

Proceedings of the Human Factors and Ergonomics Society Annual Meeting

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

Measuring the Impact of Ecological Interface Design on Operator Skill Acquisition

Dianne E. Howie and Kim J. Vicente

Proceedings of the Human Factors and Ergonomics Society Annual Meeting 1996 40: 972

DOI: 10.1177/154193129604001904

The online version of this article can be found at:

<http://pro.sagepub.com/content/40/19/972>

Published by:



<http://www.sagepublications.com>

On behalf of:



Human Factors and Ergonomics Society

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/40/19/972.refs.html>

>> [Version of Record](#) - Oct 1, 1996

[What is This?](#)

MEASURING THE IMPACT OF ECOLOGICAL INTERFACE DESIGN ON OPERATOR SKILL ACQUISITION

Dianne E. Howie
Centre for Applied Cognitive Science
Ontario Institute for Studies in Education
University of Toronto
Toronto, Ontario, Canada

Kim J. Vicente
Cognitive Engineering Laboratory
Department of Mechanical and Industrial Engineering
University of Toronto
Toronto, Ontario, Canada

This paper examines the effect of interface design on participants' adaptation with experience using a process control microworld. The experiment was conducted using DURESS II, an interactive thermal-hydraulic simulation with two interfaces: traditional (including only physical information) and ecological (including both physical and functional information). The experiment was 6 months in duration and included a variety of conditions. A group of early and late startup trials were analyzed in detail in order to investigate changes in expertise. Previously, trial completion time (or steady state time) has been used as the primary measure of performance on normal trials. In this paper, a series of additional measures were developed. These measures included: action transition graph complexity, path length in state space, deviation from the temperature goal line in mass versus energy inventory graphs, and proportion of early control actions on timelines. These performance measures revealed larger magnitude performance improvements in the ecological interface group.

INTRODUCTION

This paper introduces a series of performance measures used to analyze participants' adaptation with experience operating a process control simulation. The role of ecological interface design (EID) in shaping the participants' adaptation is examined. EID is a theoretical framework used to design interfaces for complex work environments (Vicente & Rasmussen, 1990, 1992). Rasmussen's skill-rule-knowledge (SRK) taxonomy and abstraction (or means-end) hierarchy provide the conceptual foundation of EID.

According to the SRK taxonomy, there are three levels of cognitive control: skill based, rule based, and knowledge based behavior (Rasmussen, 1983). Ecological interfaces should support all three levels. To support skill based behavior, the user should be allowed to act directly on the interface. To support rule based behavior, there should be a consistent one-to-one mapping between cues on the interface and constraints in the work environment. To support knowledge based behavior, the work domain should be represented as an abstraction hierarchy (Rasmussen, 1985). The abstraction hierarchy aids problem solving by explicitly representing goal-relevant constraints describing the structure of the system.

This study was conducted using the DURESS II simulation. DURESS II provides both an interface designed according to the principles of EID, and a traditional interface (refer to Figures A1 and A2).

There are two primary questions addressed by this paper. What measures, in addition to steady state time, may be used to assess expertise? What differences between interface groups do these measures reveal?

METHOD

Experimental Design

The analyses presented here employed a between-subjects design, with two factors: Interface and Block (early or late) with repeated measures for Trial. The early and late blocks of trials include the first 20 trials and last 20 non-fault trials over a six month period. Participants were assigned to one of two interfaces: traditional (P interface, including only physical information) or ecological (P+F interface, including physical and functional information). The interfaces are illustrated in Appendix A.

Participants

The participants were 6 males ranging in age from 23 to 32 years. Five of the six participants had a science or engineering background. The participants were matched as closely as possible in pairs according to their educational background. One member of each pair was assigned to each interface group.

Procedure

The experiment included a total of 224 trials per participant on the DURESS II simulation. The participants were gradually introduced to the experimental tasks during the first month. Initially, the participants learned to take the system from a shutdown state to steady state (startup task). Steady state was achieved when the participant met the temperature and demand goals for both reservoirs for 5 consecutive minutes. Shutdown and tuning tasks were

progressively added. In the shutdown task, participants were required to return the system to a shutdown state. In the tuning task, participants had to adjust the system to new steady state demands.

For the following five months, each trial consisted of startup, tuning, and shutdown tasks performed contiguously. Faults were distributed randomly, and infrequently, throughout the trials. Each trial had a different steady state demand pair to prevent the participants from adopting excessively simplified control methods. The participants were not informed of the results of each trial, although they could receive feedback by observing the elapsed time at any point in the trial, and any system failure messages displayed on the screen.

Whenever a participant made a control action, the component the participant adjusted, the time, and the state of all of the system variables were recorded. For each trial, these values were stored in a log file, making them available for further analysis.

RESULTS

The first 20 trials and the last 20 non-fault trials were examined in detail to investigate differences in the participants' performance with experience. Only the startup portions of the trials were analyzed. This portion of the data provides a consistent base for comparison of the participants' strategies. (For an analysis of tuning, shutdown, and fault data refer to Christoffersen, Hunter, and Vicente, 1994. For a more detailed discussion of performance measures and individual differences, refer to Howie, 1996).

Steady State Times

The mean steady state times decreased substantially between the first and last block of trials for each interface group, as shown in Table 1 ($F(1,4)=88.35, p<0.001$). Faster steady state times reflect the participants' increasing ability to reach the goals quickly, and to maintain the system in the goal state.

Table 1. Steady State Times for Startup.

	Steady state times (seconds)	
	Trials 1-22	Trials 196-217
P+F Group	687.6	395.1
P Group	602.2	393.3

There was also a significant effect for Trial ($F(19,75)=3.47, p<0.0001$), Block x Trial ($F(19,49)=5.69, p<0.0001$), and Group x Block x Trial ($F(17,49)=2.24, p<0.05$). The steady state time decreased from Trial 1 to Trial 20, and the rates of change of steady state time differed in each block of trials. The steady state times decreased rapidly in the first block of trials, and remained relatively constant in the last block of trials, as shown in Figure 1. Further, this rate of change depended on interface.

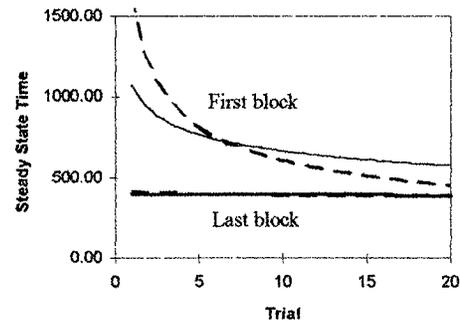


Figure 1. Average steady state times for each trial in the first and last block of trials (solid lines for P+F group and dotted lines for P group).

In previous studies, steady state times have been used as the primary measure of performance. However, as illustrated below, there are other possible measures of performance.

Action Transition Graphs

Action transition graphs reveal sequential relationships in behavior (Moray, Lootsteen, & Pajak, 1986). Each component that can be acted on is represented by a node, and those nodes that are accessed in sequence are joined by a line. The thickness of a line joining any two nodes is proportional to the frequency of that transition. Figure 2 shows the action transition graphs for one participant on his first completed trial and on the last normal trial.

The graphs for each participant generally became less complex with experience, indicating that the participants' control actions were more sequentially consistent. In order to quantify this change, complexity was operationalized as the number of lines in the action transition graph.

Table 2. Mean Action Transition Graph Complexity.

	Action Transition Graph Complexity	
	Trials 1-22	Trials 196-217
P+F Group	32.3	25.2
P Group	27.1	24.1

Table 2 gives the mean complexity for the initial and final blocks of trials. There was a marginally significant decrease in the action transition graph complexity between the first and last blocks of trials ($F(1,4)=5.11, p=0.09$). There was also a significant effect for Trial ($F(19,75)=2.56, p<0.005$), and a marginally significant effect for Block x Trial ($F(19,49)=1.72, p=0.07$). Thus, the complexity varied significantly across trials, and the rates of change of complexity differed in each block of trials, similar to the trends observed in Figure 1 for steady state time.

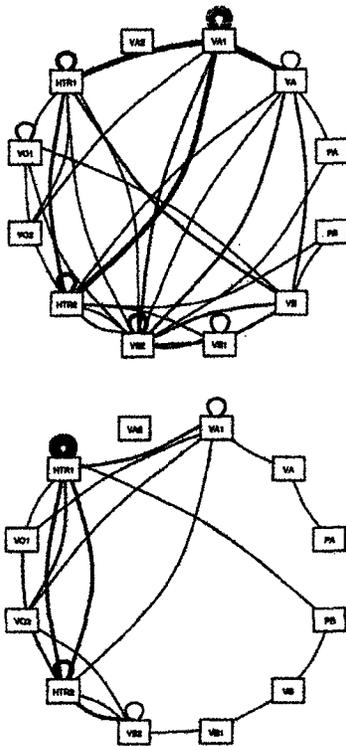


Figure 2. Comparison of action transition graphs for first (top) and last trial (bottom).

The correlations between action transition graph complexity and steady state time were examined. There were no significant correlations between for experienced participants in the P+F group ($r(13)=0.12$, $r(15)=0.07$, and $r(18)=-0.15$). However, there were significant positive correlations between complexity and steady state time for the P group ($r(18)=0.67$ and $r(17)=0.59$, and $r(15)=0.12$, *ns*). A simple action transition graph may indicate standardized operating procedures. Thus, these results may indicate that rigid procedural strategies were associated with successful performance for the P group but not for the P+F group. (This observation should be confirmed with a more direct measure of procedural variability.)

State Space Diagrams

State space diagrams portray the system state with respect to the goal state (Sanderson, Verhage, & Fuld, 1989). The system states for each reservoir in DURESS II were plotted on a graph of temperature versus demand. If both the temperatures and water outflows are normalized with respect to the goals, then the goal is reached when the demand and temperature are both equal to one. This allows the directness of the paths to the goal to be compared across trials with different goals. Figure 3 shows the state space diagrams for one participant on his first completed trial and last normal trial. Note that participants generally did not take the most direct path through the state space towards the goal. The

participants first reached the temperature goal and then the demand goal.

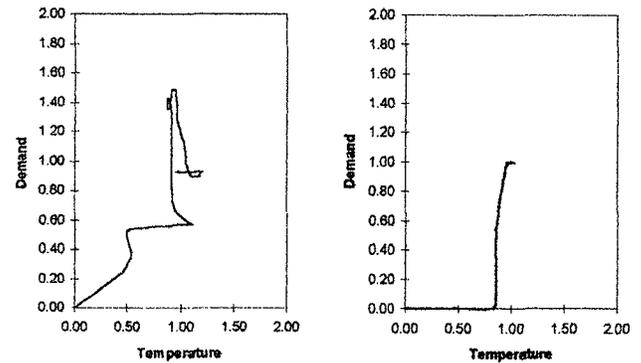


Figure 3. Comparison of state space diagrams for first (left) and last trial (right).

In order to make more quantitative comparisons, the length of a path in state space was used as a measure of directness of path. The mean path lengths in the state space diagrams became more direct with experience, particularly for the P+F group (Table 3). The paths for the P+F group were less direct initially than those for the P group, but became comparably direct with experience. This Block effect was marginally significant for reservoir 1 ($F(1,4)=5.94$, $p=0.07$). For reservoir 2, there was a marginally significant effect for Interface \times Block ($F(1,4)=5.99$, $p=0.07$), reflecting the different rates of change of path length in each block of trials. The P+F group experienced a decrease in path length between the first and last block of trials for reservoir 2, while the P group actually experienced an increase on average. There were also significant effects for Block \times Trial ($F(19,49)=3.65$, $p<0.0001$), and Interface \times Block \times Trial ($F(17,49)=3.93$, $p<0.0001$) for reservoir 2.

Table 3. Mean Path Length in State Space.

	Path Length			
	Trials 1-22		Trials 196-217	
	Reservoir 1	Reservoir 2	Reservoir 1	Reservoir 2
P+F Group	4.65	5.03	2.23	2.35
P Group	3.40	2.41	2.36	3.22

The Euclidean distance to each goal can also be plotted versus time. The area under these graphs decreased with experience. These areas also correlated with the steady state times, indicating that the two performance measures generally agree (Howie, 1996).

Mass and Energy Inventories

Graphs of energy inventory versus mass inventory are an additional way of examining the system state with respect to the goal state. The temperature goals appear as straight lines

on these graphs since temperature is proportional to the ratio of energy inventory to mass inventory. Figure 4 shows the graphs for a participant on his first completed trial and last normal trial. The divergence of the participants' paths from the temperature goal line (dotted) became less with experience for all of the participants. Further, the closeness of fit to the goal line indicated the participant's relative expertise.

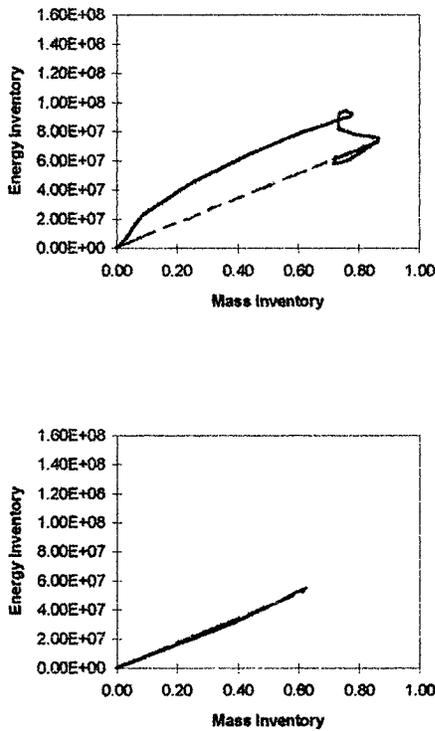


Figure 4. Comparison of energy inventory versus mass inventory graphs for first (top) and last trial (bottom).

The volume of water in the reservoir at the end of a trial (or final mass inventory) is indicated by the last point on each graph in Figure 4. With high mass inventories, the goal temperature tolerance on the energy inventory is broader, making the system easier to control. This is only directly visible on the P+F interface. Participants in the P+F group were more likely to maintain high mass inventories (above 1/3 of capacity) than participants in the P group ($\chi^2(1, N = 108) = 25.65, p < 0.01$ for reservoir 1, and $\chi^2(1, N = 108) = 2.64, ns$ for reservoir 2). In the final block, participants in the P+F group had a final reservoir volume above 1/3 for 100% of the trials, compared to only 61% for the P group. This suggests that P+F group members adapted to the goal relevant constraints visible in their interface.

Timelines

Timelines represent a sequence of actions over time (Moray et al., 1986). Control actions are plotted against time on the horizontal axis, and are grouped according to the component acted upon on the vertical axis. Each time a

component setting is changed, this control action is indicated by a square on the graph. Figure 5 shows the timelines for a participant on his first completed trial and last normal trial. With increasing experience, the participants generally made a larger proportion of their control actions near the beginning of each trial. Their later control actions were mainly heater adjustments. This replicates the work of Moray et al. (1986).

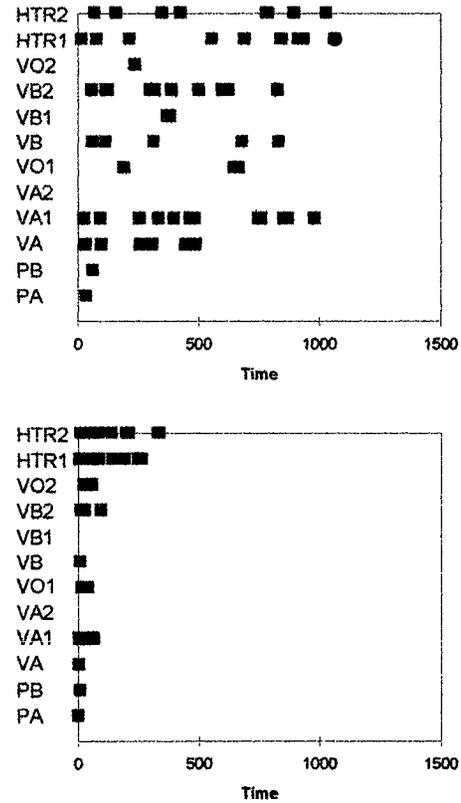


Figure 5. Comparison of timelines for first (top) and last trial (bottom).

CONCLUSIONS

A series of measures of performance were discussed. In particular, the mass inventory vs. energy inventory graphs and quantitative measures of action transition complexity and path length in state space are relatively novel. All measures showed that participants in both the P+F and P groups developed greater expertise over a six month period. Participants in the P+F group demonstrated greater improvement in steady state times, action transition graph complexity, and path length in state space, compensating for initially slower performance. Action transition graph complexity provided weak evidence that rigid procedural strategies are associated with successful performance for the P group but not for the P+F group. The mass inventories suggested that the P+F group members adapted to a goal relevant constraint visible in their interface -- namely that the system is easier to control with high mass inventories. These performance measures provide a basis for developing more comprehensive indicators of adaptation for use in actual or simulated process control environments.

ACKNOWLEDGMENTS

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

REFERENCES

Christoffersen, K., Hunter, C. N., & Vicente, K. J. (1994). *Research on factors influencing human cognitive behaviour I* (CEL 94-05). Toronto: Cognitive Engineering Laboratory, University of Toronto.

Howie, D. E. (1996). *Shaping expertise through ecological interface design: Strategies, metacognition, and individual differences* (CEL 96-01). Toronto: Cognitive Engineering Laboratory, University of Toronto.

Moray, N., Lootsteen, P., & Pajak, J. (1986). Acquisition of process control skills. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-16, 497-504.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, 13, 257-266.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Transactions on Systems, Man, and Cybernetics*, 15, 234-243.

Sanderson, P. M., Verhage, A. G., & Fuld, R. B. (1989). State-space and verbal protocol methods for studying the human operator in process control. *Ergonomics*, 32, 1343-1372.

Vicente, K. J., & Rasmussen, J. (1990). The ecology of human-machine systems II: Mediating "direct perception" in complex work domains. *Ecological Psychology*, 2, 207-249.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, 22, 589-606.

APPENDIX A

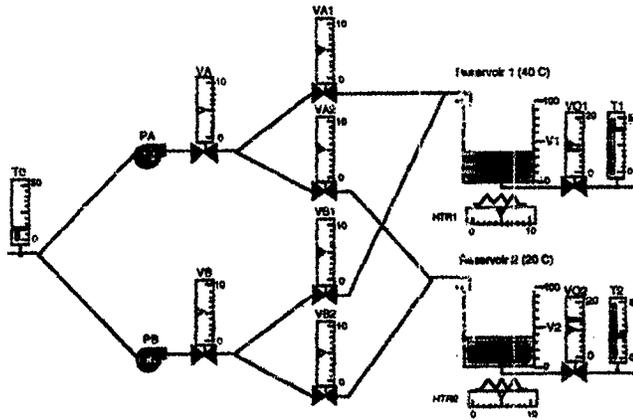


Figure A1. P Interface.

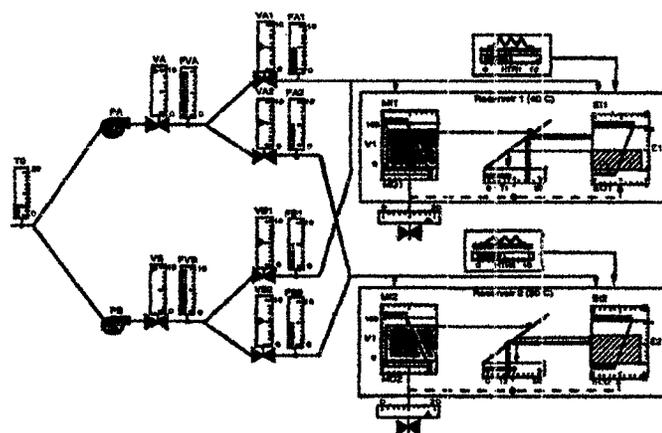


Figure A2. P+F Interface.