

Proceedings of the Human Factors and Ergonomics Society Annual Meeting

<http://pro.sagepub.com/>

Ecological Interface Design and Fault Management Performance: Long-Term Effects

Klaus Christoffersen, Christopher N. Hunter and Kim J. Vicente

Proceedings of the Human Factors and Ergonomics Society Annual Meeting 1995 39: 496

DOI: 10.1177/154193129503900912

The online version of this article can be found at:

<http://pro.sagepub.com/content/39/9/496>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Human Factors and Ergonomics Society](http://www.hfes.org)

Additional services and information for *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* can be found at:

Email Alerts: <http://pro.sagepub.com/cgi/alerts>

Subscriptions: <http://pro.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://pro.sagepub.com/content/39/9/496.refs.html>

>> [Version of Record](#) - Oct 1, 1995

[What is This?](#)

ECOLOGICAL INTERFACE DESIGN AND FAULT MANAGEMENT PERFORMANCE: LONG-TERM EFFECTS

Klaus Christoffersen, Christopher N. Hunter, and Kim J. Vicente
Cognitive Engineering Laboratory
Department of Industrial Engineering
University of Toronto

This paper presents a six-month longitudinal study of the effects of ecological interface design (EID) on fault management performance. The research was conducted in the context of DURESS II, a real-time, interactive thermal-hydraulic process control simulation that was designed to be representative of industrial systems. Subjects' performance on two interfaces was compared, one based on the principles of EID and another based on a more traditional piping and instrumentation diagram (P&ID) format. Subjects were required to perform several control tasks, including startup, tuning, shutdown, and fault management on both routine and non-routine faults. At the end of the experiment, subjects used the interface that the other group had been using to control the system. The results indicate that there are substantial individual differences in performance, but that overall, the EID interface led to faster fault detection, more accurate fault diagnosis, and faster fault compensation.

INTRODUCTION

This paper describes an experiment investigating the long-term effects of ecological interface design (EID) on fault management performance in a process control microworld. EID is a novel theoretical framework for designing interfaces for complex work environments. It is based on the skills, rules, knowledge taxonomy of levels of cognitive control (Rasmussen, 1986). The framework includes three prescriptive design principles, each directed at providing the appropriate interface support for a specific level of cognitive control (Vicente and Rasmussen, 1990, 1992). First, to support skill-based behavior (SBB), operators should be able to act directly on the interface. Furthermore, the structure of the displayed information should be isomorphic to the part-whole structure of movements. Second, to support rule-based behavior (RBB), the interface should maintain a consistent one-to-one mapping between the work domain constraints and the perceptual cues provided in the interface. Third, to support knowledge-based behavior (KBB), the interface should represent the work domain in the form of an abstraction hierarchy (Rasmussen, 1986), which can serve as an externalized mental model to support problem solving. This system model contains both physical and functional representations of the system.

Although previous empirical investigations of EID have led to an encouraging set of findings, the

long-term effect of EID on operator performance had yet to be established (Christoffersen, Hunter, and Vicente, 1994). In an attempt to address this problem, a longitudinal study was conducted over a period of an unprecedented 6 months. Subjects controlled a thermal-hydraulic process simulation or microworld, DURESS II, each weekday for approximately an hour over this period. The performance of two interfaces was compared, one based on the principles of EID and another based on a more traditional piping and instrumentation diagram (P&ID) format (see Christoffersen et al., 1994 for a description of the interfaces). This paper describes the results from the fault management performance trials in the experiment. The primary questions to be addressed are:

- 1) Is there an interface effect for routine faults?
- 2) Is there an interface effect for unfamiliar, non-routine faults after extensive practice?

METHOD

Experimental Design

A repeated measures, between-subjects design, with the type of interface (P&ID or EID) as the primary manipulation, was adopted for this experiment. Subjects were assigned to one of the two interfaces and participated for a total of six months.

Subjects

Six male subjects, ranging in age from 23 to 32 years of age, participated in this study. The subjects, with one exception, had science or engineering backgrounds. An attempt was made to assign subjects to interface groups so as to match for background. Subjects were paid for participating in the experiment.

Experimental Tasks

During the experiment, subjects were asked to perform four different types of control tasks:

- 1) **Startup.** For this task, the subject was presented with a shut down system and was asked to bring the system to steady state, meeting pre-defined setpoints consisting of temperature and demand goals.
- 2) **Tuning to New Setpoints.** In this task, the subject needed to bring the system from an on-line, steady state initial condition to a pair of new steady state demand setpoints.
- 3) **Shutdown.** During this task, the subject was required to bring the system from an on-line, steady state condition to a shut down condition.
- 4) **Fault Management.** After the introductory phase of the experiment, when subjects had had a reasonable amount of practice at controlling the system, they were occasionally presented with trials during which a fault would occur. Subjects were told their task was to detect, diagnose, and compensate for any such faults.

For all control tasks, steady state was defined as maintaining the system in the goal region for 5 consecutive minutes.

Trial Types

Trials normally consisted of the first three control tasks described above, performed either in isolation or in sequence within the same trial. For fault management trials, faults were chosen from a number of possible faults available in the simulation program, all of which are believed to be representative of faults that could occur in an actual process control plant. There were two general classes of faults: routine faults and non-routine faults (see below). Each fault occurred at some fixed time after the beginning of a trial.

Routine faults were designed to be relatively simple in nature, and could be easily compensated for once diagnosed. These were intended to be analogous to recurring failures in any industrial system, where some of the system components are inherently less reliable than others. Three faults of this type were used in this experiment. Subjects

experienced each of these routine fault types three times during the experiment. Note that subjects were not told what faults might occur during the experiment, when they would occur, or how frequently they would occur.

Non-routine faults were designed to be more complex in nature. They always consisted of two independent but interacting faults. The non-routine faults were intended to be analogous to rare, unanticipated occurrences within a system which, although they can be compensated for, are more difficult to diagnose and in some cases more difficult to compensate for. There were three non-routine faults, all included late in the experiment.

Procedure

There were four distinct phases: an introductory session where the experimental protocol was explained to subjects, an introductory practice phase during which the complexity of tasks was gradually increased, an extended practice phase characterized by repeated exposure to normal trials and occasional exposure to routine faults, and a final phase examining the long-term effects of each interface on operators' knowledge. This latter phase included the three non-routine fault trials, and six transfer trials as well. The entire experiment consisted of 224 trials per subject. Six days, referred to as test days, were also set aside during the experiment to have the subjects perform a set of knowledge elicitation tests. The experiment concluded with a debriefing session, attended by all subjects with the exception of one who had not yet completed the experiment.

Performance Measures

There were three primary sources of data in the experiment: time-stamped data logs, verbal protocols, and knowledge elicitation techniques. A detailed description of each of these measures can be found in Christoffersen et al. (1994).

RESULTS

Only the results from the fault management trials will be presented in this paper. Analysis of the fault trials was based on a process tracing method. Verbal protocols were transcribed and mapped onto 3 decision making categories: detection, diagnosis, and compensation. Individual data from each trial were represented in the form of a timeline-based data table, which allowed for a relatively structured and meaningful interpretation of the trials. A total of 54 trials were analyzed in total (6 routine faults and 3 non-routine faults for each of the 6 subjects). Data

tables for all of these trials can be found in Christoffersen et al. (1994). Only a brief summary of the results can be provided here.

Is there an interface effect for routine faults?

Fault compensation performance was evaluated by examining detection, diagnosis, and compensation. Statistical tests were not conducted because of the large proportion of missing data points and the small sample size. Therefore, only descriptive statistics will be presented.

Fault detection performance was evaluated by measuring the time between the beginning of a fault and the time at which subjects first verbalized that the system was not operating as it should. Table 1 summarizes these data across the 6 routine fault trials. There were 3 trials where subjects did not even detect a fault (indicated by a "-"), 2 where the subjects did not experience the fault because they were not using the component that failed (indicated by "n/a"), and 1 where data are not available (indicated by "?").

Fault diagnosis performance was evaluated by an arbitrary scoring scheme whereby subjects would receive 0 to 3 points for each trial. Zero points indicated that the subject did not detect a fault or that they did not verbalize symptoms that were relevant to the fault, while 3 points were awarded if subjects correctly diagnosed the root cause of the fault (e.g., reservoir 1 is leaking). One and two points were awarded for accurate but less specific diagnoses (e.g., describing the surface symptoms of a fault, or the deeper symptoms of a fault but not the root cause). Table 2 summarizes these data across the 6 routine fault trials. Again, the "-" symbol represents trials where faults were not even detected, and "n/a" represents trials where subjects were not using the faulty component.

Fault compensation performance is difficult to evaluate since it is embedded in fault detection and fault diagnosis. The only meaningful measure that was adopted was the time to successfully complete a trial (i.e., achieve steady state in the presence of a fault). Table 3 summarizes these data across the 6 routine fault trials. In this case, the "-" symbol indicates trials that ended prematurely because subjects damaged a system component via their actions (these trials are referred to as "blowups").

These tables clearly show a great deal of variability in performance, both within and across individuals. Furthermore, neither interface group was consistently superior to the other across all measures. Despite this variability, there were trends indicative of differences between interface groups. More specifically, the EID group tended to have slightly faster detection and compensation times than

the P&ID group. Furthermore, the EID group exhibited much better diagnosis scores than the P&ID group.

The process tracing analyses indicate that these differences in performance were tied to differences in fault management strategies between the two groups. The P&ID group had difficulty differentiating between the onset of a fault and a normal situation requiring minor adjustment of the components. Thus, they frequently had to resort to a trial and error strategy. For example, after a particular fault trial, subject TL, the most proficient in the P&ID group, commented: "We're back to empirical science; just seeing what works". On other occasions, the P&ID subjects detected faults by using a violated rules strategy. Again, the words of TL after another fault trial clearly describe this strategy: "Usually, in my mind I know that for a certain input water level, the heater level should be around a certain place. You make up these rules of thumb. [They were] pretty good until today!" The problem with each of these strategies is that they do not give very much insight as to what the nature of the fault is, or what to do about it. Thus, the P&ID subjects usually did not have a very accurate understanding of the root cause of the abnormality. And because they could rarely diagnose the cause of the problem, they frequently had to resort to compensating in a trial and error fashion. Sometimes, this could be done efficiently, but in other cases, it took considerable time.

The EID group tended to exhibit a different type of fault management strategy. They were more likely to detect or diagnose a failure simply by observing the feedback that was provided to them by the interface. The important benefit of detecting a fault in this manner is that the diagnostic feedback also informs and aids diagnosis and compensation. In contrast to the strategies used most frequently by the P&ID subjects, detecting faults via diagnostic feedback gives important information as to what the nature of the fault is, not just that there is a fault. Moreover, having a better idea of what the nature of the problem is can also help subjects to compensate. Some subjects had problems making this leap, but in many cases (certainly many more than in the P&ID group), EID subjects were able to leverage their understanding of the fault to develop an effective compensation plan. Furthermore, with the possible exception of subject AS, it was comparatively rare to see an EID subject adopt a trial and error compensation strategy.

In summary, the differences in strategies identified by the process tracing analyses suggest that the advantage of the EID interface results from providing subjects with more direct, diagnostic feedback than the P&ID interface. The findings are also consistent with those of Pawlak and Vicente

TABLE 1. Detection times for routine faults (VB - Valve Block, RL - Reservoir Leak, HF - Heater Failure).

		TRIAL #						
		64	94	97	165	183	187	Mean
		VB	RL	HF	RL	HF	VB	
EID	IS	11	11	79	39	25	26	31.8
	AS	3	89	230	94	100	9	87.5
	AV	2	26	119	60	61	n/a	53.6
	Mean	5.3	42.0	142.7	64.3	62.0	17.5	57.9
P&ID	TL	32	105	59	15	359	n/a	114.0
	WL	?	36	115	-	-	-	75.5
	ML	54	13	59	78	46	36	47.7
	Mean	43.0	51.3	77.7	46.5	202.5	36.0	77.5

TABLE 2. Diagnosis scores for routine faults.

		TRIAL #						
		64	94	97	165	183	187	Total
		VB	RL	HF	RL	HF	VB	
EID	IS	3	1	3	1	3	3	14
	AS	1	3	3	2	3	1	13
	AV	2	3	3	2	3	n/a	13
	Total	6	7	9	5	9	4	40
P&ID	TL	1	1	3	2	1	n/a	8
	WL	1	1	1	-	-	-	3
	ML	3	1	1	1	3	1	10
	Total	5	3	5	3	4	1	21

TABLE 3. Compensation times for routine faults.

		TRIAL #						
		64	94	97	165	183	187	Mean
		VB	RL	HF	RL	HF	VB	
EID	IS	614	577	480	752	700	879	667.0
	AS	-	839	480	1356	798	-	868.3
	AV	-	-	480	-	660	n/a	570.0
	Mean	614.0	708.0	480.0	1054.0	719.3	879.0	717.9
P&ID	TL	-	480	544	654	1068	n/a	686.5
	WL	929	611	480	720	699	-	687.8
	ML	1014	-	497	1232	790	-	883.3
	Mean	971.5	545.5	507.0	868.7	852.3	-	747.5

(1994). These differences in strategies are consistent with the predictions derived from the theoretical motivations behind the EID framework.

Is there an interface effect for unfamiliar, non-routine faults after extensive practice?

The findings obtained for unfamiliar, non-routine faults are very similar to those obtained for the routine faults, and therefore, will only be briefly summarized. Not only was the EID group able to detect faults more quickly than the P&ID group, but they exhibited faster compensation times, more accurate diagnoses, and a slightly lower number of blowups as well. It was also noted that incorrect or incomplete diagnoses frequently led to either long compensation times or blowups for both groups. These performance differences seemed to be tied to the same type of strategy differences observed on routine faults.

CONCLUSIONS

The results of this experiment, when combined with the findings of the previous empirical studies of EID, lend further weight to the value of the EID principles. Adopting the abstraction hierarchy as a work domain representation, and embedding that representation in a form that supports perception and action can lead to predictable performance benefits. The results presented here are particularly noteworthy since they were obtained from an experiment conducted over an unprecedented 6 month period, thereby increasing the potential for generalizing the results to experienced operators.

Although many critical issues still remain to be addressed, the practical relevance and potential benefits of EID have not gone unnoticed by industry. For example, both Honeywell and AECL Research have incorporated portions of the EID interface for DURESS into prototypes that are intended to represent the state-of-the-art in advanced interface design. More importantly, Toshiba in Japan has adopted EID as the basis for designing its advanced control room for a next generation boiling water reactor plant (Monta et al., 1991). It has also incorporated and adapted specific features of the EID interface for DURESS II into some of their displays. This application is notable since it has been

conducted at the scale of a full-scope nuclear power plant simulator. More recently, Mitsubishi Atomic Power Industries in Japan has also demonstrated a very strong interest in EID. It has contracted Battelle to initiate a 5 year research program, solely on EID. This success in technology transfer to industry, while limited, shows that EID is a promising candidate for designing interfaces for complex industrial systems.

ACKNOWLEDGEMENTS

This research was sponsored by a research contract with the Japan Atomic Energy Research Institute (Dr. Fumiya Tanabe, contract monitor), and by research and equipment grants from the Natural Science and Engineering Research Council of Canada.

REFERENCES

- Christoffersen, K., Hunter, C. N., and Vicente, K. J. (1994). Research on factors influencing human cognitive behaviour I (CEL 94-05). Toronto: Cognitive Engineering Laboratory, University of Toronto.
- Monta, K., Takizawa, Y., Hattori, Y., Hayashi, T., Sato, N., Itoh, J., Sakuma, A., & Yoshikawa, E. (1991). An intelligent man-machine system for BWR nuclear power plants. In Proceedings of AI 91: Frontiers in Innovative Computing for the Nuclear Industry (pp. 383-392). Jackson, WY: ANS.
- Pawlak, W. S., and K. J. Vicente, K. J. (1994). Inducing effective operator control through ecological interface design. Manuscript submitted for publication.
- Rasmussen, J. (1986). Information processing and human-machine interaction: An approach to cognitive engineering. New York: North-Holland.
- Vicente, K. J., and Rasmussen, J. (1990). The ecology of human-machine systems II: Mediating direct perception in complex work domains. Ecological Psychology, 2, 207-250.
- Vicente, K. J., and Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. IEEE Transactions on Systems, Man, and Cybernetics, SMC-22, 589-606.