CONF-860415-4

TSORT - AN AUTOMATED TOOL FOR ALLOCATING TASKS TO TRAINING STRATEGIES*

CONF-860415--4

TI86 006070

R. J. Carter
C. C. Jorgensen

Reliability and Human Factors Group
Engineering Physics and Mathematics Division
Oak Ridge National Laboratory

For Presentation at:

The ANS/ENS International Topical on Advances
in Human Factors in Nuclear Power Systems

Knoxville, Tennessee
April 21-24, 1986

*Research sponsored by the U.S. Nuclear Regulatory Commission under U.S. Department of Energy (DOE) Interagency Agreement 40-550-75 with Martin Marietta Energy Systems, Inc. under Contract No. DE-AC05-84R21400 with DOE.

ABSTRACT

An automated tool (TSORT) that can aid training system developers in determining which training strategy should be applied to a particular task and in grouping similar tasks into training categories has been developed. This paper describes the rationale for TSORT's development and addresses its structure, including training categories, task description dimensions, and categorization metrics. It also provides some information on TSORT's application.

INTRODUCTION

In a systems approach to training (SAT), a variety of analyses must be performed which require the subjective expertise of a training system developer. One major analysis is the determination of where (which type of training environment) individual job tasks should be trained and how they should be ranked relative to each other and different instructional approaches. For example, when a course of instruction is produced, a training developer needs to decide whether a task should be trained in an on-the-job setting or in a formal classroom environment. In another situation a trainer may need to identify which tasks should be given extra training if resources permit.

Due to a lack of standardized SAT support tools, such decisions are often made in a very subjective manner. Depending upon the skill of training development personnel, the resulting allocation of tasks may or may not be made properly. In an SAT, the kinds of courseware developed, the media and methods used, and the performance evaluations made are directly influenced by decisions based on a general training strategy. There is thus a "ripple effect" from poor decisions which may have been made early in the process.

A tool called TSORT (task sort) was developed at Oak Ridge National Laboratory for the Office of Nuclear Regulatory Research at the Nuclear Regulatory Commission (NRC) to remedy this problem. It was originally designed to provide: (1) a standardized method to select tasks for use in NRC-sponsored training research and (2) an objective basis to determine whether industry training program developers hasigned nuclear power plant (NPP) tasks to appropriate training strategies. TSORT, however, has application for the nuclear power industry as a whole and can be used by a utility's training department to help determine which training strategy should be used for a given task and to allocate task groups within training strategies.

DEVELOPMENT DESCRIPTION

To determine which training strategy should be applied to a particular task, it was necessary to develop a tool to relate requirements of tasks to potential training approaches. This required that both the feasible training strategies and the dimensions that describe task information had to be specified.

Assessment of training strategies led to nine categories which might apply during an SAT implementation in the nuclear industry (although the tool developed need not be limited by those selections). The categories used were drawn from Department of Energy and NRC sources, as well as from previous field experience in non-nuclear SAT training developments, including the military. A study of task characteristics resulted in ten dimensions which correspond to decision-making criteria frequently applied by expert training developers when they allocate tasks to different training strategies.

In arriving at a decision, not all task dimensions are of equal importance for each strategy. Therefore, it was necessary to develop a weighting scheme, which added decison-making importance to dimensions, and to design anchored-rating scales. The result of this process was a nine-by-ten matrix whose cell values represent rating ranges above or below which a training developer would generally agree that a training dimension should play a significant role in a particular task allocation decision (see Figure 1).

After a matrix for decision criteria had been generated (through interviews with an industry trainer), it was then possible to rate particular tasks using the same set of dimensions. By comparing the rated values for a particular task to the criteria ranges for each strategy, it is possible to numerically determine whether a task "fits" better in one category or another (see Figure 2). To make the fitting process mathematically rigorous, it was also necessary to determine how a numeric score or "metric" should be produced to describe the fit. A well defined metric is important because a training developer is interested not only in which category to place a task, but also how to shift tasks among different categories as resources and time constraints change training priorities.

A numerical analysis of large numbers of tasks implied that manual procedures would be cumbersome and subject to frequent user errors. This has previously been the case with many otherwise worthwhile rating approaches. As a result, it was decided from the onset to computerize TSORT in a form that required the least possible investment of user time and effort. The result was a completely automated, menu-driven task sorting procedure capable of making recommendations for individual task allocations during training and rank ordering sets of tasks within and between allocation categories.

Since a training analyst sometimes requires information that goes deeper than a training selection, TSORT was designed to include an independent rank ordering process that considers each task dimension individually. Another option was also incorporated to demonstrate the feasibility of an SAT analysis that includes cost/benefit tradeoffs. Due to extremely high costs of safety related nuclear accidents, this capability provides the industry with an approximated insight into the cost avoidance benefits of increasing plant safety through better training.

TRAINING CATEGORIES

TSORT's nine training strategy categories include qualification, certification, refresher, and on-the-job training, and candidates for elimination, less or more training, and simulator or formal training. They are described below.

The qualification training category refers to tasks that contribute to safe plant operation but are not so critical that specific training is required. These tasks would generally require training to a clearly specified standard of performance, before a trainee would be considered as having successfully passed a course module, prior to plant operation.

Certification training is used to determine tasks whose performance is so crucial to system operation that each operator must be certified as having the ability to perform them prior to being permitted to operate in the NPP environment.

Because some tasks contain skills that tend to degrade quickly, i.e., show poorer performance over time, tasks can require periodic refresher training to assure that performance would not be compromised in a plant environment. These tasks can also include qualification or certification tasks that are seldom used, but must be capable of immediate performance should emergency needs dictate. Tasks falling into this category are generally included in institutional training.

Elimination candidates are tasks with a high probability that trainees have already been exposed to similar task demands through previous experience in NPP environments or through previous exposure, such as Navy experience or academic course-work. As a result, if time or budget pressures require tradeoffs in a plant training program, these tasks would be logical first candidates for omission.

On-the-job training tasks require site training or can readily be learned after an operator leaves formal training. Often the tasks involve simple skills and can be quickly learned through demonstration or verbal instructions. In some cases they may be difficult tasks that require close monitoring and are not amenable to standard classroom instructional methods.

Candidate for less training tasks have a high probability of previous exposure as a result of normal plant operations. They are liable to be both familiar and well practiced.

These tasks differ from "elimination candidates" in that they must still be included in training programs due to their importance to plant operations but can receive less emphasis should an instructor need to reallocate training time to areas that had not originally been planned.

Candidate for more training is a task which is so important that if any extra time is available, an instructor would want the trainee to repeat and reemphasize that task to assure that the subtask steps have been thoroughly practiced. Tasks for certification or qualification are often appropriate for this type of emphasis but are not the only cases that can occur.

Candidate tasks recommended for simulator training comprise a special instance because simulators have some unique advantages over less complex forms of training equipment such as slide projectors or mock-ups. Simulator tasks are best suited for situations that require dynamic behavior or real time performance with heavy interactions among plant systems and operators. These tasks are also generally associated with static test score measurements that are difficult to generalize to dynamic emergency or accident scenarios.

The candidate for formal training can include a dynamic task but is particularly sensitive to special task requirements including high skill-acquisition difficulties and knowledge acquisition. Formal training tasks can require a large knowledge base which must be drawn upon during plant operation.

TASK DIMENSIONS

TSORT's ten task dimensions consist of skill-acquisition and performance difficulty, immediate performance need, safety and economic consequences, previous nuclear experience, normal and emergency operation performance, plant delay tolerance, and regulatory requirement. Each is outlined in the following paragraphs.

The first rating dimension concerns the difficulty a student would have in acquiring skills required for performing a task. The skills used by a task may be drawn from either in-plant experience, SAT analysis procedures, or NRC/Institute for Nuclear Power Operations task analysis data bases.

The skill-performance difficulty dimension differs from skill acquisition in that the skills for some tasks may be very easy to acquire but very difficult to perform. For example, manual control of the feedwater system during a reactor startup is easily learned but requires skillful manipulation of controls.

Although it is not likely that a new operator will face a critical safety-related situation without supervision, it is possible, and special consideration should be given in the training program to such tasks. These tasks should be evaluated based on the potential need for immediate performance shortly after training. Typically, tasks requiring immediate performance are those that emphasize safetyrelated actions in response to plant transients.

The task dimension, safety consequences, asseses the potential impact of poor task performance in terms of radiation release to the public. An example is a failure to perform preventive maintenance scheduled for reactor scram relays that trigger rod drop into the core (as occurred at Salem Nuclear Generating Station, Unit 2 in 1983).

The previous nuclear experience dimension is the amount of prior task experience which an anticipated training population is likely to have. In some cases there may be considerable general experience but with the wrong type of procedures for large commercial reactors. In other cases the task may involve the use of common equipment which would have a high probability of previous utilization.

Some tasks may be performed frequently during normal operations and hence are likely to be well practiced by operators even without requalification training. Other tasks may seldom occur. This dimension aids in identifying tasks which will be well practiced and hence less subject to skill decay.

Some tasks are performed frequently in both normal and emergency operations so that in an emergency an operator would probably already be familiar with the tasks. Others occur only in emergency operations and are infrequent.

Another dimension is how tolerant a plant is to an operator response delay. Plant sensitivity to some tasks may be very low; others may require rapid operator decisions. Those tasks needing quick operator responses within a particular plant must be well practiced.

The ninth dimension is whether or not a task is mandated for training through regulatory requirement. This dimension by itself could require the equivalent of test certification training if its rating is high enough.

Economic consequences must also be considered in determining training tradeoffs. Although dollar cost will often correlate with safety consequences, the two can diverge dramatically depending upon whether or not poor performance would cause radioactive release from the plant containment.

METRIC DEVELOPMENT

Although decision making is still a subjective process in an SAT, it is evident upon close examination that expert training developers actually apply several criteria in their decisions. One technique that has received attention as an appropriate method to capture such judgements is called multiattribute utility theory. It has as an underlying principle that complex subjective decisions can be considered as an additive series of individual value judgements which when summed form an overall measure of desirability of one course of action over another. Because each element in a decision may have differential importance, some schemes also add weighting values to increase the sensitivity of the process.

In TSORT, the ten task dimensions are the decision elements with one training category preferred over another based on a total score achieved across all dimensions. Such rating decisions are not usually "all or none", so a mechanism had to be created to account for the partial applicability of a dimension. For example, a task could have low skill-performance difficulty but still be important if performance occurs in a situation with severe safety consequences. Therefore, it was necessary to expand the ten dimensions along a scale rather than use a single value. Nine rating levels were selected to provide adequate range for subjective differences, although the choice is arbitrary and could be more or less if warranted by predefined criteria such as maintenance intervals or cost accounting categories.

Since raters frequently differ as to a correct subjective value to assign, anchored-rating scales were developed for each of the ten dimensions. By using the rating scales it becomes possible to assign ten values to each task, one for each of the task dimension decisions. Figure 3 shows a sample rating sheet for "skill-performance difficulty". The type of information required of the rater is stated at the top of the form and 3 anchoring tasks are listed at the bottom left-hand side that are drawn from plant situations which correspond to low, medium, and high numeric levels of skill-acquisition difficulty.

Before this information can be used to choose a particular training category, each task's values must be compared to appropriate acceptability criteria stored in the matrix that defines the categories (Figure 1). If a dimension falls into the acceptability range it is considered in the final training decision; otherwise it is omitted because its inclusion would bias the final number used to rate all categories against each other. These cases are shown by the letters N/A in Figure 1. If a user wishes to change the criteria, all that needs to be done is to modify the cell values.

METRIC TYPES

One general metric to measure the "fit" of a task to a training strategy might be the sum of all acceptable dimension values divided by the total number of values appropriate for a particular sort category. Such a metric however makes a hidden assumption. It is that a training analyst is not interested in how "close" a task comes to fitting into a category, only that it has a large enough average value to be acceptable. As a result, this metric was relaxed for TSORT so that it instead counts how many of the criteria are acceptable. This relaxed metric represents a total of how many task ratings fall within an acceptable range. A "count metric" is then the first type of metric which can be used in TSORT for producing category recommendations. It works well when the primary concern is about a training strategy choice for individual tasks.

When the topic of interest is the rank order between tasks or ranking of possible strategies relative to each other, the all-or-none nature of a count metric can result in a loss of useful numeric information. This occurs because the metric will not record how far a task is from meeting a criterion. Thus, a task might be a single rating point below the cutoff and be missed as a potential category candidate. For this reason, a second metric was developed which keeps track of positive and negative deviations from the cutoff score for each dimension. This "stringent metric" has a value which is: zero if there is a perfect fit, positive if the majority of values exceed the minimum cutoff criteria for the decision, and negative if the majority of values fall below the criteria. Using this metric, it then becomes possible to rank order sets of tasks and categories in descending order from the best choice to the worst. The choice of which metric to select in a given situation is left up to the analyst.

AUTOMATED TOOL

TSORT has been programmed for use with an IBM personal computer. Extensive utilization has been made of color graphics. The computer program has been designed so that user interaction is turn-key. The user is also prompted step-by-step through data input to the results. For an entire session, or at any point in the session, a simultaneous hard copy printout of all questions asked, answers given, and results can be generated.

All options in the sort program are presented to the user as either questions or menus. In the case of questions, a necessary response is indicated and is entered by typing appropriate text or numbers. For menus, the correct response is to type the number of a desired menu element.

The main menu presents the analysis options (see Figure 4). The first two menu items produce ranked categories for each task using either the count or stringent metric. Options 3 and 4 rank order all tasks relative to each other within a single training strategy. This information is useful in determining which tasks, in order of preference, should be selected or eliminated from a particular strategy. Menu items 5 and 6 provide recommended categories for each task. These two options assume that there are no constraints in regards to which task is placed in which category.

Menu item 7 in the main menu presents the user with a new menu containing 11 other options (see Figure 5). Items 1-10 are designed to generate special sorts of task dimension information and to permit a user to quickly determine tasks which have unusual values along dimensions. When the user selects option 11, "special economic analysis", a series of questions about plant costs are displayed.

POSSIBLE FUTURE DIRECTIONS

TSORT embodies a highly general method for capturing subjective judgement processes by breaking a decision into a series of dimensions. These dimensions need not only be for training; they could also be very effective for a number of different purposes. To reach its full potential, the TSORT code should be made more flexible. Flexibility could be increased by making the dimensions, strategies, rating criteria, and rating logic interactively defined. TSORT could then be easily customized to a host of new problems. The next logical step in expanding the present effort would be to assemble a group of training development personnel to formally define and agree upon rating values for the matrix in Figure 1 and to add, subtract, or modify dimensions and categories as appropriate.

Training Strategy Categories

	Dimensions*	1 Qualification Training	2 Certification Training	3 Refresher Training	4 Elimination Candidate	5 On-the-Job Training	6 Candidate Less Training	7 Candidate More Training	8 Candidate Simulator Training	9 Candidate Formal Training
1	Skill Acquisition Difficulty	N/A	N/A	>3	<3	<3	<3	>7	>5	>7
2	Skill Performance Difficulty	N/A	N/A	>3	<3	<3	<3	>7	>5	>7
3	Immediate Performance Need	>3	>5	>8	N/A	<3	<3	>7	N/A	>5
4	Safety Consequences	>3	>7	>3	<3	<3	<2	>7	N/A	>7
5	Previous Nuclear Experience	N/A	<2	<3	>7	>7	>5	<5	<3	<5
6	Normal Operation Performance	>3	N/A	<3	N/A	>5	>7	<3	<3	<3
7	Emergency Operation Performance	>5	N/A	>5	<3	<3	<3	>3	>5	>3
8	Plant Delay Tolerance	<2	>7	>7	<3	<3	<3	>7	>5	N/A
9	Regulatory Requirement	>7	>3	>3	<1	N/A	N/A	N/A	>3	N/A
10	Economic Consequences	>5	N/A	N/A	<2	<2	N/A	>7	>5	>3

*N/A means the dimension was rated not relevant by nuclear training personnel based on the scale criteria used. Scale values range from 0 to 9 — < means less than and > means greater than.

Figure 1. Ten Task Dimensions by Nine Training Strategy Categories.

Sample Task Rating		Refresher Training		Formal Training	
Dimension	Actual Task Rating	Ideal Criteria	Deviation	Ideal Criteria	Deviation
1	3	<3	+0 (true)	>7	−4 (false)
2	7	>3	+4 (true)	>7	0 (false)
3	1	>8	−7 (false)	>5	−4 (false)
		Total (Σ)	−3		−8

Result: Category 1 (refresher training) fits the sample task better than Category 2 (formal training) because the total sum of deviations for Category 1 was less negative than for Category 2.

Figure 2. Establishing Whether One Task "Fits" a Category Better Than Another (Two Categories with Three Dimensions Each Taken from the First Three Dimensions of Figure 1).

Skill Performance Difficulty

Defined in terms of physical and cognitive effort or degree of precision required

VALUE	CRITERIA
0	Easily performed with trivial effort (> 99% can perform)
1	Easily performed with little precision
2	Easily performed with some precision
3	Some performance difficulty, no decision making
4	Some performance difficulty, occasional decision making
5	Requires some physical effort or cognitive effort with decision making
6	Definite physical effort or cognitive effort with decision making
7	Same as #6 with some precision
8	Heavy cognitive and/or physical effort with precision
9	Extended physical effort, heavy decision making, and stringent performance requirements

VALUE ANCHORS	SAMPLE TASK
0	Read a digital water level meter out loud
5	Determine that a reactor scram was caused by a normal turbine trip
9	Align fire system for core cooling following a LOCA and loss of all normal and ECS makeup

Figure 3. Sample Rating Sheet for Skill Performance Difficulty.

```
WHAT KIND OF ANALYSIS DO YOU WISH TO PERFORM?

  TYPE 1 RANKED CATEGORIES FOR EACH TASK USING MATCH VALUES
  TYPE 2 RANKED CATEGORIES FOR EACH TASK USING AVERAGE VALUES
  TYPE 3 RANKED TASKS FOR EACH CATEGOPY USING MATCH VALUES
  TYPE 4 RANKED TASKS FOR EACH CATEGORY USING AVERAGE VALUES
  TYPE 5 RECOMMENDED CATEGORIES FOR EACH TASK USING MATCH VALUES
  TYPE 6 RECOMMENDED CATEGORIES FOR EACH TASK USING AVERAGE VALUES
  TYPE 7 SPECIAL INPUT DATA SORTS
PLEASE SELECT A NUMBER FROM THE ABOVE LIST? 1
```

Figure 4. The Main Sorting Program Menu.

```
  TYPE 7 SPECIAL TASK RATING SORTS
PLEASE SELECT A NUMBER FROM THE ABOVE LIST? 7
SPECIAL OPTIONS ARE:
 1 A RANKED ORDERED LIST OF SKILL ACQUISITION DIFFICULTY ON TASKS
 2 A SIMILAR LIST OF SKILL PERFORMANCE DIFFICULTY
 3 A RANK ORDERED LIST BASED ON IMMEDIATE PERFORMANCE
 4 A SIMILAR LIST BASED ON PUBLIC SAFETY RISK
 5 A RANK ORDER OF TASKS BASED ON PREVIOUS EXPERIENCE LIKELIHOOD
 6 A RANKED LIST OF TASK PROBABILITY IN NORMAL OPERATIONS
 7 A SIMILAR LIST FOR EMERGENCY OPERATIONS
 8 A RANKED LIST OF TASKS BASED ON PLANT DELAY TOLERANCE
 9 TASKS RANKED IN TERMS OF REGULATORY REQUIREMENT CONSTRAINTS
10 TASKS RANKED ON TERMS OF POTENTIAL ECONOMIC CONSEQUENCES
11 SPECIAL COST BENEFIT ANALYSIS MODULE
PLEASE CHOOSE A NUMBER OR USE 12 TO QUIT?
```

Figure 5. A Listing of Special Menu Options.