A Direct Perception Interface for Nuclear Power Plants

Neville Moray, John Lee, Kim J. Vicente, Barclay G. Jones and Jens Rasmussen
Proceedings of the Human Factors and Ergonomics Society Annual Meeting 1994 38: 481
DOI: 10.1177/154193129403800905

A DIRECT PERCEPTION INTERFACE FOR NUCLEAR POWER PLANTS

Neville Moray[1], John Lee[2], Kim J.Vicente[3], Barclay G. Jones[1], Jens Rasmussen[4]
[1]University of Illinois at Urbana-Champaign, [2]Battelle Human Resources, Seattle,
[3]University of Toronto, [4]HURECON, Smørum,Denmark

Following the suggestions of Beltracchi (1987) a direct perception interface for the thermal hydraulic systems of a pressurized water nuclear power reactor (PWR) was developed. It presents operators with an animated graphic of the Rankine heat cycle describing the functional relations of steam generation in a PWR. The ability of students of thermal and nuclear systems to recall system states, and detect and diagnose nine transients was compared to that of experienced nuclear power plant operators. The results were compared to a display representing traditional analog meters. The direct perception interface supported better diagnostic performance, but did not improve memory for quantitative information. Problems in evaluating such displays are discussed, in particular concerning choice of scenarios, and investigation of failure modes of advanced displays.

INTRODUCTION

In large complex systems the number of displays may cause great difficulties for operators who must monitor the system. A common example is the control room of a PWR where there may be several hundred displays. Beltracchi (1987) suggested that a direct perception interface (DPI) could be developed based on the Rankine heat cycle which is a canonical representation of the thermodynamics of steam generation in power plants. Such an interface is shown in Figure 4. We have compared the efficacy of such a display with a set of simulated analog meters, using both naive and expert monitors, and for a variety of transients requiring detection and diagnosis. Full details of this work can be found in Moray et al. (1993).

METHOD

Realistic NPP data for nine transients were presented to three groups of observers. The data were obtained from a full scale PWR training simulator and showed the values taken by 35 variables during transients lasting approximately 5 minutes. Three groups of 42 observers each were used: they were undergraduate mechanical engineers with one or two semesters of thermodynamics courses, graduate students in thermodynamics or nuclear engineering, and professional PWR operators. Each observer viewed the nine transients in a different order.
Each group was divided into three subgroups of 14 observers. One subgroup monitored a display simulating 35 analog meters, one for each variable, a "single-sensor-single-instrument" (SSSI) display.

The second monitored the Beltracchi Rankine Cycle display (Beltracchi, 1987). The third subgroup monitored a display modified to include a temperature-pressure dynamic graphical plot, but the results of this group will not be discussed for want of space.

At the end of the transient the display was blanked out. Each group of observers was divided into two further subgroups of 7 observers. The first, the "Quantitative Recall" group, was asked to recall the exact value of the 35 state variables, to state whether there had been anything abnormal, and then, if so, to diagnose the nature of the transient. They indicated their memory of the state variable values by pointing with a mouse to each of 35 analog meters on the screen to indicate the value they recalled.

The second, "Qualitative Recall" group, was asked qualitative questions such as, "Did the primary coolant remain sub-cooled for the duration of the trial?", and then asked whether anything abnormal had happened, and if so to diagnose the nature of the transient.

The use of memory tests as a method of evaluation was suggested by the work of DeGroot (1965), where it was found that experts were·able to recall more details of a complex peatter when the pattern was related to their expertise, and was presented in a way which was semantically coherent. It was thought that DPIs would provide more semantic coherence than individual meters, and hence would support better recall and better diagnosis.

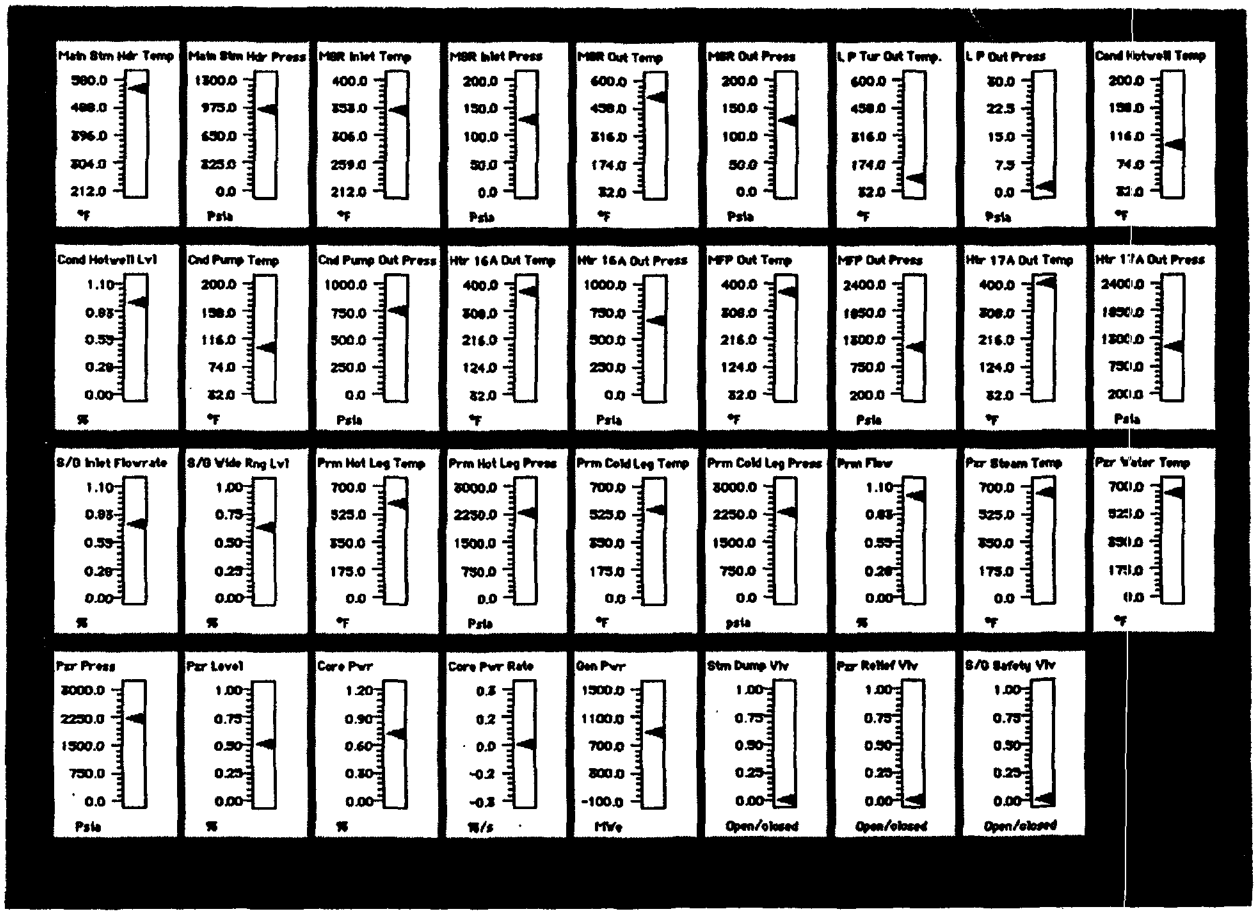

Figure 1. Simulated analog meter display

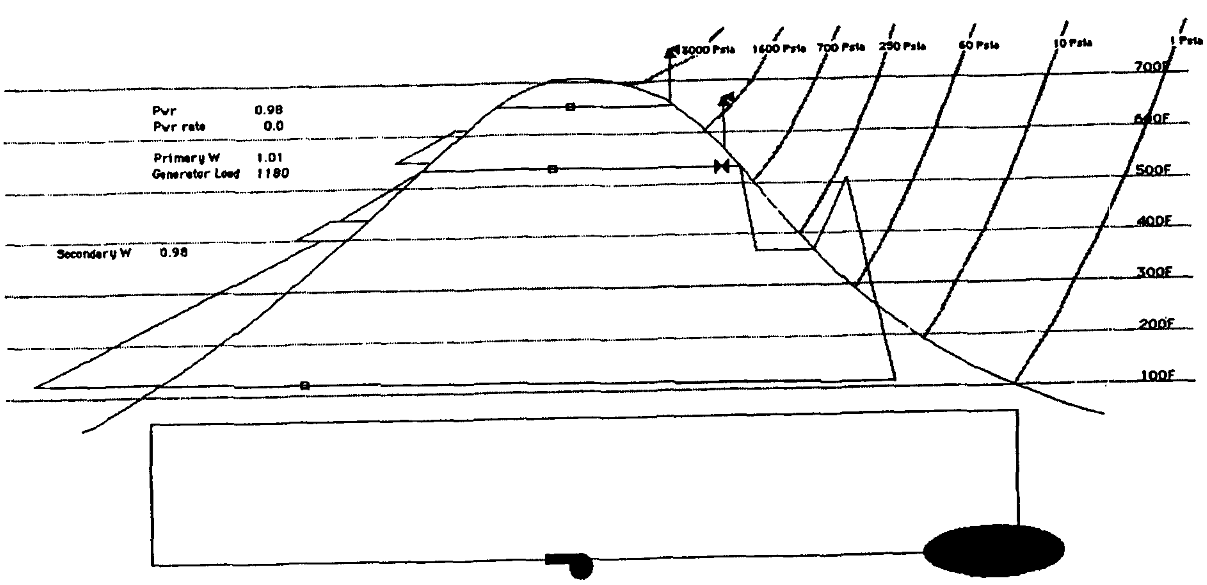

Figure 2. Rankine cycle display

Each observer saw nine transients. The formal statistical design was thus a repeated measures design over transients, and a between-subjects design by display type and recall type.

RESULTS

The main results of our experiments are shown in Figures 3 and 4, which show the scores for recall and diagnosis as a function of display type and expertise. Scores on both left and right ordinates of the figures are plotted so that poorer performance is upwards.

Recall, diagnosis and expertise.

Recall scores are not in general reliable ways to predict diagnostic performance. Figure 3 shows a decrease in incorrect diagnoses from undergraduates to graduates to NPP operators. This difference is

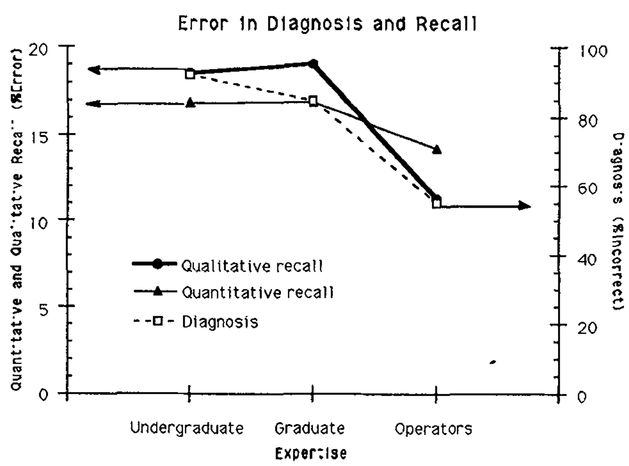

Figure 3. Performance as a function of expertise

highly significant statistically. The difference between either of the student groups and the NPP operators is very large, approximately 25%, indicating the efficacy of the NPP training program. There is no difference in the quantitative recall scores between the undergraduates and graduate students, and a small but significant reduction in quantitative recall from the students to the NPP operators. There is no difference in the recall of qualitative information between the student groups, but there is a very large and highly significant difference in the recall of qualitative information between the student groups and the NPP operators. These results suggest qualitative recall is more sensitive to changes in expertise and that operators' expertise consists in thinking about transients in terms of system state rather than in terms of the value of state variables. The relative insensitivity

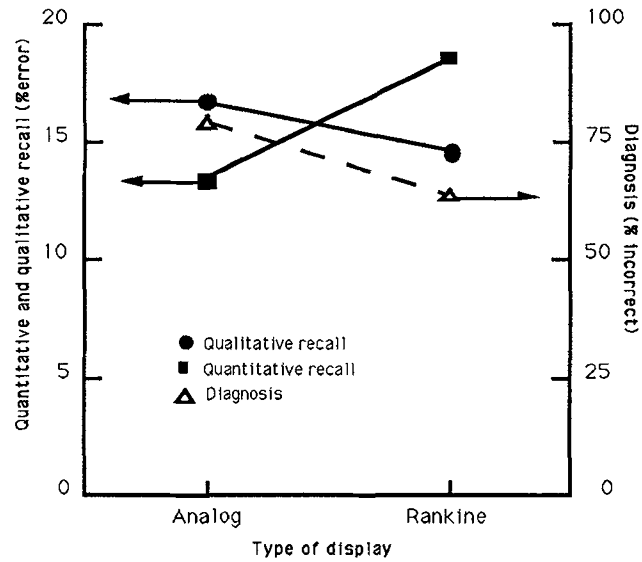

Figure 4. Performance as a function of display type

to changes in expertise of quantitative recall suggests that this measure is not a good measure of expertise.

Display type and performance

Figure 4 shows the relation of recall to diagnosis as a function of display type, and reveals dissociation among the measures of performance. The Rankine cycle display is significantly and substantially better than the analog displays in supporting diagnosis. Quantitative recall scores on the other hand are best for the Analog (SSSI) display, . Recall of qualitative state information shows no significant difference between the displays, although there is a nonsignificant improvement with the Rankine display.

Quantitative recall is not a good predictor of diagnostic performance. There are significant improvements in the recall of quantitative information in situations where diagnostic ability declines. Diagnosis is more strongly coupled to the recall of qualitative state information. This would be expected if recall of qualitative state information is a precursor of labeling the state of the whole system qualitatively as being in a particular transient state, and hence should be related to diagnosis.

CONCLUSIONS

Evaluation of DPI.

Our conclusion is that a recall test in which the observers are asked to reconstruct from memory the values of state variables is not a reliable and

Our conclusion is that a recall test in which the observers are asked to reconstruct from memory the values of state variables is not a reliable and

sensitive indicator of performance, where good performance is defined as correct diagnosis of transients. On the other hand, memory for qualitative state changes, where what is recalled is not particular values of variables, but whether or not subsystems entered abnormal states, may correlate well with diagnostic performance. The latter test showed the improved expertise of trained NPP operators compared with students who are knowledgeable about thermodynamics but not about NPP operation. It is clear that the DPI supports better diagnostic ability when used by experts. It is also clearly necessary, if such displays are implemented, to support in some secondary way the provision of accurate quantitative information for situations where it is needed.

Effects of scenarios.

There were large differences in diagnostic and other performance measures as a function of the scenarios. (See Figure 5.)

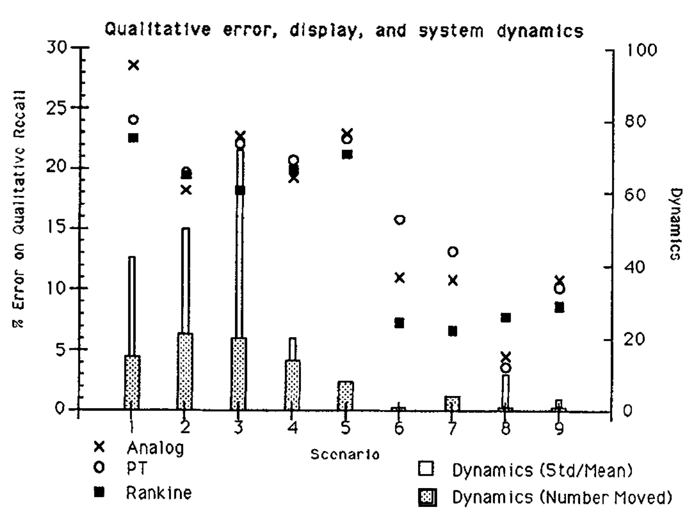

Figure 5 Relation between the qualitative memory score and scenario dynamics. The thin open bars are a measure of the range over which the value of the variables changed, namely the standard deviations of the values of all the variables divided by their means. The thick dotted bars are the number of variables (out of 35) which changed at all during the scenario.

When many of the displayed variables changed their values diagnosis was much harder than scenarios where few variables changed. These differences did not alter the rank order of the performances as a function of the type of display or the type of observer; but they raise important questions about

the choice of scenarios when simulators are used as part of the licensing process, when scenarios are used to assess operators working as members of teams. If only one or two scenarios are used imperfections in the skill of an operator might be concealed by his team-mates in an easy scenario, or his performance hindered by team mates in a difficult scenario. Hence the choice of scenarios for use in evaluation becomes critical, and research is needed to provide a scientific basis for the choice of scenarios.

DPI failure modes

Figures 6 and 7 show two Rankine displays, one for a loss of coolant accident (LOCA) and the other for a failed sensor. In the latter there is no fault in the plant operation, but the instrumentation has failed. The LOCA results in a coherent and meaningful display, but the failed sensor makes the geometry of the Rankine cycle display become physically meaningless, and leads to a display which is extremely hgard to interpret, although only a single variable is faulty. In this situation the loss of a single variable in the analog display left 34 varoiables displayed in a way which allowed them to be used for assessing plant state, but the DPI collapsed into a format which would have been extremely difficult to use, even though the plant state was in fact normal.

The calculations involved in coupling information from many sources to produce the DPI are extremely vulnerable to certain classes of failures. Little or no research exists on this aspect of DPIs, and is urgently needed. The real advantages of such displays during normal operation and during many classes of transients may be more than offset if they collapse under other classes of abnormalities.

An understanding of the *failure modes of DPIs* (as distinct from the failure modes of the plant itself) is as necessary as an understanding of their design for normal conditions. Existence proofs of interface designs are no substitute for full empirical evaluations, and yet very few advanced DPIs have been exhaustively evaluated over a wide range of abnormal conditions. It is clear from our results that DPIs and/or "ecological interfaces" can fail in catastophic ways, and it is probable that such failures may be particularly dangerous in large richly coupled systems with many degrees of

ACKNOWLEDGEMENTS

This work was supported by the U.S. Nuclear Regulatory Comission under contract NRC-04-09-094.

References

Beltracchi, L. 1987. A direct manipulation interface for water-based Rankine cycle heat engines. *IEEE Transactions of Systems, Man and Cybernetics. SMC-17*, 478-487.

DeGroot, A. D. *Thought and choice in chess.* Den Haag: Moughton. 1965

Moray, N.P., Jones, B., Rasmussen, J., Lee, J., Vicente, K., Brock, R., and Djemil, T. 1993. *Development of a Performance Indicator for the Effectiveness of Human-Machine Interfaces for Nuclear Power Plants.* UILU-ENG-92-4007. Washington, D.C.: U.S. Nuclear Regulatory Commission. NUREG/CR-5977.

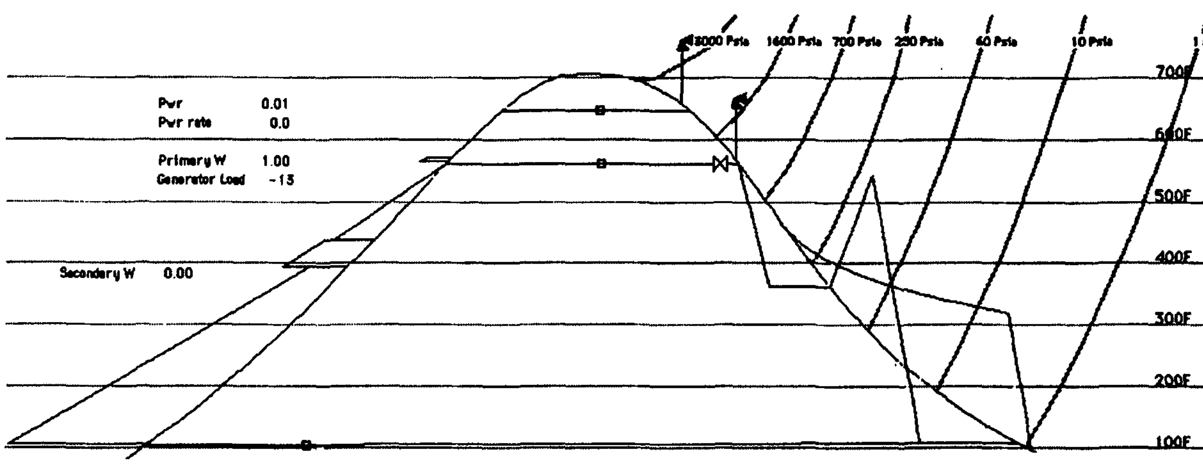

Figure 6. Rankine cycle display for loss of coolant transient

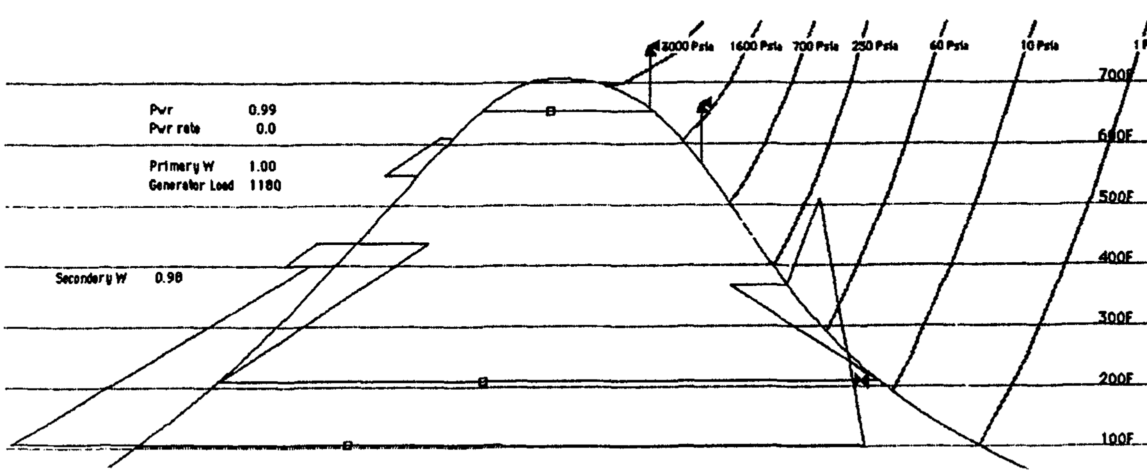

Figure 7. Rankine cycle display for failed sensor transient