Attention Allocation within the Abstraction Hierarchy

Michael E. Janzen and Kim J. Vicente

Proceedings of the Human Factors and Ergonomics Society Annual Meeting 1997 41: 274
DOI: 10.1177/107118139704100162

ATTENTION ALLOCATION WITHIN THE ABSTRACTION HIERARCHY

Michael E. Janzen & Kim J. Vicente
Cognitive Engineering Laboratory
Department of Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada

Rasmussen (1985) proposed the abstraction hierarchy, consisting of physical and functional system models, as a basis for interface design for complex human-machine systems. In this study, subjects used an interface consisting of four windows, each representing a level of the abstraction hierarchy, to control a thermal-hydraulic process simulation. The goal was to investigate the relationship between attention allocation strategies and performance under normal and abnormal conditions. Subjects controlled the process for about one hour per weekday for approximately one month. The results indicate that subjects who made more frequent use of functional levels of information exhibited more accurate system control under normal conditions, and more accurate diagnosis performance under fault trials. Moreover, subjects who made efficient use of functional information exhibited faster fault compensation times. These results provide specific evidence of the advantages of functional information in an abstraction hierarchy interface.

INTRODUCTION

Rasmussen's (1985) abstraction hierarchy (AH) is a framework for representing complex human-machine systems. Each level of the hierarchy provides a different language for describing a system (see Bisantz & Vicente, 1994). In process control, five levels have been found to be of use: a) Goals - the purposes for which the system was designed; b) Principles - the first principles (i.e., mass and energy conservation laws) describing system behavior; c) Flows - the flow functions built into the system; d) Settings - the state (i.e., settings) of the equipment comprising the system; e) Form - the spatial location and visual appearance of the equipment.

The study described here investigated how subjects allocated attention within the AH in a process control microworld. This allows us to determine which attention allocation strategies lead to enhanced performance. If the rationale behind the abstraction hierarchy is correct (see Vicente & Rasmussen, 1992; Christoffersen, Hunter, & Vicente, 1997), subjects who demonstrate greater use of functional levels of information should exhibit better performance. One way to test this hypothesis is to divide the content of an interface based on an AH representation of a process into multiple windows, each representing a level of the AH. By allowing subjects to view only one level at a time, we can determine when, and perhaps why, subjects consult specific levels of abstraction. As far as we know, this is the first experiment to investigate this issue using objective measures of performance, rather than verbal protocol measures.

METHOD

Research Vehicle

The testbed used in this study, DURESS (DUal REservoir System Simulation) II , is illustrated in Figure 1. The system consists of two pumps (PA, PB), eight valves (VA, VA1, VA2, VO1, VB, VB1, VB2, VO2), two heaters (H1, H2), and two reservoirs. Subjects manually controlled these components in order to achieve two goals for each reservoir, a temperature setpoint (T1, T2) and an output demand setpoint (D1, D2) (see Figure 4).

Multilevel Interface

The multilevel interface used by subjects consists of 4 windows, illustrated in Figures 1 to 4, each corresopnding to one level of the AH (for a more detailed description of how the interface was design and how it works, see Vicente & Rasmussen, 1990 and Pawlak & Vicente, 1996, respectively). The Settings window only displays the state of the components. The Flows window only indicates the state of the functions that each component is intended to satisfy. The Principles window represents each reservoir in terms of a mass and energy balance using emergent features displays. The Goals window only shows the state of the goal variables, reservoir temperature and output demand. Subjects had to view only one window at a time but could switch freely between windows. The fifth level of the AH, Form, was not included in this interface because a live video feed would be required to display the location and appearance of the equipment.

Introductory Sessions

During the first session, each of six subjects (primarily male engineering graduate students) was presented with a description of the experiment, and a technical description of DURESS II. During the second session, subjects were provided with an explanation of the interface.

Task

On subsequent days, subjects were required to start up the system for each experimental trial. This required that the system be brought from a shut-down state to a steady-state condition, meeting the pre-defined temperature and outflow demand goals. For all trials, steady-state was defined as having maintained the system goals (temperature and outflow demand,

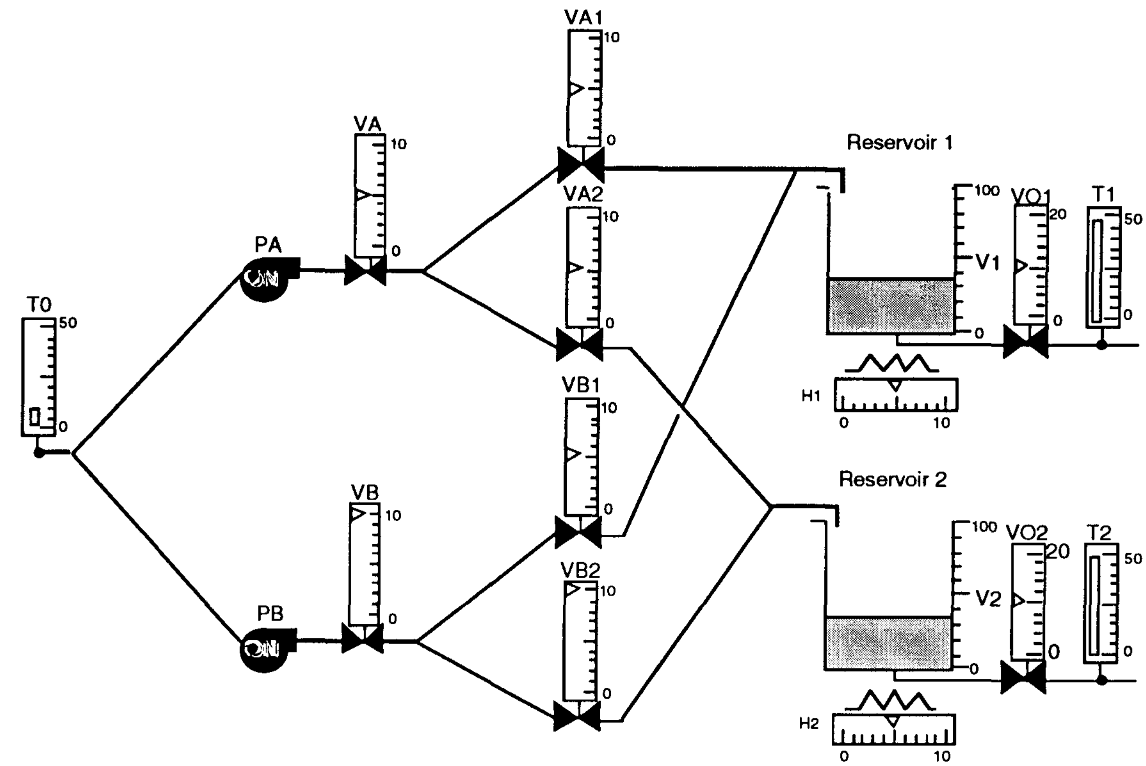

Figure 1. The Settings window in the multilevel AH interface used in this study.

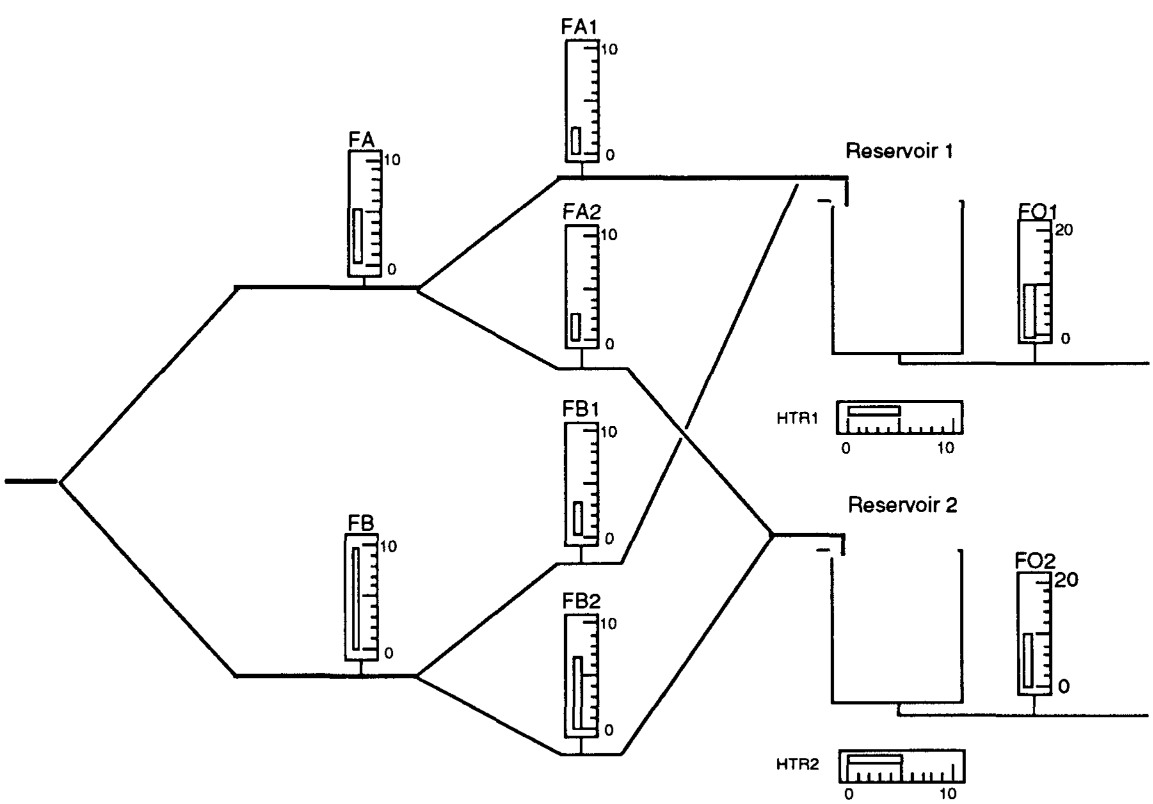

Figure 2. The Flows window in the multilevel AH interface used in this study.

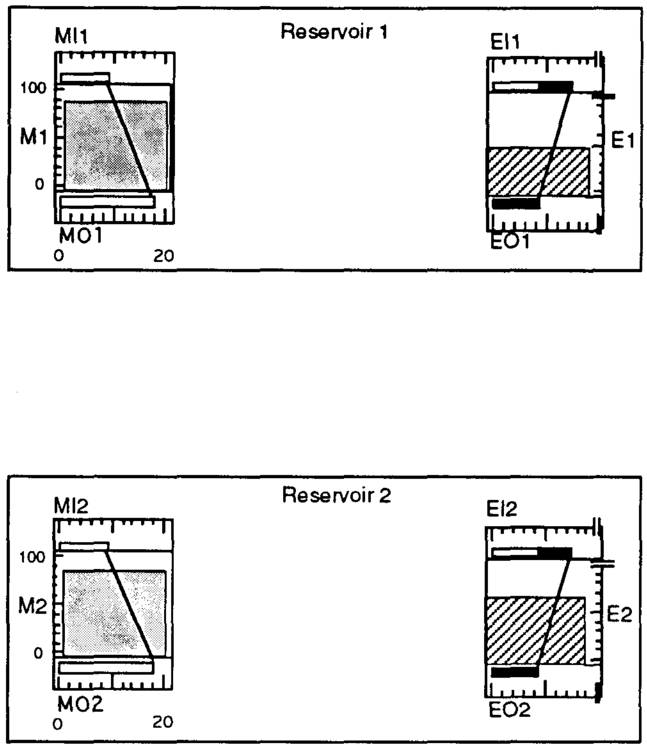

Figure 3. The Principles window in the multilevel AH interface used in this study.

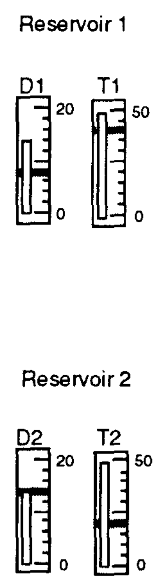

Figure 4. The Goals window in the multilevel AH interface used in this study.

see Figure 4) for both reservoirs for 5 consecutive minutes. There were three trial types: normal, routine fault, and non-routine fault. During normal trials, the system operated according to a physical model based on conservation laws. During routine fault trials, a simple fault occurred (e.g., reservoir leak). During non-routine fault trials, a simulated complex fault occurred. If subjects had not achieved steady state within 30 minutes, the trial was terminated by the experimenter. If subjects performed actions that damaged system components then the system "blew up" and the trial was automatically terminated. Several trials were presented to each subject on each day. In sum, each subject was required to control the system for about an hour per weekday for approximately one month, for a total of 67 trials.

Performance Measures

There were two primary sources of data: a log of subjects' actions and system variables, and verbal protocols collected during fault trials.

RESULTS

Four multiple, linear regression analyses were conducted to determine which attention allocation measures predicted subjects' performance in normal or fault trials. One analysis was conducted for each of the following measures: normal trial completion time, fault detection time, fault diagnosis accuracy, and fault compensation time. A total of 14 attention allocation measures were used as predictors in each analysis. For each window, the following attention allocation measures were included: Frequency of Visits, Relative Frequency of Visits, Dwell Time, and Relative Dwell Time (because of missing data, Dwell Time and Relative Dwell Time for the Flows window were omitted). Frequency of Visits measured the average number of visits per minute to each window for each trial for each subject. Dwell Time measured the average time spent per visit to each window for each trial for each subject. Frequency of Visits were normalized for each window for each subject by dividing by overall Frequency for each subject and each trial. We also conducted analyses on normalized Dwell Time. For each window, Relative Dwell Time was calculated as the ratio of Dwell Time in that window to overall Dwell Time for each trial for each subject. Only variables significant at the $p < 0.05$ level were entered into the regressions. All R-squared values refer to cumulative model R-squared. The primary question of interest is the proportion of variance accounted for by the predictor variables, not the precise quantitative parameters in the regression equations. Thus, only the results relevant to this question are presented below.

Normal Trial Completion Time

Completion times measured the time that subjects took to bring the system from a shut-down to a steady-state condition. Blowup trials were counted as missing data. In order not to further reduce the number of data points, a conservative completion time of 30 minutes was recorded for timeouts.

To study the relationship between attention allocation and asymptotic performance on normal trials, a stepwise linear regression was performed with data from the last 10 normal trials. Only one variable entered into the model:

- The (negative) predictor was Relative Frequency of Visits to Principles ($F(1, 58) = 15.77$, R-squared = 0.21, $p = 0.0002$).

The more often subjects visited Principles (Figure 3), the faster their performance on normal trials.

Fault Detection Time

The regression analyses failed to indicate any significant linear relationship between attention allocation strategies and fault detection time.

Fault Diagnosis Accuracy

To study the relationship between attention allocation and diagnosis, a stepwise linear regression for fault diagnosis accuracy was performed for all fault trials. Diagnosis accuracy was extracted from the verbal protocols based on a method used by Pawlak & Vicente (1996). Each attempt at diagnosis was assigned an accuracy score, from 0 to 3. The highest level of diagnosis made for a particular fault trial was recorded as the Diagnosis Score for that fault.

To create a model to predict Diagnosis Score, a stepwise linear regression was performed on data for all faults, using attention allocation measures from the corresponding fault trials. Two variables entered into the model:

- The first (positive) predictor was Relative Frequency of Visits to Flows ($F(1, 53) = 26.90$, R-squared = 0.34, $p = 0.0001$).
- The second (positive) predictor was Frequency of Visits to Principles ($F(2, 52) = 4.47$, R-squared = 0.39, $p = 0.04$).

The more often subjects visited Flows (Figure 2) and Principles (Figure 3), the more accurate their diagnosis scores.

Fault Compensation Times

To study the relationship between attention allocation and fault compensation performance, a stepwise linear regression for trial completion time was performed for all fault trials. Trial was added as a predictor to see if the variance in times was due to idiosyncratic trial effects rather than systematic strategy differences. Type was also added (as a dummy variable) to distinguish non-routine from routine fault trials. Three variables entered into the model:

- The first (positive) predictor was Dwell Time in Principles ($F(1, 45) = 20.00$, R-squared = 0.31, $p = 0.0001$).
- The second (positive) predictor was fault Type ($F(2, 44) = 6.75$, R-squared = 0.40, $p = 0.013$).
- The third (negative) predictor was Trial ($F(3, 43) = 8.03$, R-squared = 0.49, $p = 0.007$).

Thus, dwelling briefly in Principles (Figure 3) was related to shorter completion times. Non-routine faults took longer to compensate for than routine faults, and trials occurring later in the experiment took less time than those occurring earlier.

DISCUSSION

Despite the large number of potential predictors, a relatively simple pattern of results emerged from these regression analyses, emphasizing the significant role of functional levels of the AH. For normal trial completion times, Relative Frequency of Visits to Principles was the only significant predictor, accounting for 21% of the variance. The more often subjects visited Principles, the faster their performance on normal trials. This is probably because the vertical lines showing the rate of change of mass and energy (see Figure 3) provided a useful cue for stabilizing mass and energy, which in turn stabilized temperature. The sooner subjects effectively stabilized the system, the quicker they would be able to complete the trial.

For fault diagnosis accuracy, there were two significant predictors: Relative Frequency of Visits to Flows and Frequency of Visits to Principles, accounting for 34% and 5% of the variance, respectively. Subjects who visited Flows and Principles more frequently exhibited more accurate diagnoses. Thus, these two levels of the AH, which contained functional information, helped subjects to trace through the multiple levels of abstraction until the source of the fault was identified. For example, a heater failure can be best diagnosed by comparing the heater state in the Settings window with the heat transfer rate presented in the Flows window. The relationship that normally links these two variables is violated in the case of a heater failure, so the functional information in the Flows window is essential for diagnosis. This result indicates that functional information is important for fault diagnosis performance (cf. Christoffersen et al., 1997).

For fault compensation times, three predictors were significant: Dwell Time in Principles, fault Type, and Trial number, accounting for 31%, 9%, and 9% of the variance, respectively. The fault Type effect indicates that non-routine faults took longer to complete because they are more difficult than routine fault trials. The Trial effect indicates that trials occurring later in the experiment took less time to complete due to learning. The third, and most significant, effect indicates that subjects who dwelled less in Principles tended to compensate for faults more quickly. Thus, it seems that the faster subjects were able to exploit the emergent features in the Principles window (Figure 3) to extract information quickly.

CONCLUSIONS

This study shows the importance of using the AH as a basis for interface design for process control systems. To merely complete the task, the only necessary windows are Settings and Goals. We know this because, in previous studies, subjects completed the same task using an interface which had only these two levels of abstraction (e.g., Pawlak & Vicente, 1996). Subjects who were satisfied with merely completing the task allocated most of their attention to one of these windows, and most of their remaining attention to the other. In order to not only complete the task, but to complete it quickly, it was necessary to be able to control the system accurately. This required frequent visits to Principles, which allowed reliable stabilization of mass and energy, which in turn stabilized temperature. Also, to accurately diagnose faults, it was necessary to be able to identify broken constraints. Most faults resulted in broken constraints that could be identified by switching between Settings and Flows, or between Flows and Principles, making Flows a key window. Subjects who were good at diagnosis visited Flows regularly. Other subjects ignored Flows much of the time, because faults were uncommon, and because Flows were usually not relevant to the immediate goal of trial completion. Consequently, they did not perform as well at diagnosis.

In summary, these results provide specific evidence of the importance of functional information. Thus, they show the advantages of adopting the abstraction hierarchy as a basis for interface design rather than traditional interface practices which pay more attention to physical information alone.

ACKNOWLEDGEMENTS

This research was sponsored by a contract from the Japan Atomic Energy Research Institute (Dr. Fumiya Tanabe, Contract Monitor), as well as research and equipment grants from the Natural Sciences and Engineering Research Council of Canada. We would like to thank Dr. Tanabe, Dianne Howie, Darryl Minard, and Thomas Smahel for their various contributions to this research. We also thank the six subjects who participated in the study.

REFERENCES

Bisantz, A. M., & Vicente, K. J. (1994). Making the abstraction hierarchy concrete. *International Journal of Human-Computer Studies, 40*, 83-117.

Christoffersen, K., Hunter, C. N., & Vicente, K. J. (1997). A longitudinal study of the effects of ecological interface design on fault management performance. *International Journal of Cognitive Ergonomics, 1*, 1-24.

Pawlak, W. S., & Vicente, K. J. (1996). Inducing effective operator control through ecological interface design. *International Journal of Human-Computer Studies, 44*, 653-688.

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

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