Sensor Noise and Ecological Interface Design: Effects of Increasing Noise Magnitude on Operators' Performance

Olivier St-Cyr and Kim J. Vicente

Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2005 49: 417
DOI: 10.1177/154193120504900343

SENSOR NOISE AND ECOLOGICAL INTERFACE DESIGN: EFFECTS OF INCREASING NOISE MAGNITUDE ON OPERATORS' PERFORMANCE

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

We studied the impact of sensor noise on operators' performance using a display based on the Ecological Interface Design (EID) framework with a representative thermal-hydraulic process simulation. A previous study conducted by St-Cyr and Vicente (2004) showed no difference between EID and non-EID interfaces when the magnitude of sensor noise was randomly increased. In this paper, we describe a study that was designed to investigate the impact of gradually increasing the magnitude of sensor noise on EID versus non-EID interfaces. We hypothesized that as the magnitude of sensor noise increase, performance would worsen for both EID and non-EID participants. Our results suggest that increasing the magnitude of sensor noise does compromise both EID and non-EID interfaces. However, the EID group experienced a significantly larger decrease in performance. This may be explained by the fact that participants in the EID group had to deal with distorted emergent features.

INTRODUCTION

Ecological Interface Design (EID) is a framework for designing human-machine interfaces for complex systems (Vicente and Rasmussen, 1992). Over the past years, the framework has been applied to a variety of domains (see Vicente, 2002, for a comprehensive review). To implement EID interfaces, several sensors must be used to acquire data from the work domain and display them in meaningful graphical representations, creating emergent features. This paper investigates potential effects of the magnitude of sensor noise on the EID framework.

Background

A sensor can be defined as a device that receives and responds to a signal or a stimulus (Fraden, 1997). More specifically, "a sensor converts the physical dimension which is to be measured into an electrical dimension which can be processed or transmitted electronically" (Hauptmann, 1993, p. 4). A very important characteristic of sensors is accuracy, which represents the highest deviation of a value from the true input. In that sense, accuracy often means inaccuracy and is represented by a plus or minus (+/-) range in which the sensor is expected to stay. This range is often referred to as sensor noise (Fraden, 1997).

When EID was introduced, Vicente and Rasmussen (1992) pointed out that noisy sensors are a source of data uncertainty that could compromise the robustness of EID interfaces (cf. Reising and Sanderson, 2002a). Such interfaces will often include several emergent features (Bennett, Toms, and Woods, 1993) that are produced from relationships between low-level graphical elements. Because emergent features are derived from low-level data (which are normally obtained through sensors), sensor noise could adversely affect high-level constraints to be portrayed on the interface, compromising the geometric forms of configural displays (Reising and Sanderson, 2002b).

A similar argument was made by Vicente, Moray, Lee, Rasmussen, Jones, Brock, and Djemil (1996) who pointed out that sensor failures could create distortions in emergent features, affecting the behaviour of operators. Conversely, they also suggest that any distortions due to sensor failures may also create salient information to help operators in detecting problems.

Finally, Vicente (2002) also mentioned two possible effects of sensor noise on EID interfaces. First, the robustness of EID interfaces may not be compromised by sensor noise due to the redundant constraints portrayed on the interface. Second, as mentioned in Vicente, et al. (1996), sensor noise may also confuse operators in their ability to distinguish between the displayed state and the true state of the work domain.

Previous Studies

St-Cyr and Vicente (2004) first studied the impact of the magnitude of sensor noise on operators' control performance using an EID interface. Their results show that EID participants performed significantly better than non-EID participants when dealing with a level of noise corresponding to industrial averages. Then, when the magnitude of the noise was randomly increased, no significant differences were observed between the two groups.

While their results showed that the performance of the EID group was not worse than that of the non-EID group, there were several limitations to their study. First, operators were only exposed to a few trials (20 out of 80) in which the magnitude of sensor noise was changed. Second, the level of sensor noise was randomly changed, which made it difficult to investigate the magnitude at which the robustness of EID interfaces could be compromised. Third, the trials were programmed in such a way that some participants were able to derive mathematical heuristics to help them reach steady-state conditions even though sensors were noisy. In that sense, varying the magnitude of sensor noise did not affect their performance.

Current study

The current study was designed to assess some of these limitations by investigating the impact of gradually increasing the magnitude of sensor noise on operators' performance. Thus, in this study, there were more trials in which the magnitude of sensor noise was increased (50 trials). This change was introduced to ensure that participants would get enough exposure to the different magnitudes of sensor noise. Finally, the trials were also designed to make it more difficult for participants to find mathematical heuristics.

Hypothesis

Based on the literature outlined above, we predicted that as the magnitude of sensor noise increase, performance would worsen for both EID and non-EID groups. Moreover, we also predicted that the effects of the presence of an increased magnitude of sensor noise would be significantly stronger for the EID group.

METHOD

Participants

The twenty participants were engineering undergraduate students from the departments of Mechanical and Industrial Engineering and Chemical Engineering at the University of Toronto, contacted by local advertisements. Participants (8 females and 12 males; 8 CHEM Eng. and 12 MIE) were between the ages of 17 to 22 years old (mean = 19.6). They were all selected based on their willingness to participate, their expertise with computer systems, and their cognitive style (see Torenvliet, Jamieson, and Vicente, 2000 for more details). All participants had taken at least two courses in physics. Each participant was paid at a maximum rate of $10 per hour ($5 for each hour, $3 for completing the study, and $2 for good performance).

Apparatus

The study was conducted using the DURESS II simulation, a representative thermal-hydraulic process simulation operated through a visual display. The goals of the simulation were to keep the two reservoirs at a prescribed temperature (**T1ºC** for Reservoir 1 and **T2ºC** for Reservoir 2) and to satisfy two outputs demand flow rates (**D1** and **D2**). Two different interfaces (P and P+F) for the same microworld were developed (Vicente & Rasmussen, 1990; Pawlak & Vicente, 1996). The P interface (Figure 1) displays primarily physical

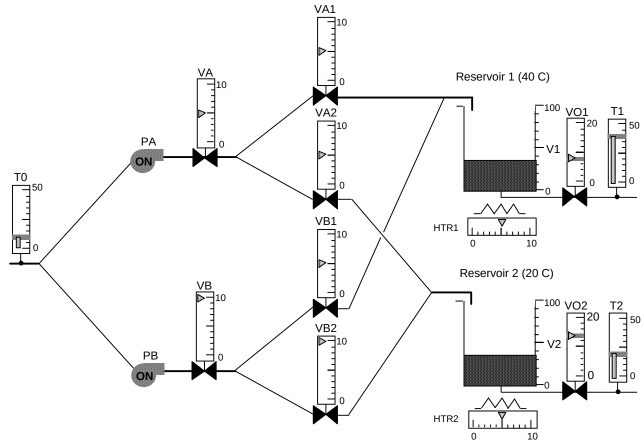

Figure 1. P Interface of DURESS II.

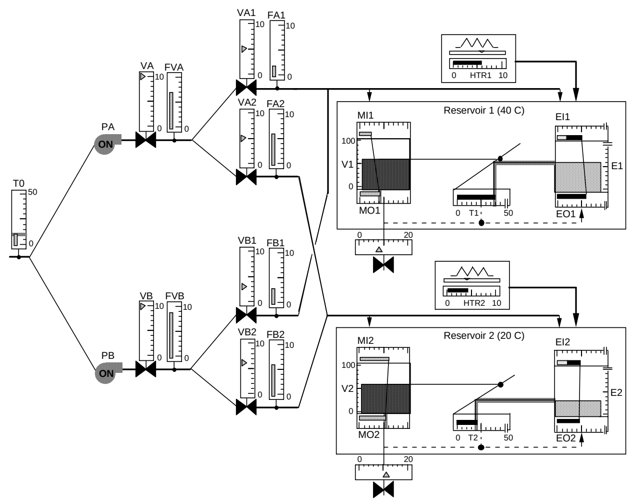

Figure 2. P+F Interface of DURESS II.

information about the system. In contrast, the P+F interface (designed using the EID framework, Figure 2) displays both physical and functional information about the system by means of emergent features based on low-level sensor data.

To study the effects of sensor noise, an updated version of the DURESS II microworld was implemented (previous versions did not incorporate sensor noise). The updated version allowed the experimenter to add sensor noise to all five sensors of the P interface and all 17 sensors of the P+F interface. White normally distributed Gaussian noise was added to true readings in the form of an accuracy range (e.g., ± 2ºC). Then, a scaling multiplier was used to increase the magnitude of the noise. The simulation ran on SGI IRIS INDIGO R4400 and SGI OCTANE R10000 machines. Participants received feedback about the state of the system through 21" high-resolution colour graphics monitors and controlled the simulation using a computer mouse.

Note that in the current study, the temperature goals were not fixed at **40ºC** for Reservoir 1 and **20ºC** for Reservoir 2 (as in

St-Cyr and Vicente, 2004), but were predetermined by the experimenter and changed from trial to trial. This modification was introduced to make it more difficult for participants to find a relationship between heater settings, temperature demands, and output demands. The experimenter presupposed that this manipulation would provide a more complete picture of the impact of different magnitudes of sensor noise on performance.

Design and Experimental Procedure

The study followed a mixed design with interface as a between-participants factor and noise magnitude as a within-participants factor. Ten participants were assigned to either the P interface or the P+F interface. Participants first completed the Spy Ring History test (Pask and Scott, 1972; Pask, 1976). Previous research (Torenvliet *et al*., 2000) had identified an interaction between these test scores and performance with DURESS II. Spy Ring scores were used to balance the groups. Patterns of answers were scored on three dimensions: Holist, Serialist, and Neutral (Pask and Scott, 1972), providing the cognitive style tendencies of a person. Using a minimum distance algorithm, unique pairs with the lowest distance were computed and members of each pair were randomly assigned to one of the two groups. Participants were then trained for a period of two hours on how to operate the simulation using their respective interface. This training session was administered through oral tutorials. After the training session, participants completed a brief questionnaire to test their level of understanding of the simulation.

Participants had to complete 110 trials (averaging to 35 one-hour daily sessions). For each trial, participants were presented with a shutdown system and were asked to bring the simulation to a steady-state, condition in which the four system's goals (two output demands and two temperature demands) had to be met for five consecutive minutes. During the first 60 trials, a level of noise corresponding to industrial averages was introduced to all sensors of the P and P+F interfaces (learning phase). Sensor noise was then gradually increased globally (all sensors) throughout the P and P+F interfaces in blocks of 10 trials. The magnitude of sensor noise was varied between blocks based on scaling multipliers (2, 3, 5, 7, and 10). See Table 1 for a distribution of the scaling multipliers over the last 50 trials.

Table 1. Distribution of the scaling multipliers for trials 61 to 110.

Scaling multipliers	Blocks and Trials
2	Block 4: 61-70
3	Block 5: 71-80
5	Block 6: 81-90
7	Block 7: 91-100
10	Block 8: 101-110

Measures

Trial Completion Time (TCT): this measure was defined as the time (in seconds) it took participants to bring the simulation to steady-state. At the end of this time (i.e., when "steady-state" is reached), the trial ended automatically.
Number of Oscillations (NO): this measure assessed participants' control stability by counting the number of times the four goal variables (two output demands and two temperatures demands) oscillated above and below the target regions in a trial. To be conservative, only the maximum value within each trial was used in the analysis. Data were stored in time-stamped data logs that were collected from the simulation in an unobtrusive way while participants completed the study.

RESULTS

The 110 trials were divided in three blocks of 20 trials and five blocks of 10 trials. For each trial, TCT and oscillation values were extracted from the log files for each participant. Data were then averaged by blocks and interface groups. Results are shown in Figures 3 through 5. Bars around the mean values represent 95% confidence intervals. If the confidence interval bars for two means do not overlap, then the difference between means is statistically significant at the $p < .05$ level.

Results from the TCT measure suggest a learning effect over the first 60 trials. By block 3, the P+F group was significantly faster than the P group with industrial level of sensor noise. This result replicates the one obtained in St-Cyr and Vicente (2004). Once sensor noise was gradually increased from block 4 to block 8, both groups experienced a gradual increase in TCT.

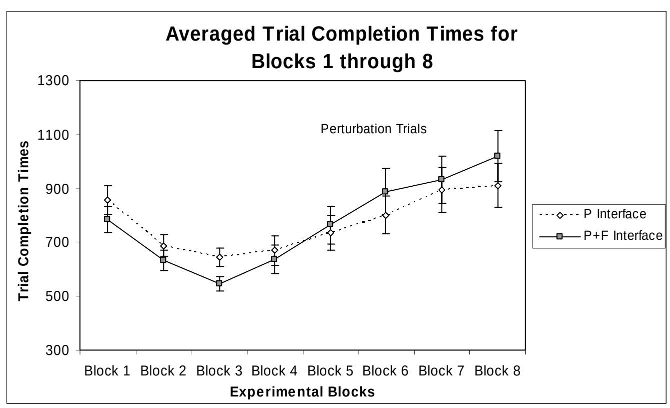

Figure 3. Averaged Trial Completion Times (TCT) for each of the eight experimental blocks.

Although the differences between groups were not significant, the gradual increases in sensor noise magnitude had a significantly larger impact on participants in the P+F group, especially from block 5 and onward, which correspond to large increases in sensor noise magnitude. Figure 4 shows the percentage of change in TCT between block 3 and all other perturbation blocks for both the P and P+F groups.

Results for the number of oscillations suggest that P+F participants seemed more stable than P participants during the first three blocks, although results were not statistically significant. In the perturbation blocks, the stability of P participants was not affected by any increases in sensor noise. As for P+F participants, they experienced non-significant increases in the number of oscillations followed by a decrease in blocks 7 and 8, suggesting potential learning on how to cope with increases in sensor noise over time.

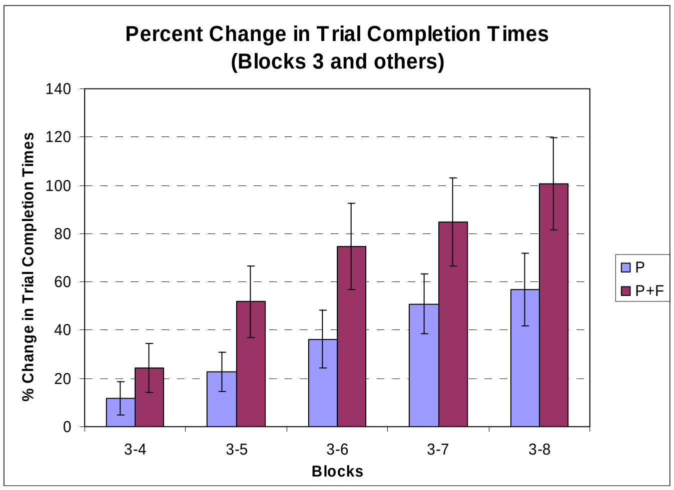

Figure 4. Percentage of change between block 3 and blocks 4 to 8.

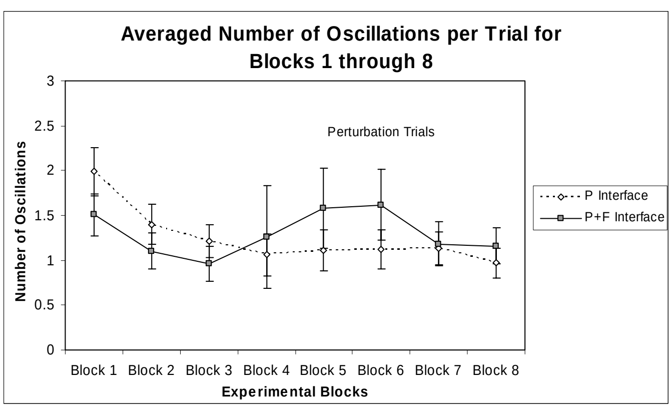

Figure 5. Averaged number of oscillations per trial for each of the eight experimental blocks.

DISCUSSION

The aim of our study was to determine the impact of sensor noise on operators' performance when the magnitude of the noise was gradually increased according to scaling multipliers. Three findings emerge out of this research. First, the results show that when sensor noise is gradually increased, both the P and P+F group experienced increases in TCT. Moreover, there was no significant difference between the two groups in the perturbation trials. This suggests, as our hypothesis proposed, that both P and P+F groups experience a decrease in performance when the magnitude of sensor noise was gradually increased.

The second finding suggests that while there was no significant difference between the two groups in each of the perturbation blocks, the P+F experienced a significantly higher percentage of change in TCT when comparing industrial averages sensor noise (block 3) to the different increases in magnitude (blocks 5 to 8). In that sense, increasing the magnitude of sensor noise had a stronger impact for the P+F group than the P group.

The third finding suggests that when sensor noise was gradually increased, the control stability of P+F participants was affected, although this trend was not observed in the last two blocks of the study, suggesting potential learning on how to cope with a noisy interface and thus, a possible regain of stability.

Altogether, these results support our hypothesis and suggest that increasing the magnitude of sensor noise does compromise the performance of both the P+F and P interfaces. However, the P+F group experienced significantly larger increase in TCT than the P group. This is especially true in blocks 5, 6, 7, and 8, which correspond to large increases in sensor noise magnitude. This may be explained by the fact that the P+F group had to deal with distorted emergent features while the P group did not.

These findings are believed to be important for the applicability of EID in industrial settings. For example, Watanabe (2001) observed that investigating the robustness of EID interfaces with noisy sensors is necessary before industry would be confident enough to apply the framework in real industrial settings. Our results suggest that the performance and stability of the P+F interface, while not inferior than that of the P interface, were compromised by gradual increases of sensor noise magnitude, especially when dealing with substantial amount of noise.

Nonetheless, as Vicente, et al. (1996) pointed out, another important aspect of EID and sensor noise is to investigate the effects on operators' abilities to interpret and to cope with uncertain data. The next step in our research program will be to analyze control strategies used by operators when the magnitude of sensor noise was increased. These strategies were collected using control recipes, in which participants had to write a set of instructions on how to operate the DURESS II simulation and will constitute further research investigations.

A final study will also investigate the impact of sensor noise on local sensors under more representative conditions. Results from this study may help in improving the design process of interfaces for complex systems, with respect to sensor verification and validation for EID interfaces

ACKNOWLEDGEMENTS

Funds to support this work came from the Natural Science and Engineering Research Council of Canada (NSERC) and the E. W. R. Steacie Memorial Fellowship. The first author would like to express his gratitude to the University of Toronto Glynn Williams Fellowship foundation. We would also like to thank Ridha Ben Mrad and especially Dal Vernon Reising for his thoughtful comments and discussions of the ideas presented herein.

REFERENCES

Bennett, K. B., Toms, M. L., & Woods, D. D. (1993). Emergent features and graphical elements: Designing more effective configural displays. *Human Factors*, *44*, 71-97.

Fraden, J. (1997). *Handbook of modern sensors: physics, designs, and applications, Second Edition*. Woodbury, NY: American Institute of Physics.

Hauptmann, P. (1993). *Sensors: Principles and applications*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology*, *46*, 128-148.

Pask, G., & Scott, B. C. (1972). Learning styles and individual competence. *International Journal of Man-Machine Studies*, *4*, 217-253.

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

Reising, D. V. C. & Sanderson, P. M. (2002a). Work domain analysis and sensors I: Principles and simple example. *International Journal of Human-Computer Studies*, *56*, 569-596.

Reising, D. V. C. & Sanderson, P. M. (2002b). Work domain analysis and sensors II: Pasteurizer II case study. *International Journal of Human-Computer Studies*, *56*, 597-637.

Reising, D. V. C. & Sanderson, P. M. (2004). Minimal instrumentation may compromise failure diagnosis with an ecological interface. *Human Factors*, *46*, 316-333.

St-Cyr, O., & Vicente, K. J. (2004). Sensor Noise and Ecological Interface Design: Effects on operators' Control performance. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, pp. 538-542.

Torenvliet, G. L., Jamieson, G. A., & Vicente, K. J. (2000). Making the most of ecological interface design: The role of individual differences. *Applied Ergonomics*, *31*, 395-408.

Vicente, K. J. (2002). Ecological interface design: Progress and challenges. *Human Factors*, *44*, 62-78.

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

Vicente, K. J., Moray, N., Lee, J. D., Rasmussen, J., Jones, B. G., Brock, R., & Djemil, T. (1996). Evaluation of a Rankine Cycle Display for Nuclear Power Plant Monitoring and Diagnosis. *Human Factors*, *38*, 506-521.

Watanabe, O. (2001). *Issues toward practical application of ecological interface to real world*. Paper presented at the International Workshop on Ecological Approach to Human Work Interface, Tokai, Japan.