ANALYTICAL EVALUATION OF COMPUTER-BASED DECISION AIDS

R. A. Kisner
Oak Ridge National Laboratory*
Oak Ridge, Tennessee

CONF-850903--11

DE85 017157

ABSTRACT

This paper introduces a method for evaluating decision aids for nuclear power plant operators. The method involves a two-stage process of classification and analytical evaluation of display form and content. The classification scheme relates each specific aid to one or more general decision-making tasks. Evaluation then proceeds using a normative top-down design process based on the classification scheme by determining or deducing how various design issues associated with this process were resolved by the designer. The result is an assessment of the "understandability" of the aid as well as identification of the training and display features necessary to ensure understandability.

INTRODUCTION

A means is needed to objectively evaluate the usefulness of proposed computer-based decision aid systems** prior to and as an adjunct to empirical simulation or in situ testing. The method introduced here is a two-stage process of classification and evaluation. Analytical assessment and empirical evaluation should be used together for a more accurate judgment of the usefulness of a proposed decision aid. Ideally, evaluation begins with analytical assessment of the aid and culminates in empirical validation that the aid achieves the design objectives pursued.

The approach taken in this work is a purely paper evaluation by a knowledgeable analyst, based mainly on design documentation. This work is limited to developing analytical methods for determining minimum information requirements [1]; other recent efforts have produced a comprehensive methodology for empirical evaluation [2]. The completeness of design documentation greatly influences the amount of effort required to perform the analytical assessment. This factor should be considered during the early phases of decision aid development.

The first stage, classification, maps any particular decision aid to one or more general decision-making tasks. (The taxonomy of general decision-making tasks employed in the mapping is based on a conceptual model of human decision making.)

The second stage of the proposed method, evaluation, is based on a normative, top-down view of a system design. In general, an aid is evaluated by first assuming that it was produced using a normative design process such as proposed by Frey et al. [3], and then determining how the various design issues associated with this process

*Operated by Martin Marietta Energy Systems, Inc., for the U.S. Department of Energy under Contract No. DE-AC05-84OR21400.

**A decision aid system is defined as a computer-based analysis and display system that provides information to help operators make correct decisions in the course of normal or off-normal plant operation.

were resolved by the designer. More specifically, an aid is evaluated in terms of the situations and tasks for which it was intended, the forms of information appropriate for each situation, the prototypical messages required to support each task, and the knowledge necessary to understand the messages. The result of the evaluation is an assessment of "understandability" as well as the training and display features necessary to ensure understandability.

OPERATOR DECISION MAKING

A recent survey of the decision-making literature [4] concludes that virtually every aid reviewed is aimed at supporting one or more of three general, continuous decision-making tasks: (1) execution and monitoring, (2) situation assessment, and (3) planning and commitment. The relationship among these tasks is shown in Fig. 1.

The first decision-making task, execution and monitoring, involves implementing the current plan, observing its consequences, and evaluating any deviation of observed consequences from expectations. Most operator activities are dominated by execution and monitoring. The majority of the time, differences between observations and expectations are minor; consequently, situation assessment or planning and commitment are not required. However, when they are required (i.e., when the deviations are unacceptable), the role of the operator becomes central to ensuring continued system operation and safety.

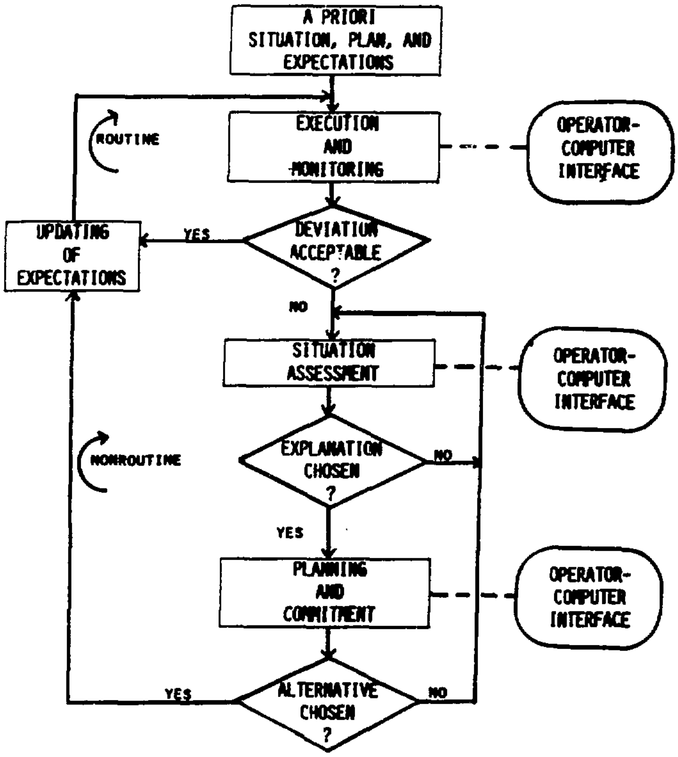

Figure 1. Relationships among decision-making tasks.

Situation assessment is required when the information received by an operator differs from his expectations in an unacceptable manner. Unexpected deviations prompt the operator to question the validity of a priori assumptions regarding the status quo, which in turn leads to a search for an explanation of what has happened, is happening, or may happen. As the term implies, the goal of situation assessment is to identify and assess the conditions that produced the unexpected information.

The third general decision-making task, planning and commitment, involves generating, evaluating, and selecting among alternative courses of action relative to criteria that reflect tradeoffs between possibly competing objectives (e.g., availability versus safety). When situation assessment reveals situations that were not anticipated in the design of procedures or were not considered in the design of training, operators may be required to pursue planning and commitment.

General decision-making tasks can be further subdivided into the subtasks listed in Table 1. Each decision aid is classified using this list and is then evaluated to verify that it supports one or more of the 13 subtasks.

Table 1. Decision-making tasks and subtasks

Task	Subtask
EXECUTION AND MONITORING	1. Implementation of plan 2. Observation of consequences 3. Evaluation of deviations from expectations 4. Selection of acceptance or rejection
SITUATION ASSESSMENT: INFORMATION SEEKING	5. Generation/identification of alternative information sources 6. Evaluation of alternative information sources 7. Selection among alternative information sources
SITUATION ASSESSMENT: EXPLANATION	8. Generation of alternative explanations 9. Evaluation of alternative explanations 10. Selection among alternative explanations
PLANNING AND COMMITMENT	11. Generation of alternative courses of action 12. Evaluation of alternative courses of action 13. Selection among alternative courses of action

ANALYTICAL APPROACH TO EVALUATION

Following classification by decision task (stage one), the evaluation proceeds to a determination of whether or not the aid provides the information necessary to perform the decision tasks. Based on design documentation and discussions with the designer, the designer's intentions can be classified in terms of support of the decision subtasks. The evaluative question then becomes whether or not the decision aid provides the necessary information for the operator to perform the decision subtasks asked of him. It is possible to consider the types and forms of required information, but it is not feasible within a general framework to specify the particular variables that must be presented.

Design Framework

The types and forms of the required information are then evaluated based on the types of situations that could be encountered (classification of familiarity and frequency of specific situations) and the types of strategies used by operators (symptomatic or topographic). Most situations are familiar and frequent--familiar in that the possibility of their occurrence has been anticipated, and frequent in that the operator has had considerable experience in dealing with them. Hence, the course of action is immediately apparent. In contrast, familiar but infrequent situations usually do not allow for immediate action because the operator does not have much experience with them even though their possibility for occurrence was anticipated. As a result, a person may rapidly hypothesize a course of action but collect a variety of information before pursuing it. Unanticipated situations that are unfamiliar and infrequent are experienced so seldom that available procedures may be inadequate or inappropriate.

Decision makers approach these three types of situations differently. Familiar situations call upon the human's pattern recognition abilities. Therefore, for these situations problem-solving strategies tend to be symptomatic in the sense that observed patterns are mapped directly to likely solutions. At the other extreme, unfamiliar situations call upon human analytical reasoning abilities. For these situations, problem-solving strategies tend to be topographic in the sense that system functions and the relationship among those functions are explicitly considered in the search strategy.

Types of information can be described in terms of two dichotomies: patterns versus elements and current versus projected. The distinction between aggregate patterns and discrete elements is important in determining how the system state should be displayed. For symptomatic strategies, system state should be displayed as an aggregate pattern such as an iconic display [5]. In contrast, topographic strategies require that the system state be displayed as discrete elements (mimic and block diagrams). This is because the values of variables such as temperatures, pressures, and valve positions are usually needed to trace through the topography of the system.

The appropriateness of the forms of information can be determined by comparison with the types of situations and types of strategies. Figure 2 illustrates that the choice of task and type of situation dictate the form of information and, hence, the choice of how the information is displayed (e.g., analog versus digital, trend plots, or mimic displays).

Form is only one attribute of information display. Of greater interest is display content, which is application specific. Knowing the task and situation, one can determine the nature of the message that must be transmitted to support each task and situation. The result will be identification of missing or irrelevant information components. The 13 general decision-making tasks are re-formed into the prototypical messages shown in Fig. 3. A set of prototypical messages is used as a framework with

DECISION-MAKING TASKS	TYPE OF SITUATION		
	FAMILIAR & FREQUENT	FAMILIAR & INFREQUENT	UNFAMILIAR & INFREQUENT
EXECUTION & MONITORING	CURRENT PATTERNS	CURRENT PATTERNS	CURRENT PATTERNS & ELEMENTS
SITUATION ASSESSMENT	—	CURRENT PATTERNS & PROJECTED ELEMENTS	CURRENT & PROJECTED ELEMENTS
PLANNING & COMMITMENT	—	—	CURRENT & PROJECTED ELEMENTS

Figure 2. Appropriate forms of information.

which to compare the actual messages or displays of the decision aid under evaluation. In order for an aid to support a particular task, the display must provide at least one of the prototypical messages associated with the task.

Evaluation

Three evaluation levels are possible: compatibility, understandability, and effectiveness. These levels form a hierarchy in which the level above cannot function without support from the level below. For example, understandability of a decision aid is not meaningful if the aid is not compatible with the sensorimotor abilities of human users; similarly, an aid cannot be effective if the data are not understandable. Issues of compatibility such as readability and reachability of displays and controls can be addressed by the guidelines of NUREG-0700 [6]. The evaluation ofeffectiveness, which goes beyond understandability, is an assessment of the degree to which an aid supports achievement of an overall plant performance objective. This assessment will require empirical testing. The evaluation level described here concentrates on the understandability level.

An aid is understandable to the extent that the information communicated to users is meaningful to them. To assess understandability one must first determine the knowledge that users must possess in order to understand the messages displayed. Once these knowledge requirements have been identified, one must then assess the extent to which users can be expected to have this knowledge. Any knowledge that is lacking can be designated as presenting a potential limit to understandability.

IMPLEMENTATION:

1. THE COMPLETE [STEPS / PROCEDURES / GOALS] ARE . . .

2. THE CURRENT [STEP / PROCEDURE / GOAL] IS . . .

3. THE NEXT [STEP / PROCEDURE / GOAL] IS . . .

OBSERVATION:

THE CURRENT [STATE] IS . . .

GENERATION/IDENTIFICATION:

THE POSSIBLE [INFORMATION SOURCES / EXPLANATIONS / COURSES OF ACTION] ARE [A* / B / C] BECAUSE. . .

EVALUATION:

1. DEVIATION OF [STATE] IS [WITHIN EXPECTATIONS / OUTSIDE OF EXPECTATIONS]

2. CONFIDENCE IN [INFORMATION SOURCE / EXPLANATION / COURSE OF ACTION] IS [D* / E / F] BECAUSE. . .

3. CONSEQUENCES OF [EXPLANATION / COURSE OF ACTION] WILL BE [G* / H] BECAUSE. . .

4. RESOURCES FOR [INFORMATION SOURCE / COURSE OF ACTION] WILL BE [I* / J] BECAUSE. . .

5. COMPARISON OF [INFORMATION SOURCES / EXPLANATIONS / COURSES OF ACTION] IN TERMS OF [K* / L / M] YIELDS RANK ORDERING OF...

SELECTION:

1. DEVIATION OF [STATE] IS [ACCEPTABLE / UNACCEPTABLE] BECAUSE. . .

2. THE BEST [INFORMATION SOURCE / EXPLANATION / COURSE OF ACTION] IS [A* / B / C] BECAUSE. . .

*A = DISPLAY ELEMENTS, DISPLAY PAGES, MANUALS, CREW MEMBERS
B = FEASIBLE SET OF FAILURES, POSSIBLE SITUATIONS, LIKELY CONTRIBUTING EVENTS/FACTORS
C = PROCEDURES, PLANS
D = ACCURACY, RELEVANCE
E = COMPLETENESS, APPROPRIATENESS
F = SUFFICIENCY, LIKELY SUCCESS
G = RESULTING STATE, RESULTING SITUATION, PROCEDURE IMPLIED
H = RESULTING STATE, RESULTING SITUATION
I = TIME, PERSONNEL
J = TIME, PERSONNEL, EQUIPMENT, INVENTORY
K = CONFIDENCE, RESOURCES
L = CONFIDENCE, CONSEQUENCES, URGENCY
M = CONFIDENCE, CONSEQUENCES, RESOURCES, URGENCY

Figure 3. Prototypical messages.

Knowledge requirements can be classified into three categories: (1) display (coding), (2) command (dialogue), and (3) plant (functions and locations). Using the finer grained classification provided in Fig. 4, one can consider each type of message as it is manifested on the display. Knowledge requirements in each category can then be identified, usually by or with the help of the designer or other individuals who are knowledgeable regarding the specific application for which the aid is intended.

Once the knowledge requirements have been identified, one must assess the extent to which users will possess this knowledge. One approach is to employ a database such as the Job and Task Analysis Database developed by the Institute of Nuclear Power Operations (INPO) [7], which is particularly useful for assessing whether or not typical operators will have particular elements of plant knowledge. For display and command knowledge, one may have to consider what conventions are employed in the environment where the aid will be used.

If one cannot ensure that particular knowledge elements have been provided by typical operators' experience and training, then look elsewhere. Two other sources are possible: operator training for using the aid and other displays intrinsic to the aid or elsewhere in the control room. Knowledge requirements not satisfied by any of the above sources are deemed unsatisfied and, hence, potential limits to understandability.

PLANT KNOWLEDGE

WHAT: CHARACTERISTICS (LOCATIONS, UNITS, CONTENTS, DEFINITIONS, DESIGN CHARACTERISTICS, INPUTS, OUTPUTS, SOURCES, LIMITS)
RELATIONSHIPS (SOURCES, INPUTS, OUTPUTS, INTERLOCKS, ORGANIZATION, DIFFERENCES)
PATTERNS (STATES, TRENDS, SEQUENCES, ALIGNMENTS)
SITUATIONS (STATES, MODES)
CRITERIA (PRIORTIES, LIMITS)
ANALOGIES (SIMILARITIES, DIFFERENCES)

HOW: FUNCTIONS (CAUSES, EFFECTS)
PROCEDURES(OPERATIONS)
STRATEGIES

WHY: REQUIREMENTS (PURPOSE, REASONS)
OBJECTIVES
OPERATIONAL BASES
LOGICAL BASES
PHYSICAL PRINCIPLES/THEORIES
MATHEMATICAL PRINCIPLES/THEORIES

DISPLAY KNOWLEDGE

TERMINOLOGY (LABELS, WORDS, ABBREVIATIONS)
SYMBOLOGY (SYMBOLS, CODING)
ELEMENTS (HOW TO READ OR INTERPRET ELEMENTS)
ORGANIZATION (RELATIONSHIPS AMONG DISPLAYS, CURRENT LOCATION)
DELAYS (DATA UPDATE, REDRAW TIME)

COMMAND KNOWLEDGE

TERMINOLOGY (COMMANDS, ARGUMENTS, ABBREVIATIONS)
SYMBOLOGY (SYMBOLS, CODING)
DEVICES (HOW TO USE DEVICES)
MODES (WHEN TO USE COMMANDS)
FEEDBACK (WHAT TO EXPECT)

Figure 4. Classification of knowledge requirements.

Assessment of understandability proceeds as follows: (1) the knowledge requirements for understanding each message as manifested on the display are identified; (2) the extent to which typical operators possess this knowledge from experience, training, or other displays is assessed; and (3) knowledge requirements not satisfied are deemed to reflect design inadequacies. The list of these inadequacies is the product of evaluation that is used to improve an aid.

CONCLUSION

An analytical approach to evaluation of computer-based decision aids should allow NRC regulators, plant designers, and others to determine the effects of an aid's proposed design on the operator's understanding of plant processes and how that understanding influences important operator decisions concerning plant safety.

REFERENCES

1. W. B. Rouse et al., "A Method for Analytical Evaluation of Computer-Based Decision Aids," Oak Ridge National Laboratory Report NUREG/CR-3655, ORNL/TM-9068, July 1984.

2. W. B. Rouse, "Computer-Generated Display System Guidelines: Vol. 2, Developing an Evaluation Plan," EPRI Report NP-3701, Vol. 2, September 1984.

3. P. R. Frey et al., "Computer-Generated Display System Guidelines: Vol. 1, Display Design," EPRI Report NP-3701, Vol. 1, September 1984.

4. W. B. Rouse and S. H. Rouse, "A Framework for Research on Adaptive Decision Aids," Search Technology, Inc., Report AFAMRL-TR-83-082, October 1983.

5. J. Rasmussen, "Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models," IEEE Trans. Systems, Man, Cybernet., SMC-13(3), 257-66, May/June 1983.

6. U.S. Nuclear Regulatory Commission, "Guidelines for Control Room Design Reviews," NUREG-0700, September 1981.

7. Institute of Nuclear Power Operations, "Job and Task Analysis Retrieval Manual," INPO Report, September 1983.