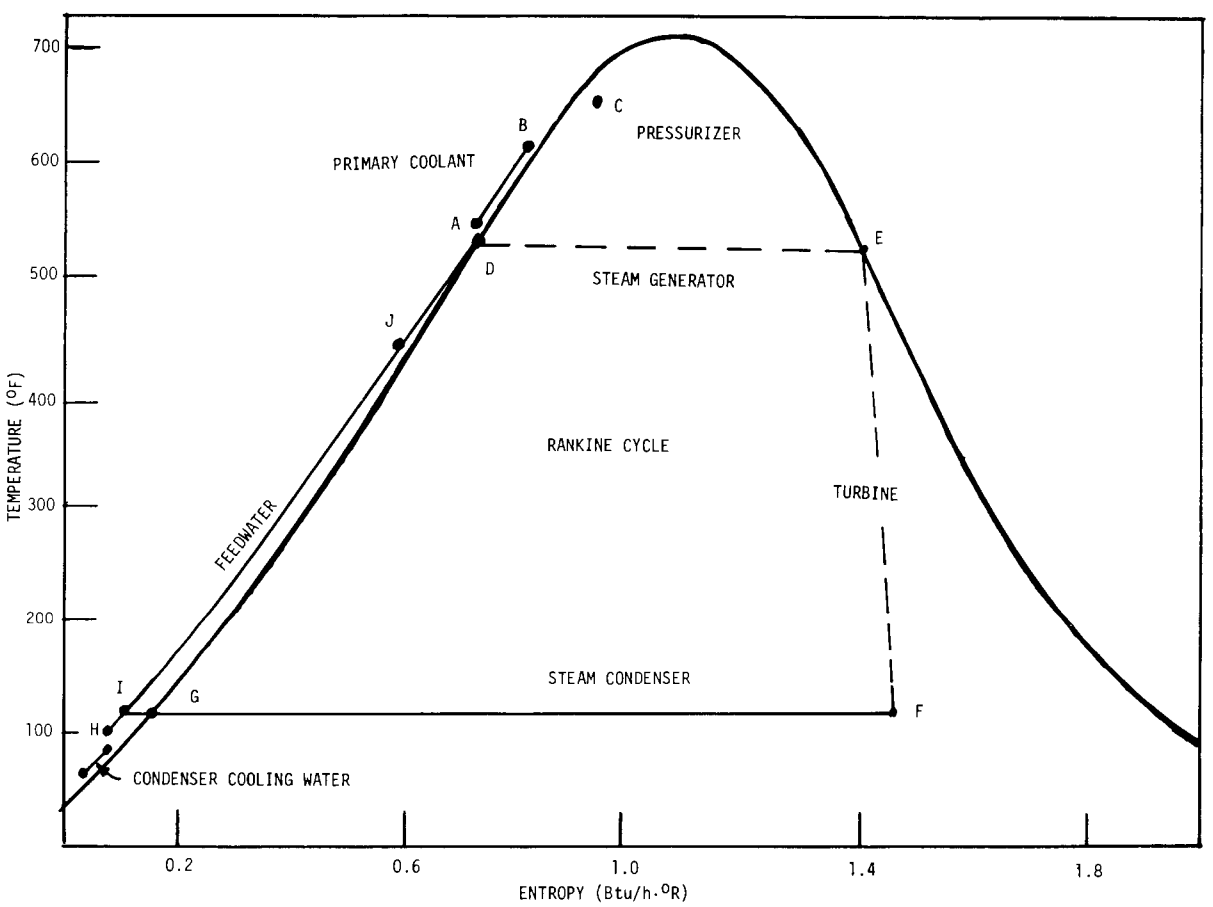


Fig. 1. Iconic Display: Heat Source to Heat Sink (Design Power).

With a display designer's poetic license, the issue of subcooled water data may be resolved. Figure 2 presents a modified display wherein the magnitude of primary coolant subcooling is shown as a bar of variable width. For example, at core exit, the coolant is 40 degrees subcooled, whereas at core inlet, it is 100 degrees subcooled. Steam generator feedwater and condenser cooling water are also shown in a similar manner.

Figure 2, as an iconic display for the nuclear power plant process, conveys a large amount of information to the control room operator. It shape codes all coolant temperatures into an easily recognized functional pattern that models the heat transfer from source to sink. Since the shape coding of the information allows for easy pattern recognition by the operator, he can, at a glance, ascertain the status of the primary coolant subcooling. For example, a narrow bar means the loss of margin to saturated liquid. Operator action may be required to restore margin. Changes in the shape of the subcooling bar are most important during transients, which is discussed later.

The iconic display also supports pattern recognition of failures in the heat transfer process. For example, the magnitude of the space between the cold primary coolant temperature and the steam generator temperature (saturated water) is indicative of the magnitude of heat transfer from the primary coolant to the secondary coolant under normal conditions of operation. An increase in the magnitude of the space might indicate a failure of heat transfer between the primary and secondary system. It might also indicate a large increase in power, which at design power would be undesirable and a reactor trip should occur. Overheating of the primary system is also undesirable. This pattern in the iconic should be easily detected by the operator and it defines a point where diagnosis of the problem should begin.

Figure 2 also integrates safety information and normal process information. One concern is whether wet steam might damage the turbine. Sensors measuring the steam quality at the turbine inlet provide data in the display. Similarly, core inlet and exit coolant temperatures are provided, from which the operator can determine whether the reactor core is being cooled.

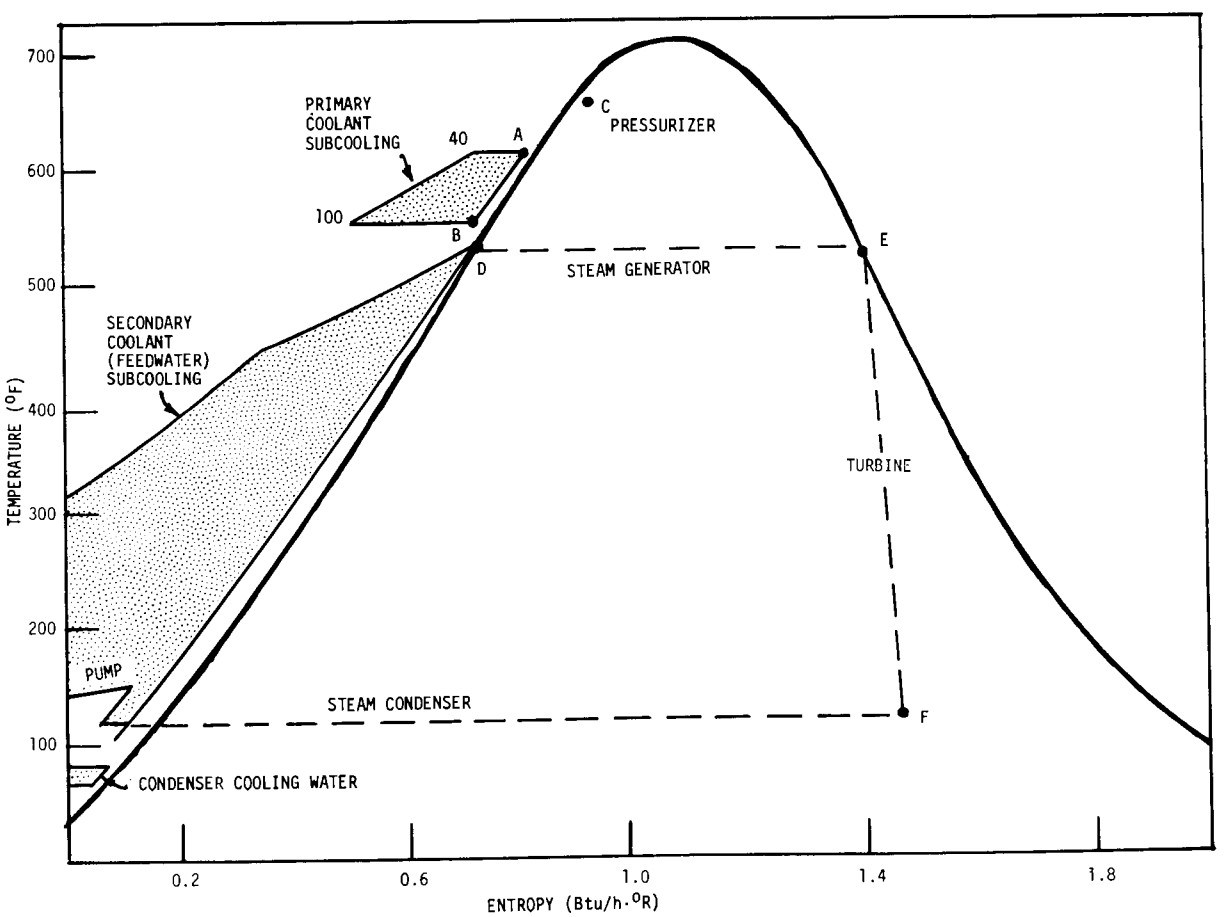


Fig. 2. Modified Iconic Display: Heat Source to Heat Sink (Design Power).

Without specific data for transients, it is not possible to structure the exact form of the iconic display for transient conditions. However, some general trends can be discussed and these will prove useful.

Figure 3 approximates a post-reactor-trip condition with normal afterheat removal. The primary coolant, about isothermal, is subcooled and the steam in the generator dumped to the condenser via a turbine bypass. The magnitude of the subcooling is now coded as the length of the line terminating at A. Should the length of the line decrease to zero, then the primary coolant is saturated liquid and localized boiling of water may be occurring at core hot spots.

The status of the heat-transfer process from the primary coolant to the secondary coolant may also be evaluated over a period of time with the patterns from Fig. 3. During afterheat removal, the space between the primary coolant temperature and saturated water temperature in the steam generator should decrease with time, reflecting the decay of afterheat. Should this space increase with time and the primary coolant temperature increase, this pattern indicates the reactor core is heating up.

From a conceptual viewpoint, the heat source to heat sink afterheat removal display pattern is easily evaluated. Deviations from this pattern should be easily identified and could be used as a starting point to diagnose the situation. After successful diagnosis, the operator selects a course of action to mitigate the consequences of the event. The impact upon the process from these operator actions can also be monitored from the iconic. Thus, the iconic serves to aid in the detection and diagnosis of events and also serves as a monitoring device to evaluate mitigation efforts.

The iconic changes when significant boiling occurs in the reactor core, and a mixture of steam and saturated water is transported through the hot leg of the primary system. Figure 4* presents the situation. Under these circumstances, cooling the core may become a critical problem when the primary coolant is solely a vapor. In terms of process information, the two-phase flow in the hot leg is important information to the operator. Continued two-phase flow in association with the loss of primary coolant mass eventually results in loss of the ability to cool the core. Detecting continued two-phase flow in the hot leg would be useful to the operator as a symptom for initiating measures which result in greater cooling of the core. This information early in a event would be much more timely and useful than waiting until one detects and measures core water level.

Other paths of heat transfer from the core to the environment are available. No attempt has been made to evaluate if the iconic display would be feasible for these paths.

The above type of model supports a mechanical engineer in the analysis of the plant process. An operator's model of the plant may be entirely different. However, the Rankine cycle presents a concise model of the process which can be easily understood by operators if properly trained. One benefit from this would be the design of a data base. Such a data base could be divided into the individual functions within the cycle. Furthermore, sensor readings and system status could also be organized into function-oriented data sets. This type of data base would provide a firm foundation upon which to add predictive type displays because functions within the plant processes may be easily determined.

Another potential benefit from this type of iconic display is the coding of information from the heat source (the reactor) to the heat sink (the environment). By integrating data into one display, the operator's workload in assessing the plant process is greatly reduced.

Mimic diagrams of the plant process are also used as models of the plant. Adaptive mimics which reflect the measured process variables and the flow of coolant are also useful to the operator. For these latter type of mimics, added data to indicate two-phase flow in the hot leg of the primary coolant could prove very useful to operators in assessing whether the core is being adequately cooled.

NRC Decision Maker. Since we have not issued a requisition for this display and had not anticipated purchasing this particular item, I do not have available a set of criteria to compare it with. But since it is related to safety, we are, of course, very much interested in it.

I would like to know what prompted you to design this display. Did it come from a spark of genius, or did you go through a systematic process? If you have actually considered some of the cognitive tasks of an operator during a control room procedures review, perhaps going over that review would help me realize that the operator's needs are not currently being met and I would have some basis for evaluating your proposal.

We probably need an HRA (human reliability assessment). It should be a low-level assessment. Don't get into higher-level cognitive abilities. Orient it toward the equipment operator rather than the system manager. Some work in the cognitive area could give the regulator criteria to evaluate a fancy HRA that does look at some of these. We have a consultant in cognitive science who may be able to help us evaluate your design.

NRC Consultant. My role is to attempt to explain how this display could be used. The level we focus on is the way we envision this could be used at a higher level than fundamental physics. I would focus on questions such as: How do you imagine an operator would conceive entropy? Give me a word recipe. What knowledge structure does an engineer believe would underlie this? What information is potentially communicated to the operator? A feature of the display is that it is presenting several dimensions in an integrated way.

*Figure 4 — Exact iconic display of this transient may not be feasible.

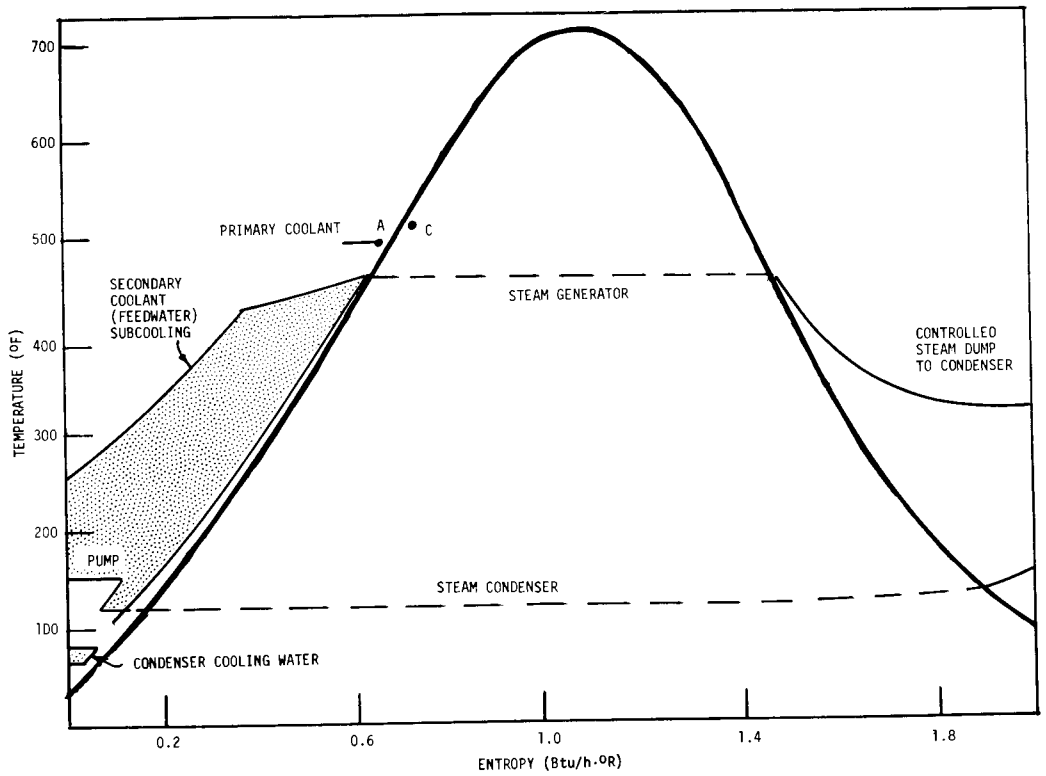


Fig. 3. Iconic Display: Post Reactor Trip (Approximate).

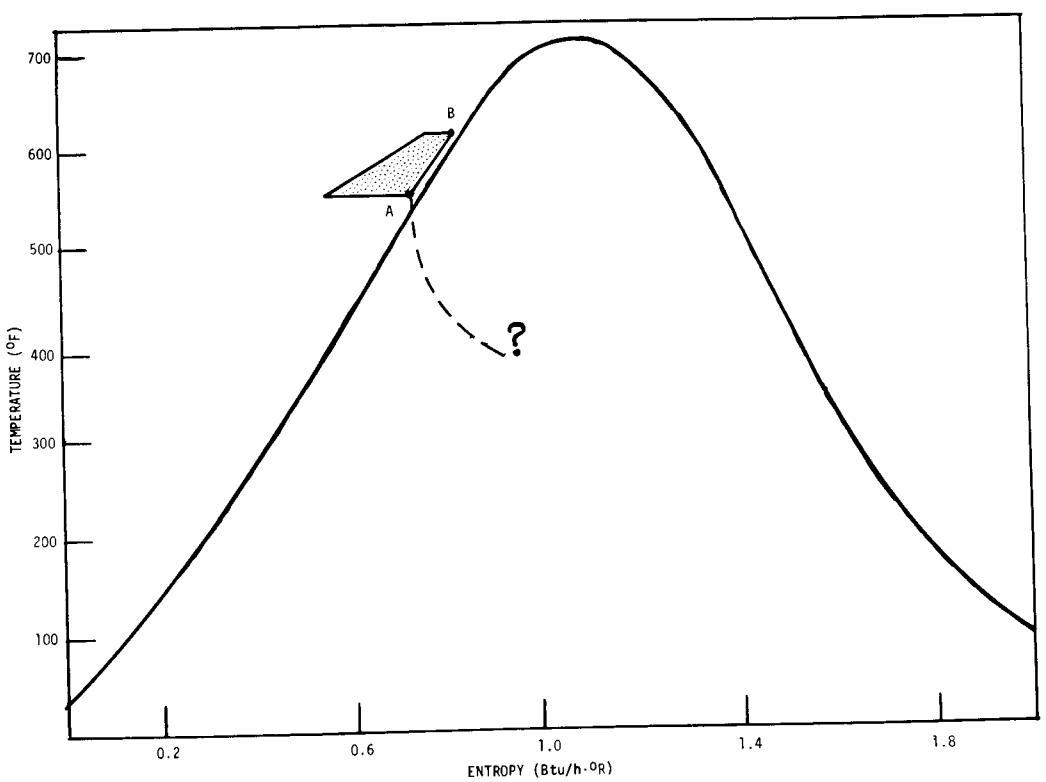


Fig. 4. Three Mile Island — Small LOCA.

We also need to ask questions about the training requirements necessary to help the operator understand the display. We should collect protocols from the designer and operator. We should ask the operator to imagine that a steam tube rupture has occurred and to tell us how he would interpret the information given on the display and what information he would seek next. We would try to understand the relationships of the pieces of information.

Then we should attempt to structure this as well as we could in terms of decision processes -- still at the paper and pencil stage. After that we would call on our supervisory control model (for only a portion of the plant - generic primary and secondary heat transfer). We would abstract the classes of information we believe are being presented by this display, and we would then add to the model an additional source of information (subcooling) as closely as possible to the way we think the operator understands the display.

By this process we would see whether in a dynamic sense this added information changes the way the plant is brought under control. We could perform sensitivity studies. I won't argue that this kind of modeling captures in exact detail the whole process, but it can allow us to evaluate whether it captures the essential features.

In summary, a combination of task analysis, protocol analyses, and understanding of cognitive processes, in conjunction with plant procedures, is how we would deal with this acceptance review.

Addendum C to Report of Working Group No. 4:

OPERATOR TRAINING AND OPERATING PROCEDURES -- ISSUES AND POSSIBLE ANSWERS FROM COGNITIVE SCIENCE

H. Halff
Office of Naval Research

In our working group, and throughout this workshop, we have been reminded that we should address the needs of NRC with respect to the operator in the control room of a nuclear power plant. We have placed those needs in two categories: those related to the procedures used in the control room; and those related to the training of the operators. What I report here is what we put together in two hours on a task that Don Norman thinks a formal task group will require six months to do.

Under the category of "Procedures" we first ask: Can we make the certification of proposed procedures more systematic, more precise, and more economical? If we can do this, then we would reduce the possibility of the regulatory agency and the industry going head to head on judgment decisions.

Next, we ask: Can we evaluate global approaches (for example, symptom-based versus system-based) to the design of procedures? As a cognitive psychologist, I do not think this is relevant, but it's the way NRC wants to go.

One thing cognitive psychology can do is to deal with procedures. We can use a formal (computer) language to represent procedures, and we can determine the human (machine) requirements for running procedures. We can also establish the sufficiency of a set of procedures for dealing with a set of problems. Finally, we can estimate the attentional resource demands for procedure execution. Of course, these are not sufficient alone, but they can reduce procedures to a technical exercise to help regulators.

Under the category of "Training" we ask: How can we determine requirements for system instruction, procedure instruction, simulator training, part-task training, simulator fidelity, novel instructional technology, etc.? How can we ensure the relevance of training? And how can we ensure the quality of training?

We would not ask questions like "how many hours?" or "what is the proper mix of simulator in training?"

Cognitive psychologists have built a body of empirical methods to evaluate training. We use qualitative, empirical methods to characterize stages of learning under particular regimes. We use discourse and cognitive task analysis to evaluate training materials and methods. We use cognitive models to establish knowledge requirements. And we use computer simulation to model learning and performance.

What differentiates novices from experts in real world problem solving? It has to do with knowledge representation. Powerful concepts are more central in the expert's knowledge structure, in his problem perception, etc. It is difficult to do a cognitive task analysis of an operator in the control room, but it is not difficult to determine how a student understands a diagram.

Cognitive models are important. Questions such as those cited here are being asked about them, but there are answers that people in cognitive psychology can supply.

Addendum D to Report of Working Group No. 4:

COGNITIVE MODELING IN PROBABILISTIC RISK ASSESSMENT/HUMAN RELIABILITY ANALYSIS

D. Embrey
National Centre of Systems Reliability, UKAEA

Cognitive modeling in PRA/HRA has many deficiencies, but it is generally applied as a way to turn qualitative issues into quantitative aids in decision making.

The current methods of modeling the human in PRA are "inverted." That is, the operator is modeled as a passive/reactive component at the bottom of the fault tree. Cognitive failures, such as decision making and diagnosis, that are higher up the fault trees are much more significant than simple errors. In cognitive models for PRA/HRA, we are concerned with the abnormal condition rather than the normal condition. The major concern is with human actions that initiate transients and disturb protective functions and with the degradation of ESF (engineered safety features) through scheduling and maintenance errors.

"Common cause" failures (of the human) typically are due to cognitive errors. The pressurized thermal shock scenario illustrates the importance of cognitive perspective. Many potential initiators and event sequences are human failure events, with sequences generally of 30- to 120-min duration. The operator actions are crucial in that they may prevent damage in a threatening sequence or they may cause damage in a benign sequence. There is conflict between undercooling and overcooling -- that is, whether or not to throttle the high-pressure injection system. To illustrate this, let's look at a possible accident sequence following a main steam line break of a typical plant. As detailed in the accompanying box, the sequences numbered 6, 9, and 10 all have human factor components.

MAIN STEAM LINE BREAK (ORNL Sequences SB1 to SBA2)

Summary Description:

The accident sequence initiates with a 34-in. steam-line break coincident with reactor and turbine trips. The forcing function for the accident is the delay by the operator in isolating the feed water flow to the affected steam generator, coupled with failure to throttle the high-pressure injection flow and delay in restarting one reactor coolant pump in each loop after 50°F subcooling is attained.

Initial Conditions:

1. Full reactor power.
2. Nominal temperatures and pressures in primary/secondary.
3. Decay heat: 0.5 times the ANS standard.
4. Pressurizer spray/heaters operate as designed.

Sequence of Events:

1. Reactor trips coincident with break of 34-in. steam line.
2. Turbine trips, turbine steam valves close.
3. Integrated control system attempts to run back main feed water, as designed.
4. Protection systems on hotwell, condensate booster, and main feed water pumps function as designed.
5. High-pressure injection actuates at set point (1500 psig).
6. Operator trips reactor coolant pumps 10 seconds after high-pressure injection actuation.
7. Emergency feed water pumps start at low main feed water pump discharge pressure.*
8. Emergency feed water system attempts to maintain 240-in. steam generator level.*
9. Operator isolates feed water to affected steam generator nine (9) minutes into transient.
10. Operator restarts one reactor coolant pump/loop ten (10) minutes after attaining 50°F subcooling in hot legs.
11. Pressurizer goes water-solid.
12. Pilot-operated relief valve opens at set point (2450 psig).
13. Safety-relief valves open at set point (2500 psig).
14. Pilot-operated relief valve/safety-relief valves reseat at their set points (2400 psig).

*Event is phenomenologically dependent.

Rationale:

Given a steam-line break (probability $\sim 10^{-4}$), the subsequent events have a moderate probability of occurring, are highly dependent upon operator action and interpretation of plant conditions, and lead to *severe* consequences with regard to thermal shock.

Calculation Restart Points:

1. High-pressure injection initiation.
2. Emergency feed water pumps start.
3. Attainment of 50°F subcooling in hot legs.

The regulatory position with respect to the pressurized thermal shock problem is that screening should be done based on the NDTT (the nil ductility transition temperature), that plant-specific studies should be performed for the affected plant, and that plant-specific fixes should be made when required. The plant-specific fixes are accomplished via operator procedures and training.

The major concern is that an event would occur that is so rare that the crew would become confused or misled and would perform in such a way as to cause severe thermal shock and repressurization. A cognitive operator model would facilitate analysis of conflicts and potential areas for operator action in rare transients.

The access and use of cognitive models (in a form acceptable to and usable by the PRA community) could lead to: (1) greater completeness of modeling through improved search strategies, (2) direct involvement of the systems analyst in operator modeling, and (3) the realization of the risk reduction potential of aids to cognitive functioning (and possible quantitative assessment thereof). Data collection strategies could be structured to allow capture of data on cognitive failures, and these data could be used for iterative refinement of models and to check the adequacy of the system model.

On the basis of the above, research in the following areas is recommended:

(1) Review cognitive models which are available now to assist the PRA analyst in qualitative modeling (short-term).

(2) Investigate how models can be classified, summarized, and presented for compatibility with PRA structure (short-term).

(3) Investigate use of dynamic models for "scenario generation" (mid-term).

(4) Consider use of models for data generation purposes (long-term).

(5) Investigate data collection requirements (mid-term).

(6) Determine the cognitive processes of the analyst. (Consider possible biases which could affect completeness of modeling.)

CLOSING REMARKS

CLOSING REMARKS

. . . by Thomas G. Ryan, U. S. Nuclear Regulatory Commission

I would like to thank Tom Sheridan for the opportunity to address this closing session of the workshop. I thank all of you for your thoughts, presentations and discussion of this timely but controversial subject. During the past three days and evenings, I have learned a great deal more about contemporary cognitive psychology and cognitive modeling, if not about their potential applications in NRC human factors work on control room operations. I hope that each one of you will be available in the future to advance the work that we have begun here.

In the short time that I have been with the NRC Office of Research, having previously been involved in military and highway safety human factors research and applications, I recognize the formidable task the NRC has ahead of it in developing a human factors technical base to adequately support the regulatory process, especially in the area of nuclear power plant control room operations. I must confess that I come to this workshop with some biases concerning cognitive modeling, having been weaned on the gospel according to Saint Art Siegel. With these thoughts in mind, I would like to comment briefly on the remarks made earlier in the workshop by Bob Pulliam (BioTechnology, Inc.), Don Norman (University of California, San Diego) and Joe Wohl (AlphaTech, Inc.).

First, what can the participants in this workshop do for the NRC in the future?

Bob Pulliam suggests that the NRC should make a commitment to cognitive modeling in the very near future to ensure that substantive progress in this arena occurs to support the regulatory process in a timely manner. I wish the NRC were presently in a position to make such a commitment. Unfortunately, the NRC is just now attempting to determine if, and how, human factors itself should be involved in the regulatory process. Since cognitive psychology and associated modeling techniques are primarily human factors tools rather than policy issues, I hope that with your assistance, we at the NRC can define the appropriate role(s) of cognitive psychology and modeling in a timely manner, coincident with our determination of the role(s) human factors will play in the regulatory process, especially as it applies to nuclear power plant control room operations.

Don Norman suggests that the NRC needs to define specific control room operator problems and requirements before determination of potential applications of cognitive modeling can be made. This is a good point. Quite often those of us in the human factors field, armed with our repertoire of ready-made solutions, go around looking for problems to fit those solutions, rather than the other way around. The participants in this workshop can be very helpful as NRC seeks to better understand the kinds of human phenomena (for example, information gathering, decision making, monitoring) that can be addressed through cognitive modeling, and begins to formulate appropriate questions concerning these phenomena within the context of nuclear power plant control room operations.

Second, what can the NRC do for cognitive psychology and cognitive modeling in the future?

Joe Wohl, spokesman for his working group, suggests that a task force be established to investigate the potential applications of cognitive modeling and associated techniques in control room operations. I agree and hope that my associates within the NRC Office of Research will make this a high priority for FY 1983. Therein, the investigation of potential cognitive modeling applications should be extended to the entire nuclear regulatory process and should involve participation by experts from academia, industry, utilities, and related regulatory agencies.

Again, I want to thank you all for participating in this thought-provoking workshop. I hope we all have the opportunity to meet again under circumstances such as these in the near future.

* * * * *

. . . by Kenneth G. Murphy, U. S. Nuclear Regulatory Commission

After spending three days listening to cognitive modeling experts, I have three observations that can be epitomized in the words "compatibility," "communication," and "cooperation."

At present, I see little *compatibility* between past cognitive modeling efforts and the needs of nuclear power plant safety. After all, we are not landing an aircraft where the time frame of emergencies can be fractions of seconds to minutes; nuclear plant emergencies, as in the case of Three Mile Island, stretch out over hours and days. This fundamental difference must influence, in large measure, the direction of cognitive research in the nuclear industry. For example, for the slow-moving nuclear emergencies, operator-machine interfaces may not be as critical as clearly written procedures and access to expert direction. We are not talking about a pilot with a joy stick in his hand.

We must also make a second contrast between the aircraft and the nuclear plant. Both have about the same number of operators (three in most cases); however, while a nuclear plant is operating, more than 300 additional people are at work modifying, maintaining, and testing equipment. A recent study of serious accident precursors shows that maintenance-induced human errors are the highest single cause of multiple component failure that bring down redundant systems. I expect, therefore, that a balanced research program must consider not only the operating staff and how they react in emergencies, but also the maintenance crew that is under pressure to keep equipment operational.

I see very poor *communication* between the nuclear safety engineer and the cognitive modeler. Nomenclature is a barrier and the two disciplines are complex. In order to start to deal with the real human problems in operating nuclear plants, the barrier to interdisciplinary understanding must be removed. We need engineers that think cognitive and cognitive specialists that think engineering. Nothing useful can happen before these barriers are broken. Research efforts should allow for this type of assimilation.

As we start research in human cognition as it affects nuclear plant safety, we need to appraise its potential value and where it can do the most good. Plant organization, worker attitude modification, training, and procedures might be the first areas to determine feasibility and practicability of cognitive research. A critical look at the deficiencies in current probabilistic risk analysis (PRA) as it concerns human activities should be high on the list of initial research. After all, PRA has helped the safety engineer prioritize his reviews of equipment and thus it should also be of help in prioritizing cognitive research. But before that can be done, the treatment of the human in PRAs must be substantially improved. (Because quantification of human error rates will, no doubt, continue to be problematic, ways must be sought to gain the needed perspectives without a great dependence on PRA quantification.)

The first tentative steps in cognitive research demands *cooperation* between the cognitive expert and safety engineers. In order to gain a balanced viewpoint and to ensure useful products, the initial research programs should be designed to ensure a team effort between these diverse disciplines.

* * * * *

. . . by Leo Beltracchi, U. S. Nuclear Regulatory Commission

This workshop has been beneficial as it has presented the state-of-the-art on cognitive modeling. From the workshop, I have concluded that cognitive modeling is needed to assist in the process of designing control rooms.

The process used to design reactor pressure vessels is well developed. The strength of materials, the design basis of the core, the static loads, dynamic loads, thermal stresses, etc. are well known and used in the design of a pressure vessel. In general, pressure vessels have been designed and built and they are operating successfully in containing the process.

In contrast to pressure vessel design, the design process for power plant control rooms with human operators and interfaces to process control systems has been less successful. A major issue is the operators, as they have been backfit into the control room after the control room has been designed and built. A cognitive model serves as a design tool, thereby allowing the designer to integrate the operator in the control room design process. Dr. Norman made a very good point when he suggested that cognitive factors be considered during task analysis. These tools would also be useful for the detailed control room design reviews.

In conclusion, with an improved control room design process, incidents such as TMI-2 can be avoided.

* * * * *

. . . by Thomas B. Sheridan, Massachusetts Institute of Technology

As we conclude this workshop, it appears that we have a concensus on several points. We agree that operating a nuclear power plant is largely a cognitive process. We also agree that there are many approaches available for modeling the cognitive activity of the operator — some qualitative and some quantitative, some normative and some descriptive, and some probabilistic and some deterministic.

We are not at the point where we can offer a complete model, so what we are doing is selecting and applying pieces of available models. Such applications can be useful in a number of ways: (1) to generalize what we think we know about how operators think; (2) as an aid in defining the problem and identifying what is important; (3) as a guide for collecting data; and, in limited cases, (4) to predict.

We know that we have a dilemma about operator errors. On the one hand, we want to prevent them; on the other hand, we want to allow the operator some personal space to try the system, to learn from small failures, and thus to gain understanding and confidence.

Finally, we recognize that we are still far short of where we would like to be in providing NRC with useful cognitive models. As the saying goes, more research is needed!

I thank each of you and very much appreciate your efforts.

* * * * *

. . . by Roger A. Kisner, Oak Ridge National Laboratory

As a control engineer, I see a need to couple humans and machines to form a total system capable of delivering the improved performance and reliability being demanded for nuclear power plant operation. However, unless cognitive modeling can help to engineer an improved total system, it will not be useful to those of us designing, constructing, evaluating, or operating complex systems. The discussions these past three days have triggered some thoughts on the application of cognitive modeling to system modeling and design, although much remains unclear to me.

Now that the workshop is over, it is easy to envision other ways that it could have been organized; however, in this limited three-day time span and with this very diverse group of individuals, some of you meeting for the first time, I believe that the format we have used has been very productive, and the proceedings of this workshop will provide useful information to the NRC in forming a plan for future research.

I thank each of you for accepting our invitation. As we leave, I think all of us will agree that similar dialogs should be carried out in the future as we work together toward solving some of the problems confronting the NRC in the area of cognitive modeling.

REFLECTIONS

REFLECTIONS

ON PROBLEMS RAISED IN LETTER OF INVITATION
(July 26)

Erik Hollnagel
Institutt for Atomenergi
OECD Halden Reactor Project

The following are some brief comments to the problems raised in the letter of invitation to the Workshop on Cognitive Modeling of Nuclear Power Plant Control Room Operators. Even though previous engagements will prevent me from attending the workshop, I would like to give my contribution in this way. The topic is central in the ongoing research on operator performance in nuclear power plant control rooms, and I see the workshop as an important step towards integrating much of this research. The topic fits well with the activities of the Halden Project in this area, and my comments thus reflect the basis of this work.

1. On Cognitive Models

Let me begin by disagreeing slightly with the title of the workshop. The concern should not be with models of operators, but instead with models describing and explaining operator performance. We are interested in models of human activity, rather than in models of humans per se. Furthermore, we are not really talking about models, but rather about a mixture of assumptions, hypotheses, and (possibly) theories. But I accept that the term "model" has come to be used for this conglomerate, and shall stick to that. By "model" we thus mean any kind of systematic analogy which we use for consistent descriptions and explanations of operator activity.

Another point is that, properly speaking, we are not discussing cognitive models, but rather models about cognition. The models can be of any kind and need not in their substance (i.e., their terms) have anything to do with cognition. But they are about cognition, i.e., those parts of human activity which refer to human cognition. We are thus really talking about models of those aspects of human activity that can be explained as cognition in a wide sense. This, by the way, indicates that Question 6 below should be extended to cover also the proper definition of the terminology.

Types of Models. When this much has been said, we may go on to list the common types of models of operator action that have been used:

1. Mathematical models.
2. Engineering models.
3. Computer models/computer programs.
4. Information-processing models. [These can actually be of two types, emphasizing either structures (various types of memories or processing stages) or processes (for example, production rules).]
5. Conceptual models. [These models are expressed in and refer to high-level cognitive functions, such as strategy selection, problem solving, etc.]
6. Psychological models. [These models include psychological theories about human cognition.]

Distinctions may further be made between descriptive and normative models, process and structural models, and generic and specific models. These dichotomies cut across the various types of models.

Models and Performance Types. It is important to note that a type of model may be particularly useful for a specific type of performance, using, for example, the distinction between skill-based, rule-based, and knowledge-based performance (the SRK-model). If we are concerned only with skill-based performance, mathematical or engineering models are often the most precise. But if we are concerned with a knowledge-based type of performance, as, for example, diagnosis, a mathematical model may be inadequate, and a conceptual or psychological model should be used instead.

We thus need a repertoire of specific models for characteristic activities, but we also need models that can describe and explain performance on an overall level. We particularly need models that describe the learning of a specific activity and the transition from one level of performance to another. Otherwise one may feel tempted to generalize a specific model to a range of phenomena for which it was not intended.

2. The Questions

After these introductory remarks, I shall comment on each of the nine questions mentioned in the letter of invitation.

Question 1: *Does the NRC need information about cognitive models for regulation of nuclear facilities and what are the major purposes of these models?*

The two parts of this question should be reversed, since we must know the purposes of the models before we can say anything about whether they are needed for regulation. Given the purpose as described above, and the state of psychological theories in general, I would say that the question is premature. It is not possible to make them sufficiently specific and detailed to support regulations. They may, however, be used as support for guidelines in various areas (cf. Question 3 below).

Question 2: *What reactor operator cognitive behaviors do we need to predict? Can these behaviors be listed in order of importance?*

I do not think we need to predict activities (certainly not "cognitive behaviors"), and if the experience from psychology is worth anything, I further do not think that it is possible to predict activity at any useful level of detail. It is, of course, the ultimate purpose of any theory to be able to give predictions of what will happen under specific conditions. But there are few psychological theories which can do that to any notable extent. For instance, the attempts to predict human decision making have failed miserably. And models of operator action are even less developed than theories of decision making. What we can do is instead to identify the possibilities for influencing activity, and hence reduce the variability of human performance. A good model of human activity may indicate how and where/when information can be provided which will change the course of actions. And that is far more useful than trying to predict behavior in a deterministic sense.

The second part of the question should accordingly be answered with the types of activity that we want to gain control of. These types may be derived from the performance models, e.g., the step-ladder model.

Question 3: *How would such predictions or other outputs from cognitive models be applied by the NRC in the areas of: (a) plant design, (b) procedures, (c) management, (d) training, (e) and licensing or other?*

(a) I take plant design to refer to the part of the plant that is relevant for the operators, in particular, the control room. If so, we may need the cognitive models as a basis for control room design in general, particularly information presentation and the exercise of control. That, of course, involves the whole of the design process as suggested by Cognitive Systems Engineering (CSE).

(b) The use of cognitive models in this area is obviously important, as shown by the experience from the EPRI/IEOP/TSAM* project. We can use the models to design procedures to be easier to understand and to use.

(c) Again, I take this to refer to the management of control room operators. This seems to be more in the line of the social psychology of small groups, where a considerable knowledge exists. But that must be combined with models of individual functions. This aspect also highlights the importance of individual/personal factors which are conspicuously absent from most cognitive models. We should, perhaps, strive to extend models of operator activity to include more than what is traditionally called cognition!

(d) The use of cognitive models for training is also important, cf. the experiences from the Scandinavian NKA/LIT† project. We can use the models to specify training (both contents and form) and to analyze the outcome of training.

(e) This is related to *(d)*. The use of cognitive models might here be as the background for a Cognitive Task Analysis (CTA), based on CSE principles. This may again be used to suggest criteria for licensing additional to those derived empirically.

Question 4: *What are the principal cognitive modeling approaches in use now, and how are they used in other applications or fields?*

This has partly been answered by the types of models mentioned above. But a more systematic treatment is, of course, required. Other applications are rather scarce, since cognitive psychology and cognitive science mostly have confined themselves to the world of the psychological laboratory. Some recent attempts to break this has been worked on, e.g. slips of action, visual momentum, and decision making. But what we generally need is good models for the variability of human performance.

Question 5: *At what stage of development do models become useful?*

The development of what? If we are talking of the control room as an integrated whole, cf. the above, models are clearly relevant from the very start. No design or development is made without a conceptual background, and the more explicit this is the better. It may also greatly facilitate the later evaluation or validation of the design.

Question 6: *What taxonomies apply (e.g. normative vs. descriptive, control vs. decision, mathematical vs. verbal, deterministic vs. stochastic)?*

The question should rather be: What part of the philosophy of science do we implicitly refer to when using cognitive models? What we really need is an epistemology of process control! Once we have made it clear what the background is, the kind of taxonomy and the set of descriptive dimensions we can use will come naturally.

Question 7: *What data are presently available to use for generating models? What data are missing? Do we have means of collecting needed data?*

The various sources of data have been systematically described in Hollnagel, Pedersen, and Rasmussen: *Notes on Human Performance Analysis* (Riso-M-2285). The data that we presently have available come from a variety of sources, some from nuclear control rooms, some from other areas (e.g., trouble-shooting in electronic equipment). What we need are more data from normal performance conditions, either from field studies or from intensive studies using simulators. There are a lot of data from training simulators that are simply thrown away, although it would require little effort to register them. We also need a consistent classification which is easy to use in a variety of conditions. We know where we can get the data we need,

*EPRI/IEOP/TSAM = International Evaluation of Operating Practices/Training Simulator Analysis Methods.

†NKA/LIT = Joint Scandinavian Research Project on Human Reliability.

but it requires an integrated effort to actually collect, and analyze it.

Question 8: *What sources of data can be used to calibrate or validate cognitive models of nuclear power plant operators (e.g. anecdotal data, LERs, simulators, etc.)?*

We can probably use any source of data, as long as the observations are systematically made and the analysis is consistent. We already have a general scheme of analysis which can be used for that, cf. Riso-M-2285. But we cannot validate a cognitive model, only calibrate or verify it. This process is, of course, easier the more specific the model is.

Question 9: *What are the current obstacles to: (a) developing the status of cognitive modeling of nuclear power plant operators, (b) calibrating and validating the models, (c) putting the models to use, (d) and learning from experience?*

What we generally need is a coordination of the efforts of data collection, analysis, and model development. Some attempts have been made with notable success, e.g., the EPRI/IEOP project and the Scandinavian NKA/KRU[‡] and NKA/LIT projects. We should endeavor to coordinate the activities at various sources (e.g., research simulators, training simulators, LERs, etc.), so that the same problems are addressed insofar as possible. Concerning *(c)*, putting the models to use, I think it is important not to promise too much but to be realistic in indicating the usefulness of cognitive models. It is, after all, an approach which is trying to establish itself.

[‡]NKA/KRU = Joint Scandinavian Research Project on Control Room Design.

3. Epilogue

The questions raised in the letter of invitation obviously reflect the practical engineering approach, and my comments probably reflect a more academic, but hopefully not impractical, point of view. My general comment is that we must not be too hasty in applying cognitive models of operator action. We have a solid basis for these models in the work that has been done over the last decade or more. We have developed the models far enough to be able to discuss them with others, and to offer them as a useful tool. But we need still more work, particularly in getting a broad empirical support, before we can confidently claim that we have a complete theory of operator action.

ON WORKING GROUP NO. 1
(August 24)

N. Moray
University of Toronto

1. I agree strongly with Shelly Baron that even relatively simple models can be of predictive use because the situation constrains the operator, and hence the degrees of freedom in his behavior are drastically reduced.

2. The models which we presented are all of proven value in predicting human performance in certain (more or less) well-defined situations. The critical question is "What are the boundary conditions for each model?" not "Is it a true (and complete) theory of human behavior?"

3. A *most* important point is that the effect of subjective expected utilities is universal. If the operator's perception of what is "good" behavior -- for example, the relative acceptability of false alarms and misses as required by management or his co-workers -- changes, then his performance will change *even if no changes occur in the quality of the displayed information.* Take, for example, the "speed-accuracy tradeoff." If the operator believes it to be in his interest to look at as many displayed variables in the shortest possible time, the accuracy of his perception will decline. *The efficiency of operator information processing* ***cannot*** *be predicted without a knowledge of his subjective payoff matrix.* This is true of *all* behavior at any level (skill, rule, or knowledge based).

4. If an operator's level of "arousal" or motivation is too low, increasing it will increase both speed and accuracy. At any fixed level of arousal or motivation, speed and accuracy are inversely related. This effect can be modeled using a modification of Sequential Decision Making.

5. The models as mentioned are all plausible candidates for modeling a subtask (part) of the behavior required of a nuclear power plant control room operator. I make this claim in the light of my experience in evaluating human factors work in U.K. nuclear power stations -- that is, not merely on supposed analogy with other industries with which I am familiar.

6. It is possible to measure operators' subjective payoffs directly by suitable questionnaire and scaling methods. Such estimates are also relevant to assessing the effectiveness of training, management/operator relations, and probabilistic risk analysis.

7. The NRC would have received a more convincing and useful picture of the techniques offered by psychology if we had been asked very specific questions -- see, for example, the questions in NII.ONSWG(82)P.1. But the psychologists would have to be briefed on the *nature and requirements of specific tasks.*

8. If a "task force" is set up, it must be arranged so as to give the psychologists adequate first-hand experience of nuclear control room "natural history."

9. The NRC should not concentrate too closely on "cognitive models." What is required is an intelligent and informal use of *anything* relevant from psychology. Cognitive modeling is a small subset of human factors which is an interdisciplinary mix of psychology, engineering, physiology, and some other minor disciplines. What are required are not "cognitive models," but "man-machine-*system*" models, whether "cognitive" or not.

10. It may be worth noting in the light of the agonizing which takes place about how much automation to include, that most European NPPs have more automation than any U.S. NPP, and the Canadian reactors (CANDUs) are fully automated and have the best load factors in the world.

ON WORKSHOP THEME QUESTIONS
(August 24)

R. Pulliam
BioTechnology, Inc.

These comments respond to the "theme questions" posed for the Workshop on Cognitive Modeling of Nuclear Plant Control Room Operators.

NRC clearly does need cognitive models and information about cognitive function to support its regulatory role. Nuclear power plants (NPPs) are built, maintained, and operated by people, whose behavior determines the usefulness and safety of the NPPs. Cognitive function is the critical determiner of behavior; until recently cognitive function has evaded scientific treatment. NRC should support R&D in applied cognitive science so that it can more effectively characterize the NPP operator (and other people in the NPP setting), and so that it can regulate those conditions which lead to safe working behavior. As John Senders observed, "An adequate understanding of human behavior is as necessary as an adequate understanding of plant behavior."

NRC's investment in cognitive science should recognize, however, that early applications are not usually to be expected. Some small-scale and low-technology applications seem feasible in the early future, but the big payoffs are several years away. Early investment should plan for a basic science and exploratory phase. In the long term, however, NRC will absolutely require cognitive science in order to understand the role of man within the high-technology control environments that are coming. NRC can either wait, and in ten years seek expertise from the universities and other technology areas, or it can begin now so that its needs will be met more effectively when the time comes.

Why Cognitive Science?

In the modern workplace, and particularly in the NPP control room, we are faced with the question: How can we usefully characterize man in his role as the more complex component of the man-machine control loop and as the transfer function which inserts human purpose into the system? Now we appear to have a new discipline with the explanatory and predictive power needed to provide this description. We need to "model" man in order to describe, understand, predict, and diagnose. Our models need to be "cognitive" because the key problem behaviors are at the levels of rule-based and knowledge-based (rather than skill-based) behaviors. Already this new science has provided models which make it possible to document and manage a category of work activity which we could not deal with before — that growing majority of work which manipulates information, rather than physical things.

Nature of a Program

So what should NRC do to exploit cognitive science? Only NRC and the Congress can decide the proper scope of a program, within NRC's charter and objectives. But some suggestions are offered:

(1) The history of R&D, and particularly of "new science" programs, shows that early applied efforts will probably be disappointing. The last time behavioral science promised a major contribution to society, it was going to revolutionize learning with "Skinner boxes." This time we should be more conservative in our expectations, and in what we promise. During their early years, most new programs produce only (a) information, and (b) trained people. It is this second resource which is most necessary to later success. NRC should decide how much it can afford to invest, on a continuing basis, in early R&D, with the understanding that this investment is made primarily to develop a corps of professionals. In a few years, NRC will desperately need a stable of practitioners who know (a) NRC, (b) NPPs, and (c) cognitive science.

(2) A key insight is Don Norman's "Cognitive Engineering." To me this suggests a developed profession with experience in converting theory to applied uses, and with a collegial memory. Until you have an institutionalized profession there is no efficient communication of research, and many applied successes are lost for want of institutional memory.

(3) Sheridan makes at least two points worth emphasizing: (a) Cognition should be broadly defined to include sensory and motor aspects. We need models

of man which are fully able to interact with models of the plant, and of the external world as a whole. (b) "Mixed models" have high utility. Simulations are achievable which are part quantitative and part verbal (descriptive, qualitative), and which along with artificial intelligence (AI) can be used to validate or exploit model structures.

(4) For the first few years, NRC should not invest in global, quantitative models of human function. Such models have a low probability of success. Theorists still disagree on the most basic questions of cognitive structure, and thus we have no basis for a formal model. Similarly we have few reliable quantitative data, and therefore lack the means to quantify a model. Neither are plant models immediately achievable. Tom Sheridan reported that the plant model may be a 47th-order differential equation. Any closed-loop man-machine model would suffer from the cumulative error of an inadequate operator model driving an inadequate plant model, which feeds back increasingly flawed data to drive the operator. Such a general model of the NPP operator would probably run, and might appear to be a success. But all that this would demonstrate is that we can write software even to false or incomplete assumptions. The modeling effort would be costly, and would be of little scientific or applied interest. Rasmussen remarked that "...in the rush to get quantitative models we may forget how to categorize...and define boundaries."

So for a time NRC should limit itself to projects of high scientific merit, and to a few low-cost, high-probability-of-success applied projects. Several such projects were described by speakers at the workshop. Some that seem useful are reiterated here:

- Better formal and descriptive models are needed. These should be developed in applied settings, where they can be validated against real-world human performance and used to explore applications. The objective would be to find more sophisticated model structures which are useful and/or valid.
- Part-task quantified models are achievable. These should be indexed to real tasks, use short program runs to limit cumulative error, and be calibrated against reality.
- Human subsystem models are achievable. For instance, Dave Woods discussed the close coupling of visual attention to the subject's need for data. This could be usefully embodied in a model.
- Certain AI applications appear to have an early payoff, including knowledge-based expert systems for decision aiding. AI may be only debatably a part of cognitive science, but that is academic. AI and cognitive science will become inseparable when it comes to matters of application.
- Models can inform task analysis. Several speakers observed that models and task lists can be developed interactively. As tasks are discovered, they can be used to specify structures or characteristics of the model. As the model develops, it provides a descriptive language for the task list, and guides the analyst in discovering tasks. Such models are particularly effective in capturing cognitive function and information flows, data which are usually missing from task analyses.

(6) Some commitment to "pure science" is desirable. Politically, this may have to be justified in terms of applied objectives. However it is to be justified, such support should be directed to where it is most effective in broadening the professional community of cognitive science. This may mean *not* supporting the most deserving researchers — those who have pioneered the field. However unfair this may seem, the really great and established researchers are certain to be supported anyway. Highly focused funding could result in a dilution of their attention and in the concentration of expertise in a very few universities and business firms. I would suggest institutional priorities as follows:

- Academic departments with a potential, but not an established tradition, in cognitive science.
- National laboratories, rather than private firms, so that more discoveries and software remain in the public domain.
- Business firms with developing, but not leading, capabilities.
- All others.

In a few years, when an applied technology is developed, it will be time to reverse these priorities and to give the action to those universities and firms which are most able to produce. At that point, there will be a broader range of choice.

(7) About validity: Several speakers suggested that we should not ask that a model be representative of cognition, nor exhibit validity except at its end-product point; that is, a model need not resemble cognition in its processes if it predicts performance. This is certainly true of engineering models. If a model predicts an optimum wing design, we do not worry about what aerodynamic theories it exercises. But this is *not* true for cognitive models developed *to advance the science* of cognition. For the most part we now look to those models for explanatory power. We do not expect to use them for quantitative prediction, or if we do, we do not yet know what they can predict or at what points in the modeling sequence data may be drawn off. Therefore, we should strive for models which have formal validity, in the sense that they represent cognition in structure and exhibit concurrent validity at the nodes of the model. For the moment, we should expect behavior to be predicted more reliably by expert judgment than by models. Later we may derive, from the theoretical models, "engineering" models which predict human performance for real-world applications.

Access to NPP Data

If cognitive scientists are to assist the NPP industry, they will need more than money. A major obstacle is the difficulty of access to real world data. In any

realistic program, NRC must address the question of access to subject matter experts, nuclear plants, simulators, and vendor design shops. This access is necessarily limited, and could become a constraint on R&D which is more limiting than the costs of research.

Summary

Yes, NRC should support cognitive science, and should expect powerful results (after some years) from a better understanding of human cognitive functions. This will depend on:

- Maturing the technology.
- A larger community of professionals.
- Not committing to a single direction prematurely. NRC should keep on asking the question asked at this workshop: "What is the best direction to go?"
- Participation of the NPP industry as a field data source.

ON MODELS, AVIATION, AND AUTOMATION
(September 13)

R. Curry
NASA Ames Research Center

What is a (Cognitive) Model, Anyway?

At the workshop I sensed a certain lack of clarity of what a model is and what it can do. Perhaps the following definition will be useful:

> *A model is a "thing" of which we ask questions about the real world. A model is adequate if the answers we get from the model are close enough to the answers we get in the real world situation.*

This definition seems to explain the lack of communication (other than vocabulary) that exists between model builders: some people are trying to answer questions about how the human's innards actually work, and others are trying to describe behavior under different task/environmental conditions. If you reexamine the workshop discussion about various models, I think you will see that most disagreements can be traced to a misunderstanding of the goals (questions to be answered) that were assumed by the various sides of the discussion.

A second reason that this definition of a model is useful is pedagogical. It can be pointed out that the NRC commissioners and staff have already been using cognitive or human performance models as a basis for their many actions regarding human factors. There also may be other reasons for building models (see, for example, Bill Rouse's book and the paper by Dick Pew and Shelly Baron for the Baden Baden conference).

Some Examples

With this definition, a model, cognitive or not, can take a wide variety of forms:

A Statement. Consider the following statement about Short Term Memory (STM) extracted from Don Norman's workshop paper:

> *Meaningful information is much more easily retained than nonmeaningful or unfamiliar material.*

This model, and the other statements on STM in Don's paper, can answer questions about how much and under what conditions information can be remembered for short periods of time.

An Equation. The following equations describe the time required to position a cursor on a CRT using two devices, a mouse and a joystick:

$$T = 1.03 + 0.096 \log (D/S + 0.5) \quad \text{mouse}$$
$$T = 0.99 + 0.220 \log (D/S + 0.5) \quad \text{joystick}$$

where D is the distance to the target, S is the size of the target, and the logarithms are base 2. These equations (models) and other devices are described in the paper by Stu Card et al., *Ergonomics*, **28**, No. 8, 1978. The article also has equations (models) that describe the learning time with the different devices. There is much behind these equations, and the article must be consulted to determine if the answer to the question "How long does it take to position the cursor?" will be valid for any given real-world situation.

Manned Simulation. A manned simulation is also a model, since it will provide answers to certain questions; confidence in the answers is usually higher than for other models because "everything" has been taken into account. This increased confidence is not always justified. However, may times it is true that if a task cannot be done in a simulation, it probably cannot be done in the real world.

Similarities in Aviation

Ken Murphy said it most succinctly: "Controlling a nuclear reactor is not like flying an airplane." True. But there are some interesting parallels involving automation that may be useful, perhaps as models.

SPDS* and Flight Directors. About 30 years ago the aviation industry recognized the high information-processing load of instrument flying: assimilation of information from a variety of sources, interpretation, and appropriate control action. These activities are similar to those of a plant operator dealing with a disturbance. The flight director was developed to compute and display steering commands based on measured position, velocity, altitude, rate of climb, etc. The flight director has reduced the processing demands of instrument flying and this allows the pilot to perform other tasks such as monitoring. Presumably a SPDS system would offer these advantages too.

Sympton-Based-Procedures and SBOs. Until ten years ago, the training of pilots and flight engineers included a large amount of detail about hydraulic, pneumatic, electrical, and other systems. This is still done for some operators who fly to remote areas without adequate maintenance, but the majority of training is done by Specific Behavioral Objectives (SBOs), i.e., symptom-based procedures: "if this light comes on, push that button." These procedures are designed to stabilize the situation and take advantage of the onboard system redundancy to continue safely or land as soon as necessary. Automating the responses by performing the decision making ahead of time has reduced training requirements and, overall, produced better and more consistent crew responses. If system-based procedures are that good, why not replace the operator with a completely automatic system? This is the key issue: how do we design the system and procedures and how do we train the operator so that he will recognize a new situation not covered by existing procedures and take appropriate action?

Automation Is Not a Panacea

Automation (e.g., SPDS, flight directors, system-based procedures) can perform well when properly designed and implemented. However, Earl Wiener and I have reviewed aviation accidents, incidents, and experiences and found that more automation is not always better (*Ergonomics*, **23**, No. 10, 1980). An important lesson to be learned from these real-world examples is that there have been some unexpected human factors side effects to automation, for example, increased training requirements, loss of skill or knowledge, operator complacency, defeating of alerting and warning systems, ignoring alarms, etc. In addition, attempts to eliminate human error by automation have introduced human error elsewhere.

System design guidelines for automated and semiautomated systems have been proposed (see our article for some), but no one has the definitive answer on how to design the interface. It is known, however, that we should be asking more questions about these side effects of automation, i.e., we should be examining automation for its human factors liabilities as well as its benefits.

ON POST-WORKSHOP THOUGHTS
(September 9)

J. M. Christensen
General Physics Corporation

I want to take advantage of the invitation to send in any "post-workshop" thoughts that may have occurred to us.

In reviewing the papers and my notes, I was impressed by all of the speakers but I feel that the ideas presented by Rasmussen, Sheridan, and the Bolt, Beranek and Newman (BBN) group hold promise of being most useful to NRC in the near future.

Rasmussen's model has an intuitive appeal and I like his remarks about different models and treatments for different purposes. He also feels that the skills and rules parts can be treated quantitatively now or in the near future. It is not surprising that he feels qualitative treatment will be required for the knowledge part of his triumvirate, but many good models start out that way.

Sheridan's "internal – external" model also has considerable intuitive appeal. If we can determine the internal model that an operator possesses and then construct a system that is congruent with that internal system, errors should be reduced and general efficiency increased. The model is also sensitive to the differences between training and human engineering. Training may be considered an attempt to modify the internal environment so that it is congruent with the external environment. Problems arise in determining the exact nature of the internal environment at any instant in time; internal models will vary from person to person (e.g., one operator may operate at Rasmussen's "rules" level while another may operate at Rasmussen's "knowledge" level), and even within individuals as their physiological state changes with physical fatigue, emergencies, etc. and their psychological state changes with training, mental fatigue, etc.

The BBN model is an attempt to apply an existing supervisory control modeling approach to nuclear power plant operations. The model consists of block diagrams

*SPDS = Safety Parameter Display System.

and appropriate algorithms for demonstrating feasibility of the approach. Time-based simulations that include dynamic representations of the system, the operators, and the environment are included. Operators are considered as elements in a closed-loop system with well-developed information processing, decision making, and implementation capabilities. A task analysis is important in the definition of jobs in the model. Stress because of task overload during a specified time period can be handled. As conditions change, their effects are reflected in changes in the "activity time lines" — overload, for example.

The model is vulnerable and, in my opinion, merits evaluation with real data, leading to further model improvement. I consider it the best of the modeling approaches for NPP work that I heard at the meeting.

I detected a certain reticence at the conference to proceed with any of the models presented. In my opinion, this is a mistake. After a certain amount of development, models need testing, followed by modification, followed by more tests and so on until the model has reached a stage that satisfies the requirements that led to its development in the first place. I believe that application of this time-tested process to the BBN model, for example, would lead to a refined model that would satisfy your requirements.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Adams, J. A., "Letter to the Editor," *Human Factors Society Bulletin*, Santa Monica, California (June 1982).
2. Aiken, L. S., and D. R. Brown, "A Feature Utilization Analysis of the Perception of Pattern Class Structure," *Perception and Psychophysics* **9**, 279-83 (1971).
3. *Aircraft Accident Report No. NTSB-AAR-77-1*, National Transportation Safety Board (NTSB), Washington, D.C. (April 27, 1976).
4. Anderson, J. R., *Cognitive Psychology and Its Implications*, W. H. Freeman, San Francisco, California (1980).
5. Arnoult, M. D., "Prediction of Perceptual Responses from Structured Characteristics of the Stimulus," *Perceptual and Motor Skills* **11**, 261-68 (1960).
6. Bainbridge-Bell, L. H., "The Standardization of Symbols and the Arrangement of Electronic Circuit Diagrams," *Journal of the British Institute of Radio Engineers* **13**, 339-53 (1953).
7. Baron, S., "A Model for Human Control and Monitoring Based on Modern Control Theory," *Journal of Cybernetics and Information Sciences* (Spring, 1976).
8. Baron, S., C. Feehrer, R. Muralidharan, R. W. Pew, and P. Horwitz, *An Approach to Modeling Supervisory Control of a Nuclear Power Plant*, Report No. 4811, Bolt Beranek and Newman, Inc., Cambridge, Massachusetts (July 1982); also NUREG/CR-2988 (November 1982).
9. Baron, S., C. Feehrer, R. Muralidharan, R. Pew, and P. Horwitz, *A Framework for Modelling Supervisory Control Behavior of Operators of Nuclear Power Plants*, Paper No. 2, these proceedings.
10. Baron, S., G. Zacharias, R. Muralidharan, and R. Lancraft, "PROCRU: A Model for Analyzing Flight Crew Procedures in Approach to Landing," *Proc. of Eighth IFAC World Congress*, Tokyo, Japan (August 1981).
11. Bartter, W. D., A. I. Siegel, and P. J. Federman, *Job Analysis for Instrument and Control Technicians in Nuclear Power Plants*, Applied Psychological Services, Wayne, Pennsylvania (1982).
12. Bateson, G., *Mind and Nature*, Elsevier-Dutton Publishing Company (1979).
13. Baum, A. S., and C. G. Drury, "Modeling the Human Process Controller," *International Journal of Man-Machine Studies* **8**, 1-11 (1976).
14. Beltracchi, L., *A Qualitative Operator-Process Introduction Model*, paper presented at Second International Computer Engineering Conference and Exhibit, American Society of Mechanical Engineers, San Diego, California, August 15-18, 1982.
15. Bennett, J. L., "Spatial Concepts as an Organizing Principle for Interactive Bibliographic Search," pp. 67-82 in *Interactive Bibliographic Search: The User/Computer Interface*, D. E. Walker (Ed.), AFIPS Press, Montvale, New Jersey (1971).
16. Biederman, I., "On the Semantics of a Glance at a Scene," in *Perceptual Organization*, M. Kubovy and J. R. Pomerantz (Eds.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1981).
17. Brooke, J. B., and K. D. Duncan, "Effects of System Display Format on Performance in a Fault Location Task," *Ergonomics* **24**, 175-189 (1981).
18. Bruns, R. A., A. C. Bittner, and R. C. Stevenson, *Effects of Target Size, Target Contrast, Viewing Distance on Dynamic Televisual Target Detection and Identification*, TP-72-24, Naval Missile Center, Point Magu, California (1972).
19. Burton, R. R., J. S. Brown and J. de Kleer, "Pedagogical, Natural Language and Knowledge Engineering Techniques in SOPHIE I, II, and III," in *Intelligent Tutoring Systems*, D. Sleeman and J. S. Brown (Eds.), Academic Press, London (1982).
20. Carbonell, J., "A Queueing Model for Many-Instrument Visual Sampling," *IEEE Trans.*, HFE-7, 157-164 (1966).
21. Card, S. K., *Studies in the Psychology of Computer Text-Editing Systems*, Ph.D. Th., Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania, May 1978.
22. Card, S. K., W. K. English, and B. J. Burr, "Evaluation of Mouse, Rate-Controlled Isometric Joystick, Step Keys, and Text Keys for Text Selection on a CRT," *Ergonomics*, **21**, 601-13 (1978).
23. Card, S. K., T. P. Moran, and A. Newell, *Applied Information Processing Psychology: The Human-Computer Interface*, Lawrence Erlbaum Associates, Hillsdale, New Jersey (in preparation).

24. Card, S. K., T. P. Moran, and A. Newell, "Computer Text-Editing: An Information-Processing Analysis of a Routine Cognitive Skill," *Cognitive Psychol.*, **12**, 32-74 (1980).

25. Card, S. K., T. P. Moran, and A. Newell, *The Keystroke-Level Model of User Performance Time with Interactive Systems*, Rep. SSL-79-1, Xerox, Palo Alto Res. Ctr., Palo Alto, California (March 1979).

26. Carroll, J. M., and J. C. Thomas, "Metaphor and the Cognitive Representation of Computing Systems," *IEEE Trans. Syst., Man, and Cybern.*, SMC-12, 107-116 (1982).

27. Chapanis, A., "Men, Machines, and Models," *American Psychologist*, **16**, 113-31 (1961).

28. Chase, W. G., and H. A. Simon, "Perception in Chess," *Cognitive Psychol.*, **4**, 55-81 (1973).

29. Chow, E. Y., *Failure Detection System Design Methodology*, LIDS-TH-1055, MIT, Cambridge, Massachusetts (October 1980).

30. Coluccio, T., and J. Wasielewski, *Electro-Optical Image Enhancement*, RADC-TR-70-35, Griffin Air Force Base, Rome, New York (1970).

31. Corcoran, W. P., *Critical Functions, the Basic Structure for an Efficient Informations Systems,* graphic presentation circulated at the Workshop on Cognitive Modeling of Nuclear Plant Control and Operators, Dedham, Massachusetts (August 15-19, 1982).

32. Crooks, W. H., M. Artof, A. Freedy, and G. Weltman, *Man/Machine Interaction in Adaptive Computer Aided Control: Analysis of Automated Control Allocation*, Report PATR-74-12/30, Perceptronics, Inc., Woodland Hills, California (December 1974).

33. de Kleer, J., and J. S. Brown, "Assumptions and Ambiguities in Mechanistic Mental Models," in *Mental Models,* D. Gentner and A. S. Stevens (Eds.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1982).

34. de Kleer, J., and J. S. Brown, "Foundations of Envisioning," *Proceedings of the AAAI National Conference on Artificial Intelligence,* 1982.

35. de Kleer, J., and J. S. Brown, "Mental Models of Physical Mechanisms and Their Acquisition," in *Cognitive Skills and Their Acquisition,* J. A. Anderson (Ed.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1981).

36. Deutsch, P. L., and B. W. Lampson, "An Online Editor," *Comm. ACM 10*, **12**, 793-99 (1967).

37. Devoe, D. B., "Alternatives to Handprinting in the Manual Entry of Data," *IEEE Trans. HFE-8*, 21-32 (1967).

38. Dougherty, E. M., *Cognitive Demands on the Reactor Operator (as Inferred from Emergency Operating Instructions)*, TEC Report R-81-014, Technology for Energy Corporation, Knoxville, Tennessee (June 1981).

39. Drury, C. G., *Human Performance in Manual Process Control, Monitoring Behavior and Supervisory Control,* T. B. Sheridan and G. Johannsen (Eds.), Plenum Press, New York (1976).

40. Duncan, K. D., "A Simulator and Training Technique for Diagnosing Plant Failures from Control Panels," *Ergonomics* **18**, No. 6, 627-41 (1975).

41. Duncan, K. D., "Training for Fault Diagnosis in Industrial Process Plant," in *Human Detection and Diagnosis of System Failures*, Jens Rasmussen and William B. Rouse (Eds.), published in cooperation with NATO Scientific Affairs Division, Plenum Press, New York and London.

42. Dutton, J. M., and W. H. Starbuck, "Finding Charlie's Run-Time Estimator," in *Computer Simulation of Human Behavior*, J. M. Dutton and W. H. Starbuck (Eds.), Wiley, New York (1971).

43. Electric Power Research Institute, *Integrating Human Factors Engineering into Nuclear Power Plant Design*, NP-1637 (Draft), Palo Alto, California (June 1982).

44. Embley, D. W., M. T. Lan, D. W. Leinbaugh, and G. A. Nagy, "A Procedure for Predicting Program Editor Performance from the User's Point of View," *Internat. J. Man-Machine Studies*, **10**, 639-50 (1978).

45. English, H. B., and A. C. English, *A Comprehensive Dictionary of Psychological and Psychoanalytic Terms*, Longmans Green, New York (1958).

46. Erman, L., F. Hayes-Roth, V. Lesser and R. Reddy, "The Hearsay-II Speech Understanding System: Integrating Knowledge to Reduce Uncertainty," *Computing Surveys,* 213-252 (1980).

47. Feigl, H., *Readings in Philosophical Analysis*, Appleton-Century-Crofts, New York (1949).

48. Fitter, M. J., and M. E. Sime, "Responsibility and Shared Decision-Making," in *Human Interaction with Computers,* H. T. Smith and T. R. G. Green (Eds.), Academic Press, London (1980).

49. Fitts, P. M., and M. I. Posner, *Human Performance*, Brooks/Cole Publishing Company (1962).

50. Fogel, L. W., *Human Information Processing*, Prentice-Hall, Englewood Cliffs, New Jersey (1967).

51. Forbus, K. D., *Qualitative Process Theory,* Massachusetts Institute of Technology A.I. Memo No. 664 (1982).

52. Forbus, K. D., and A. Stevens, *Using Qualitative Simulation to Generate Explanations,* Technical Note, Navy Personnel Research and Development Center (1981).

53. Forgy, C., and J. McDermott, "OPS, a Domain-Independent Production System Language," pp. 933-939 in *Proceedings of the Fifth International Joint Conference in Artificial Intelligence* (1977).

54. Forsyth, R., U. Saisangjam, and J. Gilmer, "Some Empirical Results Related to the Robustness of the Rasch Model," *APM* **5**, 175-86 (1981).

55. Freedy, A., F. C. Hull, L. F. Lucaccini, and J. Lyman, "A Computer-Based Learning System for Remote Manipulator Control," *IEEE Trans. Syst., Man, and Cybern.*, SMC-1, 356-363 (1971).

56. Gai, E. G., and R. E. Curry, "A Model of the Human Observer in Failure Detection Tasks," *IEEE Trans. Syst. Man and Cybern.*, SMC-6, 85-94 (1976).

57. Golder, T. V., *P-3 Pilot Errors: A Conceptual Approach*, Center for Advanced Research, Naval War College (June 1978).

58. Goodstein, L. P., *Computer-Based Operating Aids*, Design 82, Birmingham, UK (September 22-23, 1982).

59. Goodstein, L. P., and J. Rasmussen, "Man/Machine System Design Criteria in Computerized Control Rooms," abbreviated version in *Proc. of ASSOPO 80 - an IFIP/IFAC Symposium*, Trondheim, Norway, June 16-18, 1980.

60. Guilford, J. P., and R. Hoepfner, *The Analysis of Intelligence*, McGraw-Hill, New York (1971).

61. Hake, H. H., "Form Discrimination and the Invariance of Form," in *Pattern Recognition*, L. Uhr (Ed.), Wiley, New York (1966).

62. Hall, R. E., Joseph R. Fragola, and William J. Luckas, Jr., *The Man-Machine Interface and Human Reliability: An Assessment and Projection,* Conference Record for 1981 IEEE Standards Workshop on Human Factors and Nuclear Safety, August 30 — September 4, 1981, Myrtle Beach, South Carolina.

63. Hayes-Roth, F., *Artificial Intelligence for Systems Management,* P-6573, The Rand Corporation, Santa Monica, California (1980).

64. Hayes-Roth, F., *Human Planning Processes,* R-2607-ONR, The Rand Corporation, Santa Monica, California (1980).

65. Hayes-Roth, F., D. Gorlin, S. Rosenschein, H. Sowizral and D. Waterman, *Rationale and Motivation for ROSIE,* N-1648-ARPA, The Rand Corporation, Santa Monica, California (1981).

66. Hayes-Roth, F., D. Waterman and D. Lenat (Eds.), *Building Expert Systems,* Addison-Wesley, Boston, Massachusetts (1983).

67. Hochberg, J., *Perception* (2nd ed.), Prentice-Hall, Englewood, New Jersey (1978).

68. Hochberg, J., and V. Brooks, "Film Cutting and Visual Momentum," in *Eye Movements and the Higher Psychological Functions,* J. W. Senders, D. F. Fisher and R. A. Monty (Eds.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1978).

69. Hochberg, J., and V. Brooks, "Psychological Aspects of the Cinema," in *Handbook of Perception, Vol. X,* E. C. Carterette and M. P. Friedman (Eds.), Academic Press, New York (1978).

70. Hochberg, J., and L. Gellman, "The Effect of Landmark Features on Mental Rotation Times," *Memory and Cognition* **5**, 23-26 (1977).

71. Hollnagel, E., O. M. Pedersen, and J. Rasmussen, *Notes on Human Performance Analysis*, Riso-M-2285, Riso National Laboraory, Denmark (March 1981).

72. Hollnagel, E., and J. Rasmussen, *Simulator Training Analysis: A Proposal for Combined Trainee Debriefing and Performance Data Collection in Training Simulators*, Riso-M-2301, Riso National Laboratory, Denmark (July 1981).

73. Hollnagel, E., and D. Woods, "Cognitive Systems Engineering: New Wine in New Bottles," *International Journal of Man-Machine Studies,* in press.

74. A. D. Swain and H. E. Guttmann, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications,* NUREG/CR-1278, draft report for interim use and commment, October 1980, U.S. Nuclear Regulatory Commission, Washington, D.C.

75. Jenkins, J. P., *Cognitive Models for Complex Man-Machine Systems Application* (draft) (1982).

76. Johannsen, G., and T. Govindaraj, "Optimal Control Model Predictions of System Performance and Attention Allocation and Their Experimental Validation in a Display Design Study," *IEEE Trans. Syst., Man, and Cybern.*, SMC-10(5), 249-261 (1980).

77. Joos, D. W., Z. A. Sabri, and A. A. Husseiny, "Analysis of Gross Error Rates in Operation of Commercial Nuclear Power Stations," *Nuclear Engineering and Design* **52**, 265-300 (1979).

78. Kahneman, D., *Attention and Effort*, Prentice-Hall, Englewood Cliffs, New Jersey (1973).

79. Kahneman, D., P. Slovic, and A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York (1982).

80. Katter, R. V., C. A. Montgomery, and J. R. Thompson, *Cognitive Processes in Intelligence Analysis: A Descriptive Model and Review of the Literature*, ARI Technical Report 445, U. S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, Virginia (February 1980).

81. Katter, R. V., C. A. Montgomery, and J. R. Thompson, *Human Processes in Intelligence Analysis: Phase I. Overview*, Report 1237, U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, Virginia (December 1979).

82. Kendall, M. G., "Introduction to Model Building and Its Problems," in *Mathematical Model Building in Economics and Industry*, Hofner, New York (1968).

83. Kisner, R. A., P. R. Frey, *Functions and Operations of Nuclear Power Plant Crews,* NUREG-CR-2587 (April, 1982).

84. Kisner, R. A., A. M. Fullerton, and P. R. Frey, "A Taxonomy of the Nuclear Plant Operator's Role," *Enlarged Holden Progress Meeting*, Friederichstaadt, Norway (June 14-19, 1981).

85. Klahr, P., and W. Faught, "Knowledge-Based Simulation," *Proceedings of the First Annual Conference of the American Association of Artificial Intelligence* (1980).

86. Kleinman, D. D., S. Baron, and W. N. Levison, "A Control Theoretic Approach to Manual Vehicle Systems Analysis," *IEEE Transactions on Automatic Control*, AC-16, No. 6, 824-832 (1971).

87. Kleinman, D. L., and R. E. Curry, "Some New Control Theoretic Models for Human Operator Display Monitoring," *IEEE Trans. Syst., Man, and Cybern.*, SMC-7, No. 11 (November 1977).

88. Kubovy, M., and J. R. Pomerantz, *Perceptual Organization,* Lawrence Erlbaum Associates, Hillsdale, New Jersey (1981).

89. Levine, J. M., and D. Eldredge, *Effects of Ancillary Information Upon Photointerpreter Performance,* Army Research Institute for the Behavioral and Social Science, Arlington, Virginia (1974).

90. Levison, W. H., and R. B. Tanner, *A Control-Theory Model for Human Decision Making*, NASA CR-1953 (December 1971).

91. Lind, M., "Generic Control Tasks in Process Plant Operation," *Proceedings of the European Annual Conference on Decision Making and Manual Control, Bonn, Germany, 1982.*

92. Lind, M., *Multilevel Flow Modeling of Process Plants for Diagnosis and Control* (to be published in 1982).

93. Loftus, G. R., and N. H. Mackworth, "Cognitive Determinants of Fixation Location During Picture Viewing," *Journal of Experimental Psychology: Human Perception and Performance* **4**, 565-572 (1978).

94. Lord, F. M., and M. R. Novick, *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, Massachusetts (1968).

95. Mackworth, N. H., and A. J. Morandi, "The Gaze Selects Information Details within Pictures," *Perception and Psychophysics* **2**, 547-552 (1967).

96. Marr, D., "Early Processing of Visual Information," *Philosophical Transactions of the Royal Society of London* **275**, 483-519 (1976).

97. Martin, F. F., *Computer Modeling and Simulation*, Wiley, New York (1968).

98. Maynard, H. B., *Industrial Engineering Handbook*, 3rd Ed., McGraw-Hill, New York (1971).

99. McCormick, E. J., *Human Factors Engineering*, 2nd Ed., McGraw-Hill, New York (1964).

100. McLeod, S., *Photointerpreter Performance Studies*, RADC-TDR-64-326, Griffin Air Force Base, Rome, New York (1964).

101. Meister, D., "Letter to the Editor," *Human Factors Society Bulletin*, Santa Monica, California (June 1982).

102. Miller, G. A., "The Magical Number Seven Plus or Minus Two; Some Limits on Our Capacity for Processing Information," *Psychology Review* **63**, 81-97 (1954).

103. Miller, G. A., "Psychology and Information," *American Documentation* **19**, 286-289 (1968).

104. Miller, G. M., G. A. Heise, and W. Lichten, "The Intelligibility of Speech as a Function of the Context of Test Materials," *Journal of Experimental Psychology* **41**, 329-45 (1951).

105. Minsky, M., "A Framework for Representing Knowledge," *The Psychology of Computer Vision*, P. Winston (Ed.), McGraw-Hill, New York (1975).

106. Moran, T. P., "The Command Language Grammar: A Representation for the User Interface of Interactive Computer Systems," *Internat. J. Man-Machine Studies* (in press).

107. Moray, N., *Human Information Processing and Supervisory Control,* Man-Machine Systems Laboratory Report, M.I.T. (1980).

108. Moray, N., "The Role of Attention in the Detection of Errors and the Diagnosis of Failures in Man-Machine Systems," in *Human Detection and Diagnosis of System Failure,* J. Rasmussen and W. B. Rouse (Eds.), Plenum Press, New York (1981).

109. Muralidharan, R., and S. Baron, "DEMON: A Human Operator Model for Decision Making, Monitoring and Control," *J. Cybern. and Infor. Sci.* **3** (1980).

110. Myer, T. H., and J. R. Barnaby, "Tenex Executive Language Manual for Users," Bolt, Beranek, and Newman, Inc., Cambridge, Massachusetts (1973).

111. Navon, D., and D. Gopher, "On the Economy of the Human Processing System: A Model of Multiple Capacity," *Psychological Review* **86**, 214-55 (1979).

112. Naylor, T. H., J. L. Balintfy, D. S. Burdick, and K. Chu, *Computer Techniques*, Wiley, New York (1966).

113. Neisser, U., *Cognition and Reality,* W. H. Freeman, San Francisco, California (1976).

114. Newell, A., J. C. Shaw, and H. A. Simon, "Elements of a Theory of Human Problem Solving," *Psychological Review* **65**, 151-166 (1958).

115. Newell, A., and H. A. Simon, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, New Jersey (1972).

116. Newman, W., and R. Sproull, *Principles of Interactive Computer Graphics*, 2nd Ed., McGraw-Hill, New York (1979).

117. Nii, P., and N. Aiello, "AGE (Attempt to Generalize): A Knowledge-Based Program for Building Knowledge-Based Programs," pp. 645-655 in *Proceedings of the Sixth International Joint Conference in Artificial Intelligence* (1979).

118. Nisbett, R., and L. Ross, *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice-Hall, Englewood Cliffs, New Jersey (1980).

119. Norman, D. A., "A Psychologist Views Human Processing: Human Errors and Other Phenomena Suggest Processing Mechanisms," Invited Address, *Proc. of the International Joint Conference on Artificial Intelligence*, Vancouver (1981).

120. Norman, D. A., "Categorization of Action Slips," *Psychological Review* **88**, 1-15 (1981).

121. Norman, D. A., "Steps Toward a Cognitive Engineering: Design Rules Based on Analyses of Human Error," *Proc. of the Conference on Human Factors in Computer Systems*, Institute for Computer Sciences and Technology, National Bureau of Standards, U.S. Department of Commerce, Gaithersburg, Maryland, March 15-17, 1982.

122. Norman, D. A., *Steps Toward a Cognitive Engineering: Systems Images, System Friendliness, Mental Models*; paper presented at Symposium on Models of Human Performance, ONR Contractor's Meeting, La Jolla, California (UCSD), June 19, 1981.

123. Norman, D. A., "Twelve Issues for Cognitive Science," in *Perspectives on Cognitive Science*, published jointly by Erlbaum and Ablex, Norwood, NJ (1981); Chinese and Japanese translations in progress.

124. Norman, D. A., and D. G. Bobrow, "Descriptions: An Intermediate State in Memory Retrieval," *Cognitive Psychology* **11**, 107-23 (1979).

125. Norman, D. A., and D. G. Bobrow, "On Data-Limited and Resource-Limited Processes," *Cognitive Psychology* 7, 44-64 (1975).

126. Norman, D. A., and D. G. Bobrow, "On the Role of Active Memory Processes in Perception and Cognition," in *The Structure of Human Memory,* C. N. Cofer (Ed.), W. H. Freeman, San Francisco, California (1976).

127. Norman, D. A., and T. Shallice, *Attention to Action: Willed and Automatic Control of Behavior,* (Tech. Rep.), Center for Human Information Processing, University of California, San Diego, La Jolla, California (1980).

128. Nuclear Regulatory Commission, *Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants,* NRC-Report WASH-1400 (NUREG-75/014) (October 1975).

129. Oh, C. B., E. M. Dougherty, and J. L. Hamrick, *Analysis of the Operator's Role During the Onset of a Emergency,* Report R-81-004, Technology for Energy Corporation, Knoxville, Tennessee (presented to Oak Ridge National Laboratory, February 1981).

130. Pew, R. E., "Human Information Processing Concepts for System Engineers," *System Engineering Handbook,* R. E. Machol (Ed.), McGraw-Hill, New York (1965).

131. Pew, R. W., S. Baron, C. E. Feehrer, and D. C. Miller, *Critical Review and Analysis of Performance Models Applicable to Man-Machine Systems Evaluation,* BBN Report No. 3446, Bolt Beranek and Newman, Inc., Cambridge, Massachusetts (March 1977).

132. Pew, R. W., D. C. Miller, and C. E. Feehrer, *Evaluation of Proposed Control Room Improvements Through Analysis of Critical Operators Decisions,* NP-1982, Research Project 891, Electric Power Research Institute, Palo Alto, California (August 1981).

133. Pope, R. H., "Power Station Control Room and Desk Design: Alarm System and Experience in the Use of CRT Displays," *International Symposium on Nuclear Power Plant Control and Instrumentation,* Cannes, France (1978).

134. Posner, M. I., "Orienting of Attention," *Quarterly Journal of Experimental Psychology* **32**, 3-25 (1980).

135. Potter, J. E., and M. C. Suman, *Thresholdless Redundancy Management with Arrays of Skewed Instruments,* AGARDograph No. 224 (1977).

136. Pulliam, R., *Models and Tasks in Intelligence Production: A Method of Analysis and a Taxonomy of Work,* (A BioTechnology Occasional Paper), BioTechnology, Inc., Falls Church, Virginia (September 1981).

137. Ramsey, H. R., and M. E. Atwood, "Human Factors in Computer Systems: A Review of the Literature," Science Applications, Inc., Englewood, Colorado (1979).

138. Rasmussen, J., *The Human Data Processor as a System Component: Bits and Pieces of a Model,* Report M-1722, Riso National Laboratory, Denmark (1974).

139. Rasmussen, J., "Models of Mental Strategies in Process Plant Diagnosis," Preprints of *Conf. on Human Detection and Diagnosis of System Failures,* Roskilde, Denmark (August 1980).

140. Rasmussen, J., "The Human as a System Component," in *Human Interactions with Computers,* T. H. Smith and T. R. G. Green (Eds.), Academic Press, London (1980).

141. Rasmussen, J., *Notes on Diagnostic Strategies in Process Plant Environment,* Report M-1983, Riso National Laboratory, Denmark (1978).

142. Rasmussen, J., *On the Structure of Knowledge — A Morphology of Mental Models in a Man-Machine System Context,* Report M-2192, Riso National Laboratory, Denmark (1979).

143. Rasmussen, J., "Outlines of a Hybrid Model of the Process Plant Operator," *Monitoring Behavior and Supervisory Control,* T. B. Sheridan and G. Johannsen (Eds.), Plenum Press, New York (1976).

144. Rasmussen, J., *The Role of Cognitive Models of Operators in the Design, Operation and Licensing of Nuclear Power Plants,* Paper No. 1, these proceedings.

145. Rasmussen, J., "Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models," (to be published in *IEEE Trans. Syst. Man, and Cybern.*).

146. Rasmussen, J., and M. Lind, *A Model of Human Decision Making in Complex Systems and Its Use for Design of System Control Strategies,* Riso-M-2349, Riso National Laboratory, Denmark (1982); paper presented at the American Control Conference, Arlington, Virginia, June 14-16, 1982.

147. Rasmussen, J., and M. Lind, *Coping with Complexity,* Riso-M-2293, Riso National Laboratory, Denmark (1981); paper presented at the European Annual Conference on Human Decision and Manual Control, Delft, 1981.

148. Rasmussen, J., and O. M. Pedersen, *Formalized Search Strategies for Human Risk Contributions: A Framework for Further Development*, Riso-M-2351, Riso National Laboratory, Denmark.

149. Rasmussen, J., and W. B. Rouse, *Human Detection and Diagnosis of System Failures*, Plenum Press, New York (1981).

150. Reed, S. K., *Psychological Processes in Pattern Recognition*, Academic Press, New York (1973).

151. Reiger, C., and M. Grinberg, "The Declarative Representation and Procedural Simulation of Causality in Physical Mechanisms," *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, 250-255 (1977).

152. Robertson, G., D. McCracken and A. Newell, "The ZOG Approach to Man-Machine Communication," *International Journal of Man-Machine Studies* **14**, 461-488 (1981).

153. Roitsch, P. A., G. L. Babcock, and W. W. Edmunds, *Human Factors Report of the Tenerife Accident. Aircraft Accident Report: Pan American World Airways, Boeing 747, N737 AP; KLM Royal Dutch Airlines, Boeing 747, PH-BUF; Tenerife, Canary Islands, March 27, 1977*, Air Line Pilots Association, Washington (1979) (actual report not dated).

154. Rouse, W. B., p. 214 in *The Man-Machine Interface and Human Reliability; An Assessment and Projection*, 1981 IEEE Standards Workshop and Human Factors in Nuclear Safety, August 30 — September 4, 1981, Myrtle Beach, South Carolina.

155. Rouse, W. B., "Models of Human Problem Solving: Detection, Diagnosis, and Compensation for System Failures," *Automatica*, in press.

156. Rouse, W. B., *Outline of a Model of Human Problem Solving in Failure Situations*, Paper No. 8, these proceedings.

157. Rouse, W. B., *Systems Engineering Models of Human Machine Interactions*, North Holland, New York (1980).

158. Rouse, W. B., and R. M. Hunt, "A Fuzzy Rule-Based Model of Human Problem Solving in Fault Diagnosis Tasks," *Proc. of the Eighth Triennial World Congress of the International Federation of Automatic Control*, Kyoto, Japan, August 1981.

159. Schank, R. C., and R. Abelson, *Scripts, Plans, Goals, and Understanding*, Lawrence Erlbaum Associates, Hillsdale, New Jersey (1977).

160. Savitsky, S., *Son of STOPGAP*, Rep. SAILON 50.1, Stanford Artif. Intell. Lab., Stanford, California (1969).

161. Schmall, T. M., (Editor), *The Human: The Key Factor in Nuclear Safety*, Conference Record for 1979 IEEE Standards Workshop on Human Factors and Nuclear Safety, December 2-7, 1979, Myrtle Beach, South Carolina.

162. Schneider, W., and R. Shiffrin, "Controlled and Automated Human Information Processing: I. Detection, Search, and Attention," *Psychological Review* **84**, 1-66 (1977).

163. Senders, J., J. Elkid, M. Grignetti and R. Smallwood, *An Investigation of the Visual Sampling Behavior of Human Observers*, NASA CR-434 (1966).

164. Sheridan, T. B., "On How Often the Supervisor Should Sample," *IEEE Trans.*, SSC-6, 140-145 (1970).

165. Sheridan, T. B., "Understanding Human Error and Aiding Human Diagnostic Behavior in Nuclear Power Plants," in *Human Detection and Diagnosis of System Failures*, Plenum Press, New York (1981).

166. Sheridan, T. B., *Supervisory Control: Problems, Theory and Experiment for Application to Human-Computer Interaction in Undersea Remote Systems*, MIT Man-Machine Systems Laboratory Report (1982).

167. Shiffrin, R. M., and W. Schneider, "Controlled and Automatic Human Information Processing: II. Perceptual Learning, Automatic Attending, and a General Theory," *Psychology Review* **84**, 127-90 (1977).

168. Siegel, A. I., W. D. Bartter, and F. F. Kopstein, *Job Analysis for Maintenance Mechanics in Nuclear Power Plants*, Applied Psychological Services, Wayne, Pennsylvania (1981).

169. Siegel, A. I., J. J. Wolf, and A. R. Williams, *Computer Simulation of Human Performance in Electronic Processed Imagery Systems*, Applied Psychological Services, Wayne, Pennsylvania (1978).

170. Siegel, A. I., and J. J. Wolf, *Man-Machine Simulation Models: Psychosocial and Performance Interaction*, John Wiley and Sons, Inc., New York (1969).

171. Simon, H. A., "How Big Is a Chunk," *Science*, **183**, 482-88 (1974).

172. Simon, H. A., and A. Newell, "Human Problem Solving: The State of the Theory in 1970," *American Psychologist* **26**, 145-59 (1971).

173. Sloman, A., *The Computer Revolution in Philosophy: Philosophy, Science, and Models of Mind*, The Harwester Press, Sussex (1978).

174. Starbuck, W. H., "Testing Case-Descriptive Models," in *Computer Simulation of Human Behavior*, J. M. Dutton and W. H. Starbuck (Eds.), Wiley, New York (1971).

175. Steeb, R., G. Weltman, and A. Freedy, *Man/Machine Interaction in Adaptive Computer-Aided Control: Human Factors Guidelines*, NR 196-118, Office of Naval Research, Arlington, Virginia (1976).

176. Stefik, M., J. Aikens, R. Balzer, J. Benoit, L. Birnbaum, F. Hayes-Roth and E. Sacerdoti, "The Organization of Expert Systems," *Artificial Intelligence* **18**, 135-173, 1982.

177. Swain, A. D., and H. E. Guttman, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*, Draft Report NUREG/CR-1278, Office of Nuclear Regulatory Research, US NRC, Washington, D. C. (April 1980).

178. Taft, R., "Summation," in *Cognition: Theory, Research, Promise*, C. Scheerer (Ed.), Harper Row, New York (1962).

179. Thorndyke, P., *Intelligent Opponent Simulation for Tactical Decisionmaking Training*, P-82-127, Perceptronics, Woodland Hills, California (1982).

180. Thorndyke, P., D. McArthur and S. Cammarata, "AUTOPILOT: A Distributed Planner for Air Fleet Control," pp. 171-177 in *Proceedings of the Seventh International Joint Conference in Artificial Intelligence* (1981).

181. Tsach, U., T. B. Sheridan, and J. Tzelgov, "A New Method for Failure Detection and Location in Complex Dynamic Systems," *Proc. 1982 American Control Conference*, Arlington, Virginia.

182. Tulga, M. K., and T. B. Sheridan, "Dynamic Decisions and Workload in Multi-Task Supervisory Control," *IEEE Trans. Syst., Man, and Cybern.*, Vol. SMC-10, No. 5 (1980).

183. Tversky, A., and D. Kahneman, "Casual Schemas in Judgments Under Uncertainty," in *Progress in Social Psychology*, M. Fishbein (Ed.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1980).

184. Tzelgov, J., U. Tsach, and T. B. Sheridan, "Effects of Indicating Failure Odds and Smoothed Outputs on Human Failure Detection in Dynamic Systems," *Human Factors* (submitted paper).

185. Umbers, I. G., "Models of the Process Operator," *Int. J. Man-Machine Studies* **11** 263-184 (1979).

186. Van Cott, H. P., and M. J. Warrick, "Man as a Systems Component," *Human Engineering Guide to Equipment Design*, Rev. Ed., H. P. Van Cott and R. G. Kinkade (Eds.), U.S. Government Printing Office, Washington, D.C. (1972).

187. Van Melle, W., *A Domain Independent System that Aids in Constructing Consultation Programs*, Report No. STAN-CS-80-820, Computer Science Department, Stanford University (1980).

188. Wainer, H., "Estimating Coefficients in Linear Models: It Don't Make No Nevermind," *Psychol. Bull.*, **83**, 213-17 (1976).

189. Wason, P., and P. N. Johnson-Laird, *Psychology of Reasoning: Structure and Content*, Harvard University Press, Cambridge, Massachusetts (1972).

190. Waterman, D., and F. Hayes-Roth (Eds.), *Pattern-Directed Inference Systems*, Academic Press, New York (1979).

191. Welford, A. T., "The Measurement of Sensory-Motor Performance: Survey and Appraisal of Twelve Years Progress," *Ergonomics* **3**, 189-230 (1960).

192. Wesson, R. B., "Planning in the World of the Air Traffic Controller," *Proc. of the Fifth International Joint Conference on Artificial Intelligence*, MIT (August 1977).

193. Wewerinke, P., "A Model of the Human Observer and Decision Maker," *Proc. of the Seventeenth Annual Conf. on Manual Control*, Pasadena, California (1981).

194. Wherry, R. J., "The Human Operator Simulator-HOS," in *Monitoring Behavior and Supervisory Control*, T. B. Sheridan and G. Johannsen (Eds.), Plenum Press, New York and London (1976).

195. Williams, M. D., and J. Hollan, "The Process of Retrieval from Very Long Term Memory," *Cognitive Psychology* **5**, 87-227 (1981).

196. Williams, M. D., J. Hollan and A. Stevens, "Human Reasoning About a Simple Physical System," in *Mental Models,* D. Gentner and A. S. Stevens (Eds.), Lawrence Erlbaum Associates, Hillsdale, New Jersey (1982).

197. Willsky, A. S., "A Survey of Design Methods for Failure Detection in Dynamic Systems," *Automatica* **12** (1976).

198. Wohl, J. G., *Mathematical Representation of Cognitive Processes in Human-Machine Interaction* (in press).

199. Wohl, F. G., Panel Discussion on Human Performance Models and Theories, paper presented at the IEEE Standards Workshop on Human Factors and Nuclear Safety, Myrtle Beach, SC (August 30 — September 4, 1981).

200. Woodworth, R. S., and H. Schlosberg, *Experimental Psychology*, Holt, Rinehart, and Winston, New York (1954).

201. Woods, D. D., "Visual Momentum: A Concept to Improve the Cognitive Coupling of Human and Machine," manuscript submitted for publication (1982).

202. Woods, D. D., and E. Hollnagel, *Cognitive Task Analysis* (in press).

203. Wortman, D. B., S. D. Duket, D. J. Seifert, R. L. Hann, and G. P. Chubb, *Simulation Using SAINT: A User Oriented Instruction Manual*, AMRL-TR-77-61, Wright Patterson Air Force Base, Ohio (July 1978).

204. Zacharias, G. L., S. Baron, and R. Muralidharan, *A Supervisory Control Model of the AAA Crew*, BBN Report No. 4802, Bolt Beranek and Newman, Inc., Cambridge, Massachusetts (October 1981).

205. Zusne, L., *Visual Perception of Form*, Academic Press, New York (1970).

INTERNAL DISTRIBUTION

1–2. L. S. Abbott
3. J. L. Anderson
4. F. H. Clark
5. G. F. Flanagan
6. P. R. Frey
7. A. M. Fullerton
8. H. Gray
9. P. M. Haas
10. E. W. Hagen
11–20. R. A. Kisner
21. H. E. Knee
22. A. L. Lotts
23. F. C. Maienschein
24. F. R. Mynatt
25. L. C. Oakes
26. P. J. Otadvy
27. T. W. Reddoch
28. W. H. Sides, Jr.
29. D. B. Trauger
30–31. Central Research Library
32–34. Document Reference Section
35–36. Laboratory Records Department
37. Laboratory Records, ORNL RC
38. ORNL Patent Section

EXTERNAL DISTRIBUTION

39. S. Baron, Vice President, Information Sciences Division, Bolt, Beranek, and Newman, Inc., 10 Moulton St., Cambridge, MA 02238
40. Barbara Jean Bell, Staff Member, Sandia National Laboratories, Division 7223, Albuquerque, NM 87185
41. Leo Beltracchi, Senior Human Factors Engineer, Human Factors Engineering Branch, Division of Human Factors Safety, Office of NRR, U.S. Nuclear Regulatory Commission, Washington, DC 20555
42. Julian M. Christiansen, Chief Scientist, Human Factors, General Physics Corporation, Suite 240, 1010 Woodman Drive, Dayton, OH 45432
43. William R. Corcoran, Director, Plant Engineering, Nuclear Power Systems, Combustion Engineering, Inc., Mail Stop 9482421, 1000 Prospect Hill Road, Windsor CT 06095
44. Renwick Curry, NASA Ames, 239-3 Moffet Field, Mountain View, CA 94035
45. Paul E. Dietz, Project Manager, Criteria & Analysis Division, Institute of Nuclear Power Operations, 1820 Waterplace, Atlanta, GA 30339
46. David E. Embry, 1 School House, Higher Lane, Dalton, Parbold, Lancashire WN8 7RP, England
47. Donald Farr, Director, Human Factors Systems, Science Applications, Inc., 1710 Goodridge Drive, McLean, VA 22102
48. Carl E. Feehrer, Bolt, Beranek & Newman, Inc., 10 Moulton Street, Cambridge, MA 02238
49. Joseph R. Fragola, Science Applications, Inc., 274 Madison Avenue, Suite 1501, New York, NY 10016
50. John M. Gallagher, Jr., Consulting Engineer, Control & Electric Systems, Westinghouse Electric Corp., P.O. Box 355, MNC-456, Pittsburgh, PA 15230

51. Henry M. Halff, Asst. Leader, Personnel & Training Research Group, Office of Naval Research, Code 442 PT, 800 N. Quincy Street, Arlington, VA 22217
52. Lewis F. Hanes, Manager, Human Sciences Group, Building 801-3, C-8, Westinghouse R&D Center, 1310 Beulah Road, Pittsburgh, PA 15235
53. Theodore R. Hatcher, Idaho National Engineering Laboratory, EG&G Idaho, Inc., Idaho Falls, ID 83415
54. C. L. Hogan, Fairchild Camera and Instrument Corporation, 464 Ellis Street, Mountain View, CA 94040
55. Eric Hollnagel, OECD Halden Reactor Project, P.O. Box 173-N, 1751 Halden, Norway

56–70. J. P. Jenkins, Senior Engineering Psychologist, U.S. Nuclear Regulatory Commission, RES/DFO/HFB, Mail Stop 1130SS, Washington, D.C. 20555

71. Michael G. Johnson, Associate Professor, Austin Peay, Room 301-C, Department of Psychology, University of Tennessee, Knoxville, TN 37916
72. J. V. Leeds, Jr., Professional Engineer, 10807 Atwell, Houston, TX 77096
73. William J. Luckas, Jr., Group Leader, Engineering Analysis & Human Factors, Building 130, Department of Nuclear Energy, Brookhaven National Laboratory, Upton, NY 11973
74. C. A. McKay, Executive Vice President, Corporate Director of Technology, The Foxboro Company, 38 Neponset Avenue, Foxboro, MA 30205
75. O. R. Meyer, Idaho National Engineering Laboratory, Box 1625, Willow Creek Building, Idaho Falls, ID 83415
76. Thomas P. Moran, Research Scientist, Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304
77. Neville Moray, Professor, University of Toronto, Department of Industrial Engineering, Toronto, Ontario, Canada M5S 1A4
78. John Morell, JFK CMHC, Research and Evaluation Dept., 112 N. Broadstreet, Philadelphia, PA 19102
79. Jack E. Mott, General Electric Company, Advanced Reactor Systems Dept., 310 DeGuigne Ave., Sunnyvale, CA 94086
80. Kenneth G. Murphy, Jr., Nuclear Engineer, Division of Risk Analysis, Office of RES, U.S. NRC, Washington, DC 20555
81. William R. Nelson, Engineering Specialist, Idaho National Engineering Lab., EG&G Idaho, Inc., P.O. Box 1625, Idaho Falls, ID 83415
82. Roger A. Newton, Superintendent, Reactor Engineering, Wisconsin Electric Power Company, 231 W. Michigan, P.O. Box 2046, Milwaukee, WI 53201
83. Donald A. Norman, Professor of Psychology and Director of the Institute for Cognitive Science, University of California, San Diego, La Jolla, CA 29093

84. John F. O'Brien, Project Manager, Nuclear Engineering & Operations Department, Electric Power Research Institute, 3412 Hillview Avenue, P.O. Box 10412, Palo Alto, CA 94303
85. B. M. Oliver, Vice President, Research and Development, Hewlett-Packard Company, 1501 Page Mill Road, Palo Alto, CA 94304
86. Charles M. Overbey, Human Engineering Section Leader, Division of Facility Operations, RES/DFO/HFB, U.S. NRC, Washington, DC 20555
87. J. R. Penland, Assistant Vice President, Science Applications, Inc., Suite 1000, Jackson Plaza Tower, 800 Oak Ridge Turnpike, Oak Ridge, TN 37830
88. Richard W. Pew, Principal Scientist, Bolt, Beranek & Newman, Inc., 10 Moulton Street, Cambridge, MA 02238
89. Robert Pulliam, Staff Scientist, BioTechnology, Inc., 3027 Rosemary Lane, Falls Church, VA 22042
90. Leonard C. Pugh, Principal Engineer, General Electric Company, Nuclear Control and Instrumentation Dept., 175 Curtner Avenue, San Jose, CA 95125
91. Jens Rasmussen, Head Electronics Dept./RISO, % Search Technology Inc., 25 B Technology Park/Atlanta, Norcross, GA 30092
92. William B. Rouse, President, Search Technology, Inc., 25 B Technology Park/Atlanta, Norcross, GA 30092
93. Thomas G. Ryan, Senior Engineering Psychologist, RES/DFO/HFB, U.S. Nuclear Regulatory Commission, Washington, DC 20555
94. Donald L. Schurman, Principal Scientist, Applied Science Associates, Inc., 1 East Cruikshank Road, Box 158, Valencia, PA 16059
95. John W. Senders, Professor, University of Toronto, Dept. of Industrial Engineering, Ontario, Canada M5S 1A4

96–105. Thomas B. Sheridan, Professor, Massachusetts Institute of Technology, Building 1, Room 110, MIT, Cambridge, MA 02139

106. Arthur I. Siegel, Director, Applied Psychological Services, 404 E. Lancaster Avenue, Wayne, PA 19087
107. Perry W. Thorndyke, Director, Knowledge Systems Branch, Perceptronics, 245 Park Lane, Atherton, CA 94025
108. Brian Tolley, Commission of the European Communities, Directorate of Science, Research and Development, Rue de la Loi Zoo, 1049, Brussels, Belgium
109. Jussi K. Vaurio, Program Manager, Argonne National Laboratory, Fast Reactor Safety Technology, Management Center, 9700 South Cass Ave., Argonne, IL 60439
110. Michael Williams, Scientist, Cognitive & Instructional Sciences Group, Xerox Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304

111. J. G. Wohl, Vice President R&D, Alpha Tech Inc., 3 New England Executive Park, Burlington, MA 01803
112. David D. Woods, Research Psychologist, Westinghouse Research & Development, Pittsburgh, PA 15235
113. J. Wreathall, Executive Engineer, NUS, 910 Clopper Road, Gaithersburg, MD 20878
114–115. DOE Technical Information Center, Oak Ridge, TN 37830
116. Office of Assistant Manager for Energy Research and Development, Department of Energy, Oak Ridge Operations Office, Oak Ridge, TN 37830
117–536. Given distribution as shown in NRC Category RX (10 copies NTIS)

☆ U.S. GOVERNMENT PRINTING OFFICE: 1983-646-055/273