EVALUATION OF ECOLOGICAL INTERFACE DESIGN FOR NUCLEAR PROCESS CONTROL: SITUATION AWARENESS EFFECTS

Catherine M. Burns
University of Waterloo

Gyrd Skraaning Jr.
OECD Halden Reactor Project

Greg A. Jamieson
Nathan Lau
University of Toronto

Jordanna Kwok
University of Waterloo

Robin Welch
Gisle Andresen
OECD Halden Reactor Project

Address correspondence to:
Catherine M. Burns
Systems Design Engineering
University of Waterloo
200 University Ave West
Waterloo ON N2L 3G1
Canada

Keywords: ecological interface design, situation awareness, process control, nuclear power
ABSTRACT

Objective: To determine whether an ecological interface display for nuclear power plant operations supports improved situation awareness over traditional and user centered displays in a realistic environment. Background: Ecological Interface Design has not yet been fully evaluated with real operators facing realistic scenarios. Method: Ecological displays were evaluated alongside traditional and user centered “advanced” displays in a full scope nuclear power plant simulation. Licensed plant operators used the displays in realistic scenarios that either had procedural support or did not have procedural support. All three displays were evaluated for their ability to support operator situation awareness. Results: A significant three-way interaction effect was observed on two independent measures of situation awareness. For both measures, ecological displays improved situation awareness in scenarios that did not have procedural support and primarily in the detection phases of those scenarios. No other pronounced effects appear across both measures. Conclusions: The observed improvement was sufficiently large to suggest that EID could improve situation awareness in situations where procedures are unavailable. However, the EID displays did not lead to improved situation awareness in the other conditions of the evaluation, and occasionally underperformed on single measures of situation awareness. This suggests that the approach requires further development, particularly in integrating EID with procedural support. Application: This research has important findings for the ongoing development of the Ecological Interface Design approach, the design of industrial operator displays, and design to support Situation Awareness.
INTRODUCTION

Ecological Interface Design (EID) is an approach for designing graphical interfaces for complex systems (Vicente and Rasmussen, 1992; Burns and Hajdukiewicz, 2004). In laboratory studies, EID has been effective at supporting problem solving, particularly in unanticipated situations (Vicente, 2002). These studies have largely explored EID in the context of small laboratory systems with single student operators (e.g. Pawlak and Vicente, 1996; Burns, 2000; Reising and Sanderson, 2002) and with teams of student operators (Garabet and Burns, 2004). While these studies have been helpful in demonstrating how EID might be a promising approach for supporting problem solving behavior in complex environments, they fall short of demonstrating that EID can support performance by real operators in real operating environments. Although applications of EID in industrial settings are emerging, the number of representative evaluations remains small and probably insufficient to justify whole scale adoption of the framework in safety-critical industries. Still, there are cases where EID is being adopted in industry, whether in whole or in combination with existing more traditional display systems (e.g. Momtahan, Burns, Labinaz, Mesana and Sherrard, 2007; Burns, Garrison, and Dinadis, 2003). This makes evaluations of EID beyond the laboratory - with real scenarios, real operators and with a realistic implementation - very important. To date, only one example of an evaluation meeting these criteria has been reported (Jamieson, 2007).

To address these concerns, we conducted a large scale evaluation of the ecological approach, using a full scope simulator, with licensed operators working through complex scenarios. We compared ecological displays to existing displays that represent the state of practice in many control rooms. We also included a third category of displays developed by subject matter experts following a user centered approach. We implemented the ecological
displays as monitoring displays within the display system that included alarm screens, navigation, and traditional displays for control. In our experience in contemporary industry, this “hybrid” system is the most common path of adoption of ecological displays. The hybrid approach has the advantage of smoothing the transition to a new display system, as traditional displays can be retained for control, and ecological displays introduced for monitoring.

Representative measures of operator performance are an important consideration in the evaluation of EID for industrial application. To date, the measures of success for EID displays have emphasized timeliness and accuracy of detection and diagnosis (e.g. Pawlak and Vicente, 1996; Burns, 2000). While these measures show that EID is promising, they do little to explain why or to tease out the subtleties of where EID works and where the approach could be improved. To answer these questions, the product measures must be combined with relevant measures of process and cognition.

We evaluated the displays for their ability to support operator Situation Awareness (SA) as modeled by Endsley (1988; 1995a; 1995b). We chose to use SA measures for two reasons. First, the notion of SA is widely accepted in the cognitive engineering community (as reflected in the extensive literature in the area). Second, SA measures could provide insight into the processes of how operators perceive, understand, and predict system states. Although other models of SA exist, we chose the three-level model because of its prevalence in the literature and the availability of measures to guide an evaluation.

Given the prevalence of the EID and SA concepts in the cognitive engineering lexicon, the limited theoretical discussion and empirical evaluation concerning both EID and SA is surprising. Hancock and Smith (1995) provided a metacognitive and ecological perspective of SA, but without any explicit reference to the EID framework. Flach and Rasmussen (2000)
applied the abstraction hierarchy framework and skills, rules and knowledge taxonomy to provide a theoretical account of situation and awareness, respectively. From these two theoretical perspectives, our examination of the theoretical foundations of EID and the three levels of situation awareness (Endsley, 1995a) leads to three general predictions. First, we expect EID to support *perception* by making the constraints on effective action visible through graphical forms that are consistent with the perceptual capabilities and limitations of the viewer. Second, since properly comprehending constraints and variables requires an understanding of their purpose and role in system function, EID should support *comprehension* by communicating the purposeful structure of the system and displaying variables in the context of functional relationships. Third, EID should support *projection* by supporting operator manipulations of the mental model that is externalized by the ecological interface. Thus, at all levels, EID should support SA.

There is some limited empirical evidence to support these predictions. Roth, Lin, Kerch, Kenney, and Sugibayashi (2001) found that functionally designed overview displays can improve operator SA. Li, Sanderson, Memisevic, and Wong (2007) used the SA content of verbalizations to evaluate ecologically-inspired functional displays for hydropower management, finding that functional displays may increase SA. A deeper understanding of how and when EID can support operator SA would be an important contribution to both the EID and SA literature.

In the following sections we describe and give examples of the three kinds of displays that we studied. We adapted SA measures already used at the OECD Halden Reactor Project and we discuss how these measures relate to the three levels of SA (Endsley, 1995a). Finally,
we describe the evaluation of the displays with measures of SA in a full scope simulation study with licensed operators working through realistic scenarios.

**DISPLAYS**

Three different types of displays were studied; 1) “ecological displays”, 2) the existing displays for the simulator, dubbed the “traditional displays”, and 3) a set of traditional displays that had been enhanced by the process experts at the OECD Halden Reactor Project, designated “advanced displays”.

**Ecological Displays**

We designed ecological displays for the secondary side of a nuclear power plant. The secondary side of the plant contains the turbines, the condenser, feedwater systems, and electricity generation systems. In contrast, the primary side of the plant contains the reactor and its immediate feed and heat exchanging systems. We chose the secondary side for two reasons: first, to manage the scope of the project, and second to improve recruitment because more operators are qualified on the secondary side of the plant.

The ecological displays were designed simultaneously by three teams, one at the OECD Halden Reactor Project, Norway, one at the University of Toronto, and one at the University of Waterloo. The Halden team designed the feedwater subsystem displays, the Waterloo team designed the turbine subsystem displays and the Toronto team designed the condenser subsystem displays. Each team developed work domain models based on the abstraction hierarchy (Rasmussen, 1985) and followed the general design process described in Burns and Hajdukiewicz (2004). The designers shared models and design artifacts throughout the design process to develop a consistent style. More details on the models and the resulting displays are
provided elsewhere (Lau, Veland, Kwok, Jamieson, Burns, Welch, and Braseth, in review). In general, the ecological displays contained functional and relational information that is absent in the existing displays and took advantage of emergent features and more advanced configural graphics. Figure 1 shows one of the traditional displays and Figure 2 shows one of the ecological displays. The traditional display (Figure 1) primarily shows the equipment (e.g. turbines, valves and piping) and key process values (e.g. pressure readings, valve positions). In contrast, the ecological display in Figure 2 shows not only the equipment on the right hand side but also other graphics, such as a temperature profile, pressure profile, and enthalpy graph, on the left side that were derived from the work domain analysis. In some cases, the ecological displays contain new information (such as enthalpy) and in other cases, they display the same information in context (e.g., pressure values in a profile that can show critical pressure increases and decreases). In all, five ecological displays were developed, two for the turbine system, two for the condenser and condensate system, and one for the feedwater system.

*************** Insert Figure 1 here **************

*************** Insert Figure 2 here **************

The ecological displays were designed to be the main displays for the processes of the turbine, condenser, and feedwater systems. For plant processes that were not represented by the ecological displays, operators had access to their traditional displays. This meant that two different styles of displays were in use; both within the turbine operator’s purview (ecological and traditional), and between the turbine operator (ecological) and the reactor operator (traditional. This limitation was accepted because only the turbine operator was being studied
and this type of hybrid implementation is representative of current industry practice. The sidebar on the left hand side of each ecological display was the same as that which appeared in the traditional and advanced displays (see Figures 1 and 2). When operators clicked on a piece of equipment in any of the displays, the equipment status and related control variables appeared in the sidebar. Operators took control actions by keying in desired values in the sidebar entry fields. Retaining the navigation and control action patterns of the traditional display had the advantage of reducing the possibility that confusion could influence operator performance. However, this choice also had the disadvantage of potentially introducing compatibility issues between the displays and the navigation and controls. Similar to contemporary industrial control centers, operators had access to multiple screens of displays as shown in Figure 3. The operators had EID displays of their main processes, a sidebar for control actions, and had access to several screens of alarm information.

*************** Insert Figure 3 here ***************

**Traditional Displays**

The “traditional” displays were the original displays for the simulator at the Halden Reactor project. They were representative of computer-based displays available in current nuclear power plants. The displays are based on plant mimic diagrams with some trended variables. They show alarm conditions and permit the operators to take control actions from the displays. Figure 2 is an example of one of the traditional displays.
**Advanced Displays**

The “advanced” displays were developed independently of our project. They were based on the mimic diagrams in the traditional displays, but contain additional graphical features designed by process experts based on their operating experience (see, Skraaning & Nihlwing, 2008, pp. 4-5, for details). Table 1 and Figure 4 summarize the additional graphical features in the advanced displays that are not available in the traditional displays. For the purpose of having a stronger test of the EID approach, the ecological design team did not access these displays. Figure 5 shows one of these displays that employs the “mini-trends,” a fairly current development and widely adopted feature in process graphics. These mini-trends were not available on the traditional displays. The advanced displays can be considered representative of a representation aiding (Bennett, Nagy and Flach, 2006) approach to display design that does not follow an ecological approach. By this we mean that the displays contained modern configural graphics, but the rationale for where the displays have been improved is different. Instead of using a Work Domain Analysis to structure the process (Vicente and Rasmussen, 1992; Burns and Hajdukiewicz, 2004), these graphics have been selected and developed based on operator insight and opinion on where more visual support is required. The approach can be considered to be user-centered and is evolutionary, as the design changes are typical of what might be added to existing displays without undergoing an extensive new graphics project.

************ Insert Table 1 here ************

************ Insert Figure 4 here ************
METHOD

In this section we discuss the method of evaluating the three displays. We discuss the general setup of the experiment, the scenarios that were used, and the design and application of the SA measures. We hypothesized that the ecological displays would provide the greatest support for SA, particularly in novel scenarios without procedural support since these situations require the knowledge-based reasoning that EID is intended to support. Finally we describe the participants and the conduct of the study.

Experiment

The experiment was conducted at the HAMMLAB facility at the Halden Reactor Project in Norway. Figure 3 shows the configuration of the experimental control room, which had a reactor operator work station on the left, turbine operator work station on the right and a large overview display (available in all three experimental conditions) in between shared by both operators. The overview display provided a condensed mimic of both the reactor and turbine sides with key instrument readings. The traditional and ecological display conditions used the same mimic-based overview. For the advanced display condition, the overview display was supplemented with three mini-trends, a set of pie graphs, and a text-based indicator of the turbine operation mode. Only one mini-trend and the text-based indicator were pertinent to the operation of the secondary side.
Operators were informed that the experimental leader would act as a representative from plant management and could call the control room for status updates during the experiment.

**Scenarios**

In all scenarios, operators were asked to maintain the original power level and safe operations. They completed three scenarios of two types for a total of six scenarios.

**Within Design Basis Scenarios**

Within design basis scenarios consist of equipment failures and job responsibilities that are anticipated by the utility. All failures in the trials of this category can be resolved by referencing the procedures that describe the various alarms in the scenarios. The three within design basis scenarios were:

1. Leakage at the intermediate superheater at 109.3% power,
2. Drain route switch and pre-heater bypass reset at 20% power, and
3. Drift of an instrument at 44% power.

**Beyond Design Basis Scenarios**

Beyond design basis scenarios consist of equipment failures or events that are not anticipated by the utility. While procedural support was still available to the operators, none of the failures in these trials can be addressed directly by referencing the procedures. Thus, solutions depend primarily on the problem solving skills of the operators. The three beyond design basis scenarios were:

1. Turbine trip while the generator is connected to the grid at 109.3% power,
2. Leakage at the condensate cleaning building at 109.3% power,
3. Sudden and high increase of seawater temperature at 109.3% power.

Each scenario had a “detection phase” or time before the first alarm sounded, then a
“mitigation phase” that consisted of all subsequent events.

Measures

In conducting research on human performance in complex situations, it can be
challenging to establish meaningful measures of performance. In a rich and realistic scenario,
actors are not constrained to a single trajectory through the scenario. Although certain paths
may be more likely, it is possible for actors to find a preferable path that was not anticipated, or
to commit an error that makes the scenario more challenging than expected. For this reason,
measures of speed or accuracy may not always be relevant. We used measures that were
robust to the complexity of operators working through realistic scenarios and were similar to
the types of evaluations that might be employed during operator training. We took advantage
of measures that had been developed, tested and used previously in HAMMLAB since they
were expressly designed to assess nuclear power plant operator performance in complex
scenarios. We adapted these measures to extract elements of situation awareness (Skraaning
jr., Lau, Welch, Nihlwing, Andresen, Brevig, Veland, Jamieson, Burns, and Kwok, 2007). In
the next sections we will discuss the measures that we used and their connection to the
standard levels of situation awareness.

Process Overview (Level 1 SA)

For our first measure of SA, we used queries that were very similar to those used in the
SACRI method (Hogg, Follesø, Strand-Volden, and Torralba, 1995). They were presented
during freezes in the scenario in a format similar to SAGAT (Endsley, 1995b; Endsley and Garland, 2000).

The questionnaires were administered at the end of the detection and mitigation phases of each scenario. The simulator was frozen and the user interface was unavailable to the operators. The questionnaire was administered electronically asking the operators to report their perceptions of various process parameters in the scenario. Simultaneously, the process expert examined trends and data in the simulator to determine the correct answers. At each freeze, 9 to 12 questions were administered. Operators were given a score of 1 for answers in agreement with those of the process expert and 0 for those in disagreement. A process overview score was calculated as the percentage of correct responses.

Questions were designed primarily to elicit the operator’s perception of the situation. Table 2 shows some typical questions that were asked. Some questions were very specific to the developing situation while others assessed the operator’s perception of global elements of the situation. Global and local questions were presented in random order and both were considered in the calculation of the overall process overview score. Results on these questions were scored to develop a measure of Level 1 SA we refer to as Process Overview.

************** Insert Table 2 here **************

**Scenario Understanding Queries (Level 2 and 3 SA)**

For our second measure of SA, we used queries that were modeled on and adapted from SA measures including SAGAT (Endsley, 1995a; Endsley and Garland, 2000), SALSA (Hauss and Eyferth, 2003) and SA/BARS (Neal, Griffin, Paterson and Bordia, 1998; 2000).
Queries occurred at predefined points in the scenario and were disguised as telephone calls from plant management to maintain the realism of the experiment. The process-specific content for each question was predefined before the start of the experiment using an approach comparable to an SA requirements analysis (Jones and Endsley, 1996), but the questioner (a process expert) was allowed to improvise slightly to match with the scenario as it unfolded. Because the queries were disguised as calls from management, operators were free to consult their displays as they would in a real situation. This also means that the operators could respond freely to the queries, which differs from SAGAT, which typically relies on an alternative forced choice format. The simulator was frozen while operators responded to queries to reduce time pressure on the responses and operator stress from simultaneously responding to the demands of the process and answering the query. Simulator freeze times were very short, usually just a few seconds. Operators were also exposed to the queries and short freezes during a training period before the experiment.

Queries were designed to elicit the operator’s level of understanding of the situation and their ability to predict future occurrences. Table 3 below shows some of the questions mapped to the level of SA.

*************** Insert Table 3 here ***************

For each query, a process expert scored the operators’ level of SA on an anchored, four-point rating scale, similar to SA/BARS (Neal, Griffin, Patterson and Bordia, 1998, 2000). The individual scores were aggregated in a measure of “Level of Scenario Understanding” which reflected elements of Level 2 and Level 3 SA.
The process overview and scenario understanding variables were meant to be complementary indicators of situation awareness.

Hypotheses

We formulated the following hypotheses for both SA measures based on theory and prior studies of EID:

Effect of Interface

Ecological interfaces have generally shown advantages over traditional designs in supporting performance in complex systems (Vicente, 2002). We expected that the ecological displays would outperform the other two designs in supporting situation awareness. We expected that the traditional design would provide the weakest support for supporting situation awareness. We did not differentiate between supporting process overview or supporting scenario understanding.

Effects of Scenario Type

We expected that beyond design basis scenarios would be more difficult for the operators than within design basis scenarios. Historically the most severe accidents occur when unanticipated situations occur.

Interaction between Interface and Scenario Type

We expected that the ecological display would provide the best support for situation awareness in the beyond design basis scenarios. EID was developed to aid knowledge based problem solving in unanticipated situations. Other research (Jamieson, 2007) has shown that the performance advantages of ecological interfaces over traditional interfaces are more
pronounced in unanticipated situations. We did not develop a hypothesis regarding the performance of the advanced display in this situation.

**Participants**

Participants were recruited from a boiling water reactor nuclear power plant identical to the simulated process. A total of six (n = 6) licensed control room operating crews participated on a voluntary basis. Each crew consisted of one reactor operator (RO) and one turbine operator (TO). In two cases, personnel currently working as ROs acted in the role of TO in the experiment. While not ideal, this substitution was considered acceptable since ROs must already hold the TO licence. Most of the operators had been to the facility before and had worked with the traditional displays on previous visits.

**Procedure**

The experiment was conducted over a span of three weeks with two crews per week. Each week had three days of training and data collection. A one-day training session at the start of the week was held to familiarize the operators with the research facility, the simulator, and the information displays with a one hour review period on the first day of experimentation. A background demographics questionnaire was also distributed during the training session. The total training before the experiment took six hours. Participants received one hour of training on the traditional and advanced displays and three hours on the ecological displays. The received less training on the traditional displays because they were similar to those in use in their regular workplace. The remainder of the training focused on differences between the simulation and the home plant of the operators. Experimental sessions were held in the
following two days of the week. To prevent fatigue, each crew was limited to three scenarios per day.

The experiment was a 3x2x2 within subject design. There were three treatment levels of interface (traditional, ecological and advanced), two treatment levels of scenario type (within design basis and beyond design basis), and two treatment levels of scenario phase (detection and mitigation). The treatments were completely crossed and presented in a balanced Latin square design to counter balance order effects of presentation of the experimental conditions (Skraaning et al., 2007).

Each scenario was characterized by two phases, a time period of detection before the first alarm sounded, and a time period of mitigation activities starting with the sounding of the first alarm (shown in Figure 6). Between the detection and mitigation periods a freeze of 5-10 minutes occurred during which the Process Overview queries were administered. These queries were also administered at the end of the scenario/mitigation phase. Throughout the scenario overall performance was observed and scenario understanding queries could occur. The simulator was frozen for a few seconds during the scenario understanding queries to allow the operators to answer without undue time pressure.

************** Insert Figure 6 here ***********/

RESULTS

We briefly discuss overall operator performance and present detailed results for our two situation awareness measures - process overview and scenario understanding. Full details on these results are available in Skraaning et al. (2007).
Overall Operator Performance

Creating a highly representative experimental environment for complex industrial systems is a methodological challenge. The operators in this study were working with unfamiliar displays with a relatively short training period. The choice to use realistic scenarios increases the variability of the performance. Thus, it was important to assess whether the operators were performing at an acceptable level throughout all experimental conditions, thereby ensuring that the situation awareness data was meaningful for statistical analysis.

A process expert observed all teams and assessed their overall performance both qualitatively and quantitatively. Our process expert had an engineering background and had worked as a turbine operator for 4 years, a reactor operator for 4 years, and a shift supervisor for 4 years. He also had been responsible for developing simulations and training operators before working as a process expert in HAMMLAB for 9 years.

According to the process expert, the overall performance level was “close to what one would expect from licensed NPP operators working on a familiar plant with unfamiliar computerized interfaces and a reduced crew size. There was no shift supervisor to manage the work and this may have hindered the operators’ decision making compared to normal operation at their home plant. It is also likely that the short training on the computerized displays and the novelty of the beyond design basis scenarios influenced performance negatively. However, there were no situations where the operators completely lost control and/or the simulation became meaningless.”

The quantitative performance results corroborate this observation (Lau, Jamieson, Skraaning, & Burns, in review). Performance was scored using the Operator Performance Assessment System (Skraaning, 1998; 2003). The scores obtained throughout the scenarios
indicated that the operators acted with some degree of conformance to canonical solutions established a priori. From the combined qualitative and quantitative assessments we felt confident that operator performance, while influenced by the experimental manipulations, was reasonable and that we could pursue further assessment of their situation awareness.

**Correlations between Process Overview and Scenario Understanding**

The process overview and scenario understanding measures were meant to be complementary indicators of situation awareness, therefore we examined the correlations between them. The correlation between the two measures on the raw scores was not significant ($r(72) = -0.03, p=0.78$). Because this correlation coefficient could be confounded by variance due to both crews and scenarios, we computed the partial correlations using z-score standardization to isolate these effects. The correlation coefficients based on measurements standardized by crew ($r(72) = -0.0051, p=.966$) and by scenario ($r(72) = -0.0633, p=.597$) were not significant, confirming the complementary nature of the measures. (Note that available empirical evidence has yet to establish the degree of correlation between measures for the three levels of SA (c.f., Lang, Roth, Bladh, and Hine, 2002; O’Brien & O’Hare, 2007). The lack of correlation is also suggested by the opposite direction for the main effects of phase between process overview (see Table 4) and scenario understanding (see Table 5) as illustrated by the interaction plots (Figure 7 and 10 below). This is discussed in more detail below.

**Process overview**

The normality assumption of ANOVA for the process overview measure is not satisfied according to the Shapiro-Wilks’ W tests on the distributions for every level of each manipulated variable. Given the limited access to participants for representative studies, this is
to be expected. However, ANOVA is generally robust against the violation of normality except when population distributions are heterogeneous (Kirk, 1995, pp. 99) – for example, distributions skewed in opposite directions. We thus examined the distributions for every treatment level using histograms and normal probability plots and did not find any major threats to the validity of the ANOVA results. The multivariate approach to repeated measures was used to test the hypothesized effects on process overview. Table 4 shows the outcome of this test for all effects.

Effects were observed for scenario phase (p<0.03), the interaction between interface type and scenario type (p<0.01), and the three way interaction between interface type, scenario type and scenario phase (p<0.03). Operators demonstrated better process overview in the detection phase of the scenario than in the mitigation phase (Figure 7). This effect explains 11% of the total variance (ω²=.11).

The interaction between interface type, scenario type and scenario phase on process overview is presented in Figure 7. The effect explains 19% of the total variance (ω²=.19).

Figure 7 shows that the ecological interface displays enhance process overview in the detection phase of beyond design basis scenarios, but impairs the overview in the detection phase of within design basis scenarios. Advanced displays appear to improve the process
overview in the mitigation phase of beyond design basis scenarios, while traditional and ecological displays seem to facilitate process overview at similar levels in the mitigation phase of within design basis scenarios.

Post hoc tests for differences between the means on this and the scenario understanding measure are not reported for two reasons. First, post-hoc procedures marginalize the power necessary to detect important findings (i.e., type II error). This is of particular concern for a representative study in the nuclear industry, where increasing the recruitment of operators licensed on a particular plant design is impractical. Second, post-hoc test procedures lack protection against the sphericity assumption for within-subject designs (Howell, 2006). Thus, rather than rely on statistical tests, we turn to alternative sources of confirmation of the omnibus effect. These alternative sources are discussed below.

**Scenario understanding**

The normality assumption for scenario understanding is not satisfied according to the Shapiro-Wilks’ W tests on distributions for every treatment level, but we did not find any threats to the validity of the ANOVA results (i.e., heterogeneity of forms between population distributions; see above). The multivariate approach to repeated measures was used to test the hypothesized effects on scenario understanding. Table 5 shows the outcome of this test for all effects.

*************** Insert Table 5 here ***************

Significant effects were found for scenario phase (p<0.05), the interaction between interface type and scenario type (p<0.05), and the three way interaction between interface type,
scenario type and scenario phase (p<0.01). Operators demonstrate better scenario understanding in the “mitigation phase” of the scenario (Figure 9). The effect explains 21% of the total variance ($\omega^2=.21$).

The interaction between interface type, scenario type and scenario phase on scenario understanding is presented in Figure 9. This effect explains 46% of the total variance ($\omega^2=.46$).

*************** Insert Figure 9 here ***************

Figure 9 shows that ecological interface displays enhance Scenario understanding in the detection phase of beyond design basis scenarios. Traditional displays seem to improve Scenario understanding in the mitigation phase of within design basis scenarios. These tendencies are consistent with the effects on process overview (even though scenario understanding and process overview are uncorrelated dependent variables). Once again, post hoc tests for differences between the means are not reported for reasons discussed above.

DISCUSSION

Display

We hypothesized that the ecological displays would improve situation awareness over the advanced, and further over the traditional, interfaces. We anticipated this effect to be most substantial when working with the unanticipated beyond design basis scenarios. Ecological displays are designed on the premise of showing the relationships in the environment and are built on a model that emerged from studies of troubleshooters. Ecological displays have repeatedly been found to support operators in scenarios where reasoning from principles is necessary to diagnose or control a problem. We found that the ecological displays consistently
showed better performance on both situation awareness measures during the detection phase in the beyond design basis scenarios of the experiment. The convergent indications of the two SA measures and previous empirical results support the conclusion that ecological displays could perform better than other displays in the detection phase of beyond design basis scenarios.

On within design basis scenarios, or during anticipated situations, the traditional displays were generally superior in both detection and mitigation phases. In these situations operators can respond to problems according to procedures and do not need to reason in the same way as in the unanticipated scenarios. Instead, operators can match symptoms of the scenario to their procedures and then execute and monitor the procedure. These two situations pose very different cognitive challenges, and may draw on different aspects of SA. For example, high pressure in a pressurizer may map directly to a procedure named *Pressurizer Level High*. The operator must be aware of the pressure in the pressurizer and understand that the situation is a problem. Once the operator finds the procedure, they must be aware of whether the situation matches the conditions of the procedure, then following the procedure they observe certain variables, take actions and observe the results of those actions. In contrast, if a high pressurizer level were an unanticipated situation, an operator must reason as to why the level is high, how the pressurizer works, and how it could be influenced by other variables. In the unanticipated situation, the operator must be aware of the fundamental principles behind the system or machine they are operating. In the anticipated situation, the operator must be aware of the critical cues that constitute a reliable mapping between the current situation and the situation already anticipated, planned for, or experienced. The different content of awareness mentioned here is independent of levels of SA in the sense that all three levels of SA occur in each situation. Rather this suggests there may be important
differences in SA content much along the lines of Rasmussen’s Skill Rules and Knowledge model (Rasmussen, 1983). Following this argument, that the anticipated situation is distinct from the unanticipated situation, those interface features that promote cue awareness and mapping of cues to known situations become more relevant. A familiar interface, such as the traditional interface in this study, could facilitate this mapping. Similarly, the ecological interface, with its emphasis on showing higher level principles may actually detract from an operator’s ability to perceiving and mapping the cues to a known situation effectively. There are two possible explanations to consider. First, EID may be effective at supporting SA in knowledge-based situations, but less effective at rule and skill based situations and may need further development in this area. Alternatively, our methods of how we describe situations, train operators, and write procedures may need revision to better support functional displays so that their mapping is more consistent.

The advanced displays generally supported SA in a manner similar to the traditional displays with one notable exception. The advanced displays demonstrated an advantage in supporting process overview or level 1 SA during the mitigation phase of beyond design basis scenarios. These displays may have some graphical features that facilitate better awareness of the trends of key process parameters. We speculate that the mini and pop-up trends strategically inserted by process experts into the advanced displays provide this unique support for process overview without substantially improving scenario understanding. As well, this display condition had an overview display that contained a text-based indicator of turbine operation mode and a mini trend which could also have contributed to the observed advantage. Future empirical studies are necessary to identify the specific features that induced this performance difference.
Scenario Phase

Scenario phase interacted with our SA measures. Process overview decreased while scenario understanding increased dramatically from the detection phase to the mitigation phase. As the scenario changes phase, operators become more concerned with problem solving and monitoring becomes more problem-oriented. This may be the reason that process overview scores tend to decrease. As the problem develops and operators discover more relevant information, operators could understand the situation better and project future developments more accurately.

One interesting effect however occurred with the ecological interface. In the detection phase of the beyond design basis scenarios, the ecological interface improved scenario understanding so dramatically that it was essentially the same as the level of scenario understanding in the mitigation phase. In being able to see the functional aspects of the ecological interface, operators may have been able to interpret process dynamics due to unanticipated events even in the detection phase (i.e. prior to the first alarm).

Limitations

Four major limitations of the study merit consideration. First, the ecological display implementation was, in effect, a hybrid ecological-traditional interface. The hybrid nature is evident in: a) the limited scope of the ecological displays, b) the use of traditional interactivity in the ecological display condition, and c) the use of the traditional overview display in the ecological display condition. As indicated, the scope of the ecological displays was limited to the secondary side of the plant. Given that the full plant must be in operation during the experiment, the operations team was using two different types of interfaces in controlling the two sides of the plant. The hybrid nature of the implementation is further evident in that we did
not develop a new interaction scheme based on the EID framework to suit the ecological displays. Instead, we retained the same controls and interaction style as in the traditional displays. We also did not design a specific ecological overview display. Because the overview display is shared by both operators and we had not redesigned the displays on the reactor side, we retained the traditional overview display for the ecological displays. These three limitations in implementation of the ecological displays may have introduced compatibility issues which could make intervention in the ecological condition more challenging.

Although a hybrid implementation is not ideal from an empirical perspective, it is representative of industry practice. In our experience, industry tends to adopt advanced interface design techniques in phases, typically testing new concepts in limited operations while retaining the full suite of traditional controls over the process. A comparison of displays with a full ecological implementation that included ecological design of the reactor side, the overviews, and the interaction schemes would provide a more accurate assessment of the benefits or challenges of the ecological approach.

A second major limitation pertains to the artificial manipulations of the operating environment to allow for the presentation of SA queries. In particular, freezing a simulation of a continuous process is clearly not a natural occurrence. There is a risk that interrupting the operators’ workflow could alter SA by itself (Sarter & Woods, 1995). This reactivity can go in either direction, as more time for reflection would improve the operators’ understanding, while intruding and disturbing may also degrade the operators’ comprehension of the system state. On the other hand, operators do regularly provide status updates to other plant personnel in real control room environments. The impacts of intrusion on SA and reflection on performance can therefore be seen as an integrated and natural part of the operators’ every day work. Similarly,
the number of status updates per time unit may have been higher than in real life, but this issue addresses a general concern with respect to the realism of simulator studies, since simulated scenarios are often compressed representations of possible real events. Issues with using SA measures in realistic situations have been discussed by other authors (e.g. Li et al. (2007)).

A third limitation stems from the very limited training operators received in using the ecological interface. The traditional displays were very similar to those used in the current control room, so the operators would be expected to be highly familiar with such displays. Similarly, the advanced displays had the same mimic-diagrams, closer in appearance and features to the traditional displays. Thus, to observe superior situation awareness with the ecological displays in any condition compared to the traditional displays is surprising. It is therefore likely that that the benefits for ecological displays shown in this experiment are conservative estimates of the full effect of EID in this domain.

The final limitation of this study may involve the model of situation awareness that we have used. Some researchers have cautioned against blindly transporting notions of SA from one domain to another (see e.g., Pew, 1995; 2000). As well, there is a growing literature calling for a more complete development of the theoretical foundation of SA (see e.g., Rousseau, Tremblay and Breton, 2004). For the purpose of understanding how EID can influence operator coupling with the environment (cf. Flach, 1995) we have adopted the prevailing model of situation awareness as a framework to form our hypotheses, describe our measures, and communicate our findings.

**CONCLUSION**

This experiment provides evidence that EID can support SA for process control tasks. In particular, EID can improve SA during the monitoring of unanticipated events. For complete
support of both anticipated and unanticipated events however, the greatest benefits may obtain from integrating EID with traditional or user-centered approaches.

There are two clear areas for improvement in EID. EID needs to support SA more effectively in anticipated situations and during intervention. Both of these needs may be achieved by expanding the EID approach by introducing task analysis or task based displays to the ecological display suite. This supports past research by Jamieson, Ho, Miller and Vicente (2007) and recommendations in Burns and Hajdukiewicz (2004).

Similarly, traditional and other display approaches can benefit from adding EID to their design process. This study provides evidence that EID does improve SA during the monitoring of unanticipated events. Since unanticipated events are historically the most damaging accidents, improving SA during the monitoring period is very important. In conclusion there is reasonable evidence that EID does support SA and should be considered as a possible design approach to improve SA.

ACKNOWLEDGEMENTS

We would like to acknowledge an NSERC Special Research Opportunity grant to Burns and Jamieson and internal support from the OECD Halden Reactor Project. We are thankful to Christer Nihlwing for his support throughout the experiment and numerous people on the Halden Project, particularly Jon Kvalem, who made this research possible. Finally, we would like to thank the operators who took part in this study.
REFERENCES


Tables

Table 1. Graphical features specific to the advanced displays (absent in the traditional displays).

Table 2. Questions to assess Level 1 SA.

Table 3. Level 2 and Level 3 SA queries.

Table 4. Effects of variables on Process Overview.

Table 5. Effects of variables on Scenario Understanding.
Table 1. Graphical features specific to the advanced displays (absent in the traditional displays).

<table>
<thead>
<tr>
<th>Graphical Features</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mini trends</td>
<td>Small trend diagrams for one or multiple key system parameters that are embedded in the mimic diagrams</td>
</tr>
<tr>
<td>2. Pop-up trends</td>
<td>Pop-up windows containing trend information for one or multiple selected system parameters</td>
</tr>
<tr>
<td>3. Pump icons</td>
<td>Pump icons showing status and nominal flow</td>
</tr>
<tr>
<td>4. Configural displays</td>
<td>Graphics containing information from at least two related sensors (e.g., core coolant flow vs. reactor effect)</td>
</tr>
<tr>
<td>5. Mini dials</td>
<td>Small dials for quick identification of deviating parameters</td>
</tr>
<tr>
<td>6. Pump bars</td>
<td>Bar graphs representing the speeds of eight internal circulation pumps</td>
</tr>
<tr>
<td>7. Others</td>
<td>Several minor enhancements (e.g., pie charts for condenser pressure)</td>
</tr>
</tbody>
</table>
Table 2. Questions to assess Level 1 SA.

1. Recently, the position of [reference to equipment] has:
   decreased - remained stable - increased
2. Recently, the number of [reference to equipment] in operation has:
   decreased - remained stable - increased
3. Recently, the electrical output power of [reference to equipment] has:
   decreased - remained stable - increased
4. Recently, the flow from [reference to equipment] has:
   decreased - remained stable - increased
Table 3. Level 2 and Level 3 SA queries.

Level 2 SA Queries
- Why is [reference to process state]?
- What is happening with [reference to process state]?
- Has [reference to process state] occurred?
- Have you taken care of [reference to problem in the process]?
- Have you come any further with respect to [reference to problem in the process]?
- Has the process changed since my last call?
- What has changed in the process since my last call?
- What do you think is wrong in this situation?
- Is everything normal?
- How is everything?

Level 3 SA Queries
- What are you planning to do about [reference to process state]?
- Which strategy are you following with respect to [reference to process state]?
- Where are you in the work process with respect to [reference to process state]?
- Did the [reference to process event(s)] evolve as expected?
- Did the [reference to execution of task(s)] proceed as expected?
- What are the consequences of [reference to process state]?
- What will happen in the process now?
- Is there something I should know about this situation?
- Is everything going according to plan?
Table 4. Effects of variables on Process Overview.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Test</th>
<th>Value</th>
<th>F</th>
<th>Effect df</th>
<th>Error df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface type</td>
<td>Wilks</td>
<td>0.95</td>
<td>0.10</td>
<td>2</td>
<td>4</td>
<td>0.91</td>
</tr>
<tr>
<td>Scenario type</td>
<td>Wilks</td>
<td>0.94</td>
<td>0.33</td>
<td>1</td>
<td>5</td>
<td>0.59</td>
</tr>
<tr>
<td>Scenario phase</td>
<td>Wilks</td>
<td>0.33</td>
<td>10.01</td>
<td>1</td>
<td>5</td>
<td>0.03*</td>
</tr>
<tr>
<td>Interface type*Scenario type</td>
<td>Wilks</td>
<td>0.09</td>
<td>20.32</td>
<td>2</td>
<td>4</td>
<td>0.01*</td>
</tr>
<tr>
<td>Interface type*Scenario phase</td>
<td>Wilks</td>
<td>0.79</td>
<td>0.53</td>
<td>2</td>
<td>4</td>
<td>0.62</td>
</tr>
<tr>
<td>Scenario type*Scenario phase</td>
<td>Wilks</td>
<td>0.61</td>
<td>3.22</td>
<td>1</td>
<td>5</td>
<td>0.13</td>
</tr>
<tr>
<td>Interface type<em>Scenario type</em>Scenario phase</td>
<td>Wilks</td>
<td>0.18</td>
<td>9.41</td>
<td>2</td>
<td>4</td>
<td>0.03*</td>
</tr>
</tbody>
</table>
Table 5. Effects of variables on Scenario Understanding.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Test</th>
<th>Value</th>
<th>F</th>
<th>Effect df</th>
<th>Error df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface type</td>
<td>Wilks</td>
<td>0.83</td>
<td>0.40</td>
<td>2</td>
<td>4</td>
<td>0.70</td>
</tr>
<tr>
<td>Scenario type</td>
<td>Wilks</td>
<td>0.99</td>
<td>0.03</td>
<td>1</td>
<td>5</td>
<td>0.87</td>
</tr>
<tr>
<td>Scenario phase</td>
<td>Wilks</td>
<td>0.20</td>
<td>20.38</td>
<td>1</td>
<td>5</td>
<td>0.01*</td>
</tr>
<tr>
<td>Interface type*Scenario type</td>
<td>Wilks</td>
<td>0.17</td>
<td>9.98</td>
<td>2</td>
<td>4</td>
<td>0.03*</td>
</tr>
<tr>
<td>Interface type*Scenario phase</td>
<td>Wilks</td>
<td>0.42</td>
<td>2.81</td>
<td>2</td>
<td>4</td>
<td>0.17</td>
</tr>
<tr>
<td>Scenario type*Scenario phase</td>
<td>Wilks</td>
<td>0.94</td>
<td>0.32</td>
<td>1</td>
<td>5</td>
<td>0.60</td>
</tr>
<tr>
<td>Interface type<em>Scenario type</em>Scenario phase</td>
<td>Wilks</td>
<td>0.06</td>
<td>31.07</td>
<td>2</td>
<td>4</td>
<td>0.00*</td>
</tr>
</tbody>
</table>
Figures.

Figure 1. A traditional display for the steam system.

Figure 2. The ecological display for the turbine system from Kwok, 2007.

Figure 3. The operator workstation used in the study.

Figure 4. Graphical enhancements: (1) Mini-trends, (2) Pop-up trend, (3) Pump-icons, (4) Configural display, (5) Mini-dials, and (6) Pump-bars.

Figure 5. An example of an advanced display.

Figure 6. Timeline of a scenario with freeze points.

Figure 7. The interaction between interface type and scenario type, in different scenario phases.

Figure 8. The interaction between interface type and scenario type by different scenario phases.
Biographies

Catherine Burns, University of Waterloo, PhD. Human Factors Engineering, University of Toronto, 1998


Greg Jamieson, University of Toronto, PhD., Human Factors Engineering, University of Toronto, 2002

Nathan Lau, University of Toronto, BASc., Human Factors Engineering, University of Toronto, 2004.

Jordanna Kwok, University of Waterloo, MASc., Systems Design Engineering, University of Waterloo, 2007


Gisle Andresen, OECD Halden Reactor Project, cand.psychol., Psychology, University of Oslo, 1999