

NUCLEAR POWER PLANT PERSONNEL, ENTRY LEVEL QUALIFICATIONS AND TRAINING*

C. C. Jorgensen, P. M. Haas, F. L. Selby
 Engineering Physics Division
 Oak Ridge National Laboratory

J. C. Lowry
 Human Factors Branch
 Division of Facility Operations
 U. S. Nuclear Regulatory Commission

CONF-8310143--55

DE84 003294

INTRODUCTION

personnel — Along with plant design and operation, the areas of nuclear power plant qualification, education, and training are undergoing intensive study in the "post-TMI" era. The U. S. Nuclear Regulatory Commission (NRC) Offices of Nuclear Regulatory Research (RES) and Nuclear Reactor Regulation (NRR), along with the nuclear power industry (e.g., INPO) have developed programs to assess and improve nuclear power plant (NPP) operator training. Although there are many proponents of different approaches, there is now virtually unanimous agreement that improvement of personnel performance must include improvement of the training process. There also appears to be a consensus that nuclear industry training would benefit from adaptation of a "systematic approach" to the design, implementation, and evaluation of training programs.

This paper summarizes the early results and current status of a research program at Oak Ridge National Laboratory (ORNL) which is intended to provide the methods and technical basis for NRC to initiate the use of the Systems Approach to Training (SAT) in the evaluation of training programs and entry level qualifications for NPP control room personnel. The program (which has the same title as this paper) is to some extent an outgrowth of previous studies of simulator hardware and simulator training requirements under the Safety Related Operator Actions Program^{1,2,3} which recommended adaptation of a systems methodology to development and evaluation of NPP training programs. The first year's effort, initiated

*Research sponsored by the U. S. Nuclear Regulatory Commission under DOE Interagency Agreement 40-550-75 with Union Carbide Corporation under Contract No. W-7405-eng-26 with the U. S. Department of Energy.

By acceptance of this article, the publisher or recipient acknowledges the U.S. Government's right to retain a nonexclusive, royalty-free license in and to any copyright covering the article.

MASTER

in March, 1982 focused on adaptation of the SAT process to NRC needs. The current effort (initiated in August, 1983) will extend that work to adapt, develop, and field test a complete set of methods and tools for NRC to evaluate training and entry level qualification programs.

EARLY RESULTS (FY 82-83)

As noted above, the primary goal of the initial work was to examine and adapt the SAT process to provide a framework for NRC to evaluate industry programs for operator qualifications and training. Although there was an extensive body of literature, primarily from military and aerospace applications, on the design and development of training programs using the Instructional Systems Development (ISD) or SAT processes, there was considerably less available on the evaluation of the effectiveness of these programs, particularly as applied to a civilian, regulated industry. The existing SAT literature was re-examined, and with the input of industry and NRC representatives provided during two project working meetings, a high-level "model" of the SAT for use in the nuclear industry was outlined to serve as the desired framework for NRC evaluation. That model is depicted in Fig. 1. It is compatible with the ISD-based model developed by INPO⁴, with the additional emphasis that training requirements are derived from personnel performance requirements, which in turn are derived from system requirements identified in a front-end systems analysis.

Following this high-level model, twelve checklists were proposed with rating scales for NRC to use to evaluate each of the major elements of the process. The checklists developed are listed in Table 1. It is important to note that these proposed checklists have not been field tested or evaluated by end users for practicality, acceptability or utility. Such evaluation is necessary prior to full implementation, and would likely result in modification of content and/or approach.

The results of this initial effort are documented in NUREG/CR-3414.⁵ Included in that report are two other results produced during the course of development of the model and checklists - (1) preliminary identification of key variables (performance shaping factors) to be considered in NPP control room personnel qualifications and training, and (2) an illustrative media selection model that might be used in the development of an instructional delivery system for NPP control room operators.

A separate but related task carried out during this initial phase of the program was the development of a technique that can be used to rank possible plant malfunctions as to their importance to training, in particular, to simulator training. The ranking scheme is compatible with a fully implemented SAT-based training system, but is designed using subjective scales so that it can be used in the interim before

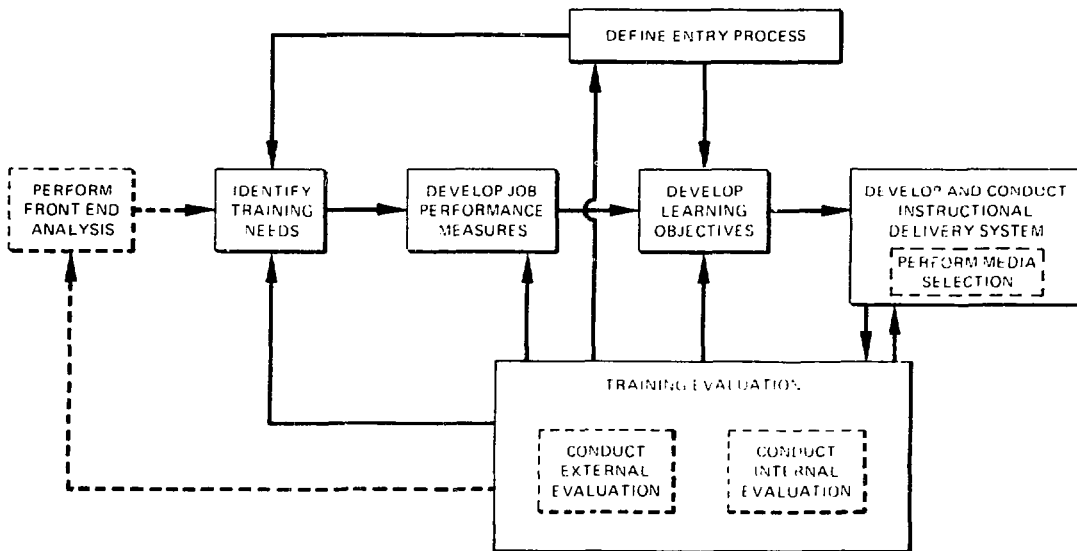


Figure 1. SAT structure for use in evaluating NPP training systems.

TABLE 1. Proposed Checklists for Evaluation of SAT-Based Training System

Job and Task Analysis Checklist

Job Performance Measure Checklist

Training Objectives Checklist

Instructional Delivery System

- Instruction Guide Checklist
- Media Selection Checklist

Training Evaluation

- Internal Evaluation Checklist-Criterion Test
- Internal Evaluation Checklist-Student and Instructor Evaluations
- Internal Evaluation Checklist-Supervisor Evaluations
- External Evaluation Checklist-Supervisor Evaluations
- External Evaluation Checklist-Review of Operational Data
- External Evaluation Checklist-Structured Questionnaires for

Evaluating Training

- External Evaluation Checklist-Review of Licensing Exam Results

complete and comprehensive data exist. Potential plant malfunctions are rated along three dimensions: consequence, frequency of occurrence, and difficulty of operator response. These parallel the difficulty, importance, frequency (DIF) criteria commonly used in SAT development. The key elements of the ranking scheme are highlighted in Table 2. Malfunctions would be considered important for training due to a high rating on any of the three categories or a combination of the three. The details of the ranking scheme and an illustrative demonstration for its application to a specific plant are described in Ref. 6. Also included in Ref. 6 is a description of how the malfunction rating process would be incorporated into the overall SAT design process.

TABLE 2. Highlights of Malfunction Ranking Scheme

Consequences: Malfunctions given a numerical rating of 1-15 depending on the consequences, primarily implication to public safety, of the malfunction; ranges from relatively minor tech. spec. violations to Condition IV events.

Frequency: Based on actual operating experience and predicted frequency of occurrence; consider plant-specific and similar plants; consider different operating modes; rating values from 1-10.

Difficulty: Based on system/task analysis data and expert judgement rate relative difficulty on 7-point scale due to:

- (1) perceptual loading
 - a) searching and receiving information
 - b) locating and identifying objects
- (2) cognitive loading
 - a) information processing
 - b) problem solving and decision making
- (3) communication loading
 - a) communication type
 - b) communication context
- (4) motor action loading
- (5) time factor loading

THE PRESENT RESEARCH (FY 83-84)

The FY 83-84 work is addressing four process areas within SAT. First, the selection of tasks for training. Second, the description of skilled operator task performance in order to identify entry level personnel requirements. Third, the evaluation of NPP simulator training effectiveness, and fourth, the evaluation of the factors influencing successful formal classroom training. This section primarily addresses work in progress for the task selection methodology. The other three tasks are briefly referenced and will be detailed in future reports.

Development of a Task Selection Method

The selection of tasks for training has traditionally been subjective and in one form or another a result of task placement into one of three categories depending on cost and available training time. Task sorts were determined by the answers to variants of the following questions:

(1) Can the task be trained by avoiding formal classroom work and relying on on-the-job resources with actual equipment?

(2) If not, can the task be trained in the least expensive modes of instruction such as traditional platform methods?

(3) If the lowest cost methods are not appropriate, can the tasks be trained on high fidelity simulation equipment rather than actual equipment reconfigured in special training centers?

Although highly pragmatic, such an approach neglects several critical aspects of training development which the SAT has shown to be important. First, single criteria sort methods usually neglect subtle task differences that impact operator performance variables, e.g., the probable rate of forgetting or skill decay rate. Second, the distinction between skill performance difficulty and skill acquisition difficulty is often lost. That is, a skill may be very easy to acquire but very difficult to execute (e.g., due to physical demands), or vice versa. Third, simple subjective sort schemes usually do not consider the full range of training options actually available to the training developer including alternative media and method combinations which can be traded off to meet development requirements.

Fourth, the implications of previous training experience are not systematically considered in the selection of specific tasks for training. As a result, instruction may be either insufficient for new student groups or tedious for highly experienced personnel undergoing periodic refresher training. Finally, the contextual importance of the task, and particularly its safety implications, are often not systematically included. Difficulty, importance, or frequency selection criteria (the DIF methods) tend to be plant- or situation-specific and may require redefinition for each training case.

In summary, a usable task selection procedure should address at least seven areas: (1) training categories into which all tasks will be sorted; (2) dimensions both subjective and objective which will define the categories; (3) mapping of the set of dimensions onto the set of task sorting categories; (4) scalar ranges along which dimensions will be defined; (5) mechanisms for data handling during the sorting process; (6) metrics defined to summarize numeric information, and (7) format which the output from the sorting procedure will take. Although the literature for this area is not extensive, there is some excellent recent work which addresses three of the above areas (Dawdy and Hawley,⁷ Zoints and Wallenius,⁸ Swezey⁹). The Dawdy and Hawley work in particular applied multiattribute utility theory to the subjective evaluation of task worth

and provides an excellent starting point for task dimensions and sort categories. By utilizing a similar logic, safety related task categories can also be developed and weighted. Scalar ranges on the other hand, are industry-specific (for example, the standardized safety categories of abnormal events) as are the mechanisms for data handling and appropriate output formats. Based on the above work and other relevant sources, e.g., Handbook for Design of Instructional Systems,¹⁰ the following task sort dimensions have been developed:

- (1) Skill acquisition difficulty
- (2) Skill performance difficulty
- (3) Need for immediate performance
- (4) Poor performance consequences
- (5) Previous nuclear experience
- (6) Task performance frequency in normal operations
- (7) Potential for performance in emergency operations
- (8) Plant performance delay tolerance.

Items #1 and #2 have already been discussed. Item #3 refers to the likelihood that a recently trained operator may be required to perform a safety task immediately after beginning work without extensive on-the-job training. Item #4 currently uses safety categories 1-4 based on the potential severity of task consequences if poorly performed. Item #5 refers to the likelihood that an anticipated student population would have previous experience in the skills required by a given task. Item #6 addresses the anticipated frequency a task may have in normal operations. This category is designed to identify tasks which may be part of emergency sequences but have little operator practice as a function of normal job activities. Item #8 refers to the tolerance of the NPP safety systems to a delay in operator performance and represents whether or not rapid task performance may be required in the event of transients.

Method

Using the above dimensions as category definition components, the actual sort categories for training may now be defined by: selecting which dimensions impact each sort category; and what range of values within that dimension correspond to the acceptable range for a particular type of training activity. For example, if the training category "refresher training" is used, a definition for refresher training is defined as having four main components: (1) skill acquisition difficulty; (2) skill performance difficulty; (3) poor performance consequences; and (4) potential for performance in emergency operations. Each dimension in

turn is given an acceptable range of values on a 9-point scale. In the above case:

- (1) (skill acquisition difficulty) > 3
- (2) (skill performance difficulty) > 4
- (3) (poor performance consequences) > 2
(e.g., FSAR Category III or IV)

and (4) (potential for performance in emergency operations) > 5

where #1 is the lowest value on a subjective scale and #9 is the highest value.

Other specialized training sort categories are defined in a similar fashion depending on specific evaluative needs by NRC. A tentative list of useful categories along with some preliminary estimates of dimensional ranges is provided in Table 3.

New task categories may be added and, after open discussion and debate, the criteria may be formalized. The actual criteria values may be obtained either subjectively or through methods such as the skilled performance model being developed in the skill identification subtask of this effort. The intent is to make all task selections as objective and verifiable as available safety research and plant operating data permits.

Once the mapping of task dimensions into sort categories has been completed, the full power of such an approach can be actualized. At least two types of metrics defined on the dimensions scalar values are possible. The first, an absolute criteria, which corresponds to whether a given task in a set of analyzed tasks had all dimensions for a sort category and fell within the acceptable range of scale values. Using the sum of the differences between actual and minimum acceptable values, all tasks may be rank ordered within sort categories. Because sort categories do not have to be orthogonal (for example, a task may be a safety-related task and difficult to perform), another useful metric is a relative value. In this case, the mean of dimensional deviations from the sort criteria (either positive or negative) are calculated for all tasks. Thus, this metric can be used to compare possible reallocations of tasks from one sort category to another as well as orders for all tasks within each category. This application becomes most important in the context of an overall SAT tradeoff analysis procedure where the use of a particular training approach may be simultaneously considered in the light of other factors such as manpower availability or program cost. Tentative metrics for both cases have been developed and coded in BASIC for use on a microprocessor. For example, the absolute metric calculation for a certification training would use the following logical steps:

TABLE 3. Tentative List of Task Sorting Categories

- (1) **Qualification Training**
 (skill acquisition difficulty) > 3
 (poor performance consequences) > 3
 (need for immediate performance) > 3
- (2) **Certification Training**
 (plant performance delay tolerance) > 7
 (poor performance consequences) > 7
- (3) **Refresher Training**
 (skill acquisition difficulty) > 3
 (skill performance difficulty) > 3
 (poor performance consequences) > 3
 (potential for performance in emergency operations) > 5
- (4) **Elimination from Training**
 (skill acquisition difficulty) < 3
 (previous nuclear experience) > 7
 (poor performance consequences) < 3
- (5) **Candidate for On-the-Job Training**
 (need for immediate performance) < 3
 (skill acquisition difficulty) < 7
 (skill performance difficulty) < 7
- (6) **Potential Task for Less Training**
 (skill acquisition difficulty) < 3
 (skill performance difficulty) < 3
 (need for immediate performance) < 3
- (7) **Potential Task for More Training**
 (skill acquisition difficulty) > 7
 (skill performance difficulty) > 7
 (plant performance delay tolerance) > 7
 (poor performance consequences) > 7
- (8) **Potential Simulator Tasks**
 (potential for performance in emergency operations) > 5
 (poor performance consequences) > 7
 (need for immediate performance) > 7
- (9) **Potential Formal Training Task**
 (need for immediate performance) > 5
 (skill acquisition difficulty) > 7
 (skill performance difficulty) > 7
 (task performance frequency in normal operations) < 3

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.

DISCLAIMER

Let A = [task rating on plant performance delay tolerance minus seven]

Let B = [task rating on FSAR consequences of poor performance minus seven]

Then if A or B < 0, exclude the task from Category 2, "Certification Training"

Else put the task in the membership set of Category 2.

An example of the same calculation for a relative metric of a certification task is:

Let A = [rating on plant performance minus seven]

Let B = [rating on FSAR consequences minus seven]

Then let task value for Category 2 = $\left[\frac{A}{9} + \frac{B}{9} \right]$ where 9 is the max. scale value.

A final requirement for the task selection procedure concerns the way in which the task sort information will be used by the training analyst. Currently, this problem is being handled through the development of a series of flexible sort algorithms which order values by tasks, along sort categories, or as a function of both. The first case permits the selection of tasks for each sort category to rank from the best task candidates to the worst. The second permits examination of which category is best for each task. The final sort permits the production of a recommended list of task allocations based on a numerical analysis of a task's ratings across all categories simultaneously, and readjusts the recommendations should any task not be placed in a recommended category. This sort is currently being coded for microprocessor use and has the potential of using either or both of the metrics described above. In summary, a powerful tool for training task selection in the nuclear industry is being developed and is currently being refined to incorporate the safety task identification needs of the NRC/RES group and at the same time permit the addition of new categories or the refinement of old criteria based on field test data.

Other Tasks

The remaining three tasks exceed the scope of this paper, but the approach can be summarized as follows. The identification of entry level skill, aptitudes, and knowledge requirements is being performed by developing a new method for linking a behavioral or in-plant task description to a taxonomic or psychological task description. This linkage permits the rapid identification of both personal aptitude requirements as well as potential measurement tests. The evaluation of simulator training is being performed through extension of techniques which assess the potential for transfer of training and simulator physical and functional fidelity requirements. Finally, the evaluation of formal classroom training is considered through an analysis of six sources of SAT degradation which impact upon the actual in-plant performance of operators and the identification of diagnostic and remedial procedures to enhance student training.

CONCLUSION

As the result of post-TMI efforts by both industry and NRC to assess and improve operator training and qualifications, the industry is moving to adaptation of the Systems Approach to Training and the NRC is in parallel moving to adapt SAT methodology in their evaluation of training and qualification. This research program is intended to provide the research support and technical basis for NRC's evaluation. A framework for an SAT-based evaluation process has been developed and efforts are underway to develop the specific methods and tools to implement the process.

REFERENCES

1. Jones, D. W., D. K. Baer, and C. C. Francis, "Nuclear Power Plant Training Simulators: Their Use in Operator Training and Requalification," NUREG/CR-1482 (ORNL/TM-395), July 1980.
2. Belsterling, C. A., S. Pandey, J. C. Praskiewicz, E. J. Kozinsky, L. Felkins, and P. M. Haas, "Specification and Verification of Nuclear Power Plant Training Simulator Response Characteristics, Part I: Summary of Current Practices for Nuclear and Non-Nuclear Simulators," NUREG/CR-2353, Vol. 1 (ORNL/TM-7985/P1), December 1981.
3. Haas, P. M., D. L. Selby, T. W. Kerlin, and L. Felkins, "Specification and Verification of Nuclear Power Plant Training Simulator Response Characteristics, Part II: Conclusions and Recommendations," NUREG/CR-2353, Vol. 2 (ORNL/TM-7985/P2), May 1982.
4. Strahm, K. A., "Excellence in Training: Our Goal for the 80's," Proceedings of the Fifth Symposium on Training of Nuclear Facility Personnel, CONF-830404, pp. 31-44 (April, 1983).
5. Haas, P. M., D. L. Selby, M. J. Hanley, and R. T. Mercer, "Evaluation of Training Programs and Entry Level Qualifications for Nuclear Power Plant Control Room Personnel Based on the Systems Approach to Training," NUREG/CR-3414 (ORNL/TM-8848), September 1983.
6. Selby, D. L., and W. T. Hensley, "A Ranking Scheme for Making Decisions on the Relative Training Importance of Potential Nuclear Power Plant Malfunctions," NUREG/CR-3523 (ORNL/TM-8950) (Draft; expected publication - November 1983).
7. Dawdy, E. D., and J. K. Hawley, "A Forecasting Method for Training Effectiveness Analysis," Proceedings of the Human Factors Society 26th Annual Meeting, Seattle, Washington (October, 1982).
8. Zoints, S., and J. Wallenius, "Identifying Efficient Vectors: Some Theory and Computational Results," Operations Research, 28, pp. 785-793 (1980).
9. Swezey, R. W., "Application of a Transfer of Training Model to Training Device Assessment," Science Applications, Incorporated Report (1977).
10. "Handbook for Designers of Instructional Systems," (AFP 50-58), Vol I-IV, Department of the Air Force (1974).