Proactive monitoring in process control using predictive trend displays

Yin, Shanqing

2012

Yin, S. Q. (2012). Proactive monitoring in process control using predictive trend displays.
Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/49971>

<https://doi.org/10.32657/10356/49971>

NANYANG
TECHNOLOGICAL
UNIVERSITY

Proactive Monitoring in Process Control using Predictive Trend Display

Yin Shanqing

School of Mechanical and Aerospace Engineering

2012

A dissertation report submitted to partially fulfill the requirements for the degree of Doctor of Philosophy

PROACTIVE MONITORING IN PROCESS CONTROL
USING PREDICTIVE TREND DISPLAY

YIN SHANQING

SCHOOL OF MECHANICAL AND AEROSPACE ENGINEERING

A thesis report presented to Nanyang Technological University
in partial fulfilment of the requirements for the
Degree of Doctor of Philosophy (Human Factors)

2012

Abstract

Process control such as that in the petrochemical industry is inherently difficult for humans to operate and monitor. Console operators need to manage hundreds of interrelated components using sluggish controls in a high-risk environment. They need to keep the process stable while optimizing production, which puts variables near plant operating limits. Any anomaly or upset has to be resolved quickly before the severity of the problem escalates. All these tasks are performed using a control console called the Distributed Control System (DCS).

This project was initiated with the goal of exploring viable information visualizations on DCS displays to support proactive monitoring in console operators. While operators may choose to be alerted of and react to problems through the alarms on the DCS, expert operators prefer to stay proactive, and seize the problem before it disrupts the stability of the process. Being proactive requires prediction, a mental process which is not well understood and difficult to perform accurately. A series of literature reviews was conducted to find out more about the concepts related to the psychology of prediction, followed by various engineering elements, particularly in process control, that aid prediction. Currently there are no explicit predictive displays for process control.

Four studies were conducted during the span of this project, each filling in knowledge gaps either not found in current literature, or provided empirical proof-of-concept for a viable predictive tool that improved control performance. The first qualitative investigation revealed how expert console operators derive, update and apply their mental models while at work. A second qualitative investigation documented the use of trend information displayed on current DCS consoles with the purpose of facilitating proactive monitoring. A simulator study was conducted which found operator performance benefits from using a trend-based predictive display with multi-variate rate-of-change cues. A second, final experiment featured a high-fidelity schematic display and a single-variate rate-of-change algorithm. Final results showed a viable prototype predictive visualization and algorithm for further industry application.

Acknowledgements

The author would like to thank some special people:

- Prof. Martin Helander, who had the major task of supervising this project
- Prof. Chris Wickens and Dr. Dal-Vernon Reising, who contributed immensely as external knowledge experts
- Mr. Jason Laberge and Mr. Andrew Trenchard of Honeywell International, as well as Mr. Brian Thompson of Engen Petroleum Ltd., who frequently participated in project discussions and provided insights into process control
- Mr. Pang Hong-Xiang for his good work as an assistant
- The good folks at the Centre for Human Factors and Ergonomics at NTU, who gave support in one way or another

Parts of this research were supported by the ASM Consortium, a group of leading companies and universities involved with process industries that jointly invest in research and development to create knowledge, tools, and products designed to prevent, detect, and mitigate abnormal situations that affect process safety in the control operations environment. Much of the contributions and collaborations would not have happened without the Consortium's support and sponsorship.

[www.asmconsortium.net]

Table of Contents

Chapter One Introduction

1.1	The DCS	3
1.2	Alarms & Monitoring by Exception	5
1.3	Proactive Monitoring – Reducing Critical Situations	7
1.4	Predictive Display to Aid Proactive Monitoring	9
1.5	Research Objectives	9
1.6	Outline of this Report	10
1.7	References	11

Chapter Two Basic Theories of Cue-based Mental Prediction

2.1	Overview	14
2.2	The Psychology of Prediction	15
2.3	Inductive & Deductive Reasoning	18
2.4	Situation Awareness	20
2.5	Mental Model	23
2.6	Mental Simulation	26
2.7	Expertise	28
2.8	Graphical Summary of Literature Review	31
2.10	References	35

Chapter Three The Engineering of Prediction: Predictive Aids

3.1	Overview	42
3.2	Types of Predictive Aids	43
3.3	Performance Limitations in Automated Predictions	56
3.4	Imperfection Automation	59
3.5	Summary	61
3.6	References	62

Chapter Four Predictive Applications in Process Control

4.1	Introduction	67
4.2	Model Predictive Control	69
4.3	Qualitative Trend Analysis	73
4.4	Challenges for Process Control Predictive Displays	79
4.5	Summary	81
4.6	References	82

Table of Contents

Continued

Chapter Five Field Studies

5.1	Overview	84
5.2	Qualitative Investigation One: Operators' Mental Models	84
5.3	Method: Ethnographic Observations	85
5.4	Results & Discussion from Ethnographic Study	86
5.5	Qualitative Investigation One: Summary	93
5.6	Qualitative Investigation Two: Trend Displays in Process Control	95
5.7	Method: Knowledge Elicitation Interviews	96
5.8	Results & Discussion from Interviews	98
5.9	Qualitative Investigation Two: Summary	102
5.10	Overall Summary	103
5.11	References	104

Chapter Six Simulator Study One: Predictive Cues on Trend Displays

6.1	Overview	106
6.2	Rate-of-Change Representation	107
6.3	Method	108
6.4	Results	116
6.5	Discussion	120
6.6	References	122

Chapter Seven Simulator Study Two: Rate-of-Change Visualizations

7.1	Overview	124
7.2	An Updated Rate-of-Change Algorithm	124
7.3	Method	125
7.4	Results	134
7.5	Discussion	142
7.6	References	147

Chapter Eight Concluding Remarks

8.1	Research Accomplishments	148
8.2	Project Limitations	153
8.3	Predicting Future Works	155
8.4	References	158

Appendix	160
-----------------	------------

List of Figures & Tables

Figure 1.1	A screenshot of a Trends display. Source: Honeywell Inc.	4
Figure 1.2	A screenshot of an Alarm Summary display. Source: Honeywell Inc.	4
Figure 1.3	A screenshot of a Schematic Graphics display. Source: Honeywell Inc.	5
Figure 1.4	Early intervention can prevent catastrophic losses (Burns, 2006)	8
Figure 2.1	People tend to extrapolate linearly and fail in predicting exponential growth.	16
Figure 2.2	Model of situation awareness in dynamic decision making (Endsley, 1995)	20
Figure 2.3	Graphical representation of the psychology of cue-based prediction	23
Figure 3.1	A storm prediction depicting Hurricane Frances' possible track into the future. Source: National Oceanic and Atmospheric Administration, USA	46
Figure 3.2	Representation of a closed-loop tracking operation	48
Figure 3.3	An illustrated navigation display found in aircraft cockpits, showing the "noodle" of the own aircraft as well as other traffic in the vicinity (Morphew & Wickens, 1998)	49
Figure 3.4	A screenshot of a tunnel-in-the-sky display featuring both the predictor and preview elements (Doherty & Wickens, 2001)	51
Figure 3.5	A screenshot of an ATC display, showing predictive lines of each aircraft and weather information to aid controllers' decision-making	53
Figure 3.6	A screenshot of the vessel navigation simulator featuring the predictor used by Sullivan et al (2006)	54
Figure 3.7	The Gaussian perturbation describes the growth in uncertainty as the span of prediction increases, in which probability of each possible directional change follows a normal distribution	58
Figure 4.1	Predictive trend display developed by Roth and Woods (1988)	67
Figure 4.2	An equal increase in one variable causes an increasing increment in another	68
Figure 4.3	Optimizing the predicted output through MPC (Garcia et al, 1989)	70
Figure 4.4	Representing a trend data using triangulation (Cheung & Stephanopoulos, 1990a)	74
Figure 4.5	Geometrical basic triangular components (Cheung & Stephanopoulos, 1990a)	74

List of Figures & Tables

Part ii

Figure 4.6	75
Primitives identified by Januz & Venkatasubramanian (1991)	
Figure 4.7	76
Assigning the primitives, forming the episodes, and eventually developing the trend signature (Januz & Venkatasubramanian, 1991)	
Figure 4.8	77
Formation of a trend variable by considering how a primitive trend pattern interacts with minute deviations and noise (Cheung & Stephanopoulos, 1990b)	
Figure 4.9	78
Illustration of Gaussian smoothing at successive scales	
Figure 5.1	96
The line graph on the top right clearly shows a decreasing trend in both variables through their slopes, as well as different rate-of-change between the two variables through their angles (Wickens & McCarley, 2008)	
Figure 5.2	98
A screenshot of the Trends displayed in the Honeywell TDC 3000 DCS	
Figure 5.3	100
The Human Intervention Framework	
Figure 6.1	109
The schematic layout of Honey Mixer	
Figure 6.2	111
The experimental display console	
Figure 6.3	112
Experimental display console showing graphical rate-of-change cues	
Figure 6.4	112
Experimental display console showing numerical rate-of-change cues	
Figure 6.5	113
The presence of ROC indicators should reduce duration outside operating envelope, with graphical being more beneficial than numerical visualization	
Figure 6.6	114
The 3 x 2 factorial experiment design	
Figure 6.7	115
The dependent variable and the spectrum of performance	
Figure 6.8	116
Mean duration of limit breach for each condition.	
Figure 7.1	125
Detailed breakdown of the filtered rate-of-change algorithm	
Figure 7.2	126
A screenshot of Honey Mixer II display	
Figure 7.3	129
The Predictive Indicator	
Figure 7.4	131
The experimental design for Simulator Study Two	

List of Figures & Tables

Part iii

Figure 7.5	134
Graph showing Alarm Present Measurement scores	
Figure 7.6	136
Graph showing Percentage Duration of performed scenarios with No Alarms	
Figure 7.7	137
Graph showing transformed average absolute deviation between participants' predictions and actual parameter values 1-minute later	
Figure 7.8	138
Graph showing transformed average absolute deviation between participants' predictions and FROC-derived parameter values	
Figure 7.9	139
Graph plotting participants' prediction deviations with their Alarm Presence Measurement scores. Solid line indicates best-fit linear trend	
Figure 7.10	140
Graph showing participants' average prediction time	
Figure 7.11	141
Graph showing the number of control movements	
Table 2.1	21
Presence of SA levels while achieving perception, comprehension, and projection	
Table 3.1	55
Categorization of predictive aids	
Table 5.1	87
Typical process control operator's weekday dayshift routine	
Table 6.1	119
Mann-Whitney U Test results of people who responded to post-Simulator Study One survey	
Table 7.1	128
Five types of FROC visualization, with progressively increasing "data precision"	

Chapter One

Introduction

Process control in the chemical industry involves human operators who manage complex chemical and industrial processes in large plant facilities. Process controllers work in teams and shifts, supervising the plant and production process through displays and computer controls from a control room. Their responsibilities can be summarized under two general categories: **routine supervision** and **failure management**. During normal process control, controllers have to monitor system instruments and periodically manipulate control settings so as to maintain desired production quantities. When abnormal events occur, controllers have to troubleshoot quickly before production conditions deteriorate further, and yet symptoms to abnormal situations can be difficult to detect and accurate diagnosis can hard to perform, especially under time pressure (Lees & Sayers, 1976). Overall, the controllers aim to optimize production through effective process control, but managing these chemical plants has been described as “hours of boredom punctuated by a few minutes of pure hell” (Wickens & Kramer, 1985).

Process control is inherently challenging due to several issues (Wickens & Hollands, 2000):

1. Sluggish display and control, as the system may respond only after seconds or minutes have passed. Therefore the controllers must compensate for the time lags during the control of processes. There is also the complexity of deciding how long to leave the process alone before concluding on whether an action is necessary, or that the implemented control maneuver was appropriate or ineffective (Crossman & Cooke, 1974).
2. Continuous and analog production processes which must be controlled in a discrete and symbolic fashion. For example, inputs are set at specific levels rather than constantly being adjusted to the desired output. This creates a conflicting operational understanding and the possibility of creating deviations in an otherwise perfectly normal operation due to the “bang-bang” or impulse

control strategy often employed on systems with large lags (Young & Meiry, 1965; Gaines, 1969, Wickens, 1986)

3. A large number of interrelated variables, where changes in one variable may affect several other variables;
4. High risks and severe consequences during error events, which adds to the pressure that operators face when making decisions and actions in their effort to balance productivity against safety.

Each process plant is a complex system consisting of thousands of components and instruments. As there are many interactions among components, subsystems and instrumentations, it can be very difficult to derive an accurate and comprehensive understanding of the current status of the plant. Furthermore, the context on how these interrelated variables interact with one another changes frequently, as there will always be components missing, broken, or working imperfectly (Mumaw, Roth, Vincente, Burns, 2000). Despite these minor imperfections, the plant can still function safely due to redundancy. Controllers constantly have to work with alternative and changing mental models of the plant, which makes it challenging to distinguish between normal and abnormal situations and decide on the appropriate actions.

Controllers are aided by automation and computer systems which display various process information. The control task typically takes place in a remote control room. Since the system can sense what is going on, some of the control can be automatic, but some requires problem solving that must be handled by operators. Technology has come a long way, from the traditional analog panel board instruments to the Distributed Control Systems (DCS) that are widely used today. These automated systems facilitate controller performance in difficult process control tasks. As such, the DCS offers important areas for research to understand how control tasks should be divided between operators and automation, and how new types of displays can be designed that capture information that can improve controller performance.

1.1 THE DCS

Controllers generally rely substantially on the Distributed Control System (DCS) for process control. The DCS is a network-integrated computer system console, which provides operators with information regarding the production process as well as control abilities to manipulate process variables. Each DCS console has a set of displays that provides a variety of information that is relayed from various sensors located throughout the plant. This information can generally be presented using the following modes:

- Trends Display where the readings of sensors over time are plotted on a line graph
- Alarms Summary Display which logs alarms in chronological order
- Graphics (Schematic) Display which maps out the spatial location of various equipment as well as their health status and current sensor readings

Each of these presentation modes has their advantages and drawbacks. A Trends page (Figure 1.1) illustrates linear patterns of data readings over a period of time, but will typically not indicate alerts or provide spatial information about the equipment in the process system. The Alarms Summary Display (Figure 1.2) gives crucial audio-visual alerts on plant problems as well as a time-based listing of each warning, but it lacks spatial representation. A Graphics page (Figure 1.3) presents spatial information and current readings of process components, but typically does not provide any historical data or dynamic behavior indications. While these displays complement one another, the limited space on the DCS console means that operators must be proficient in selecting which of hundreds of data pages to monitor. The complexity of the process plant, with many different components and sensors, reflects a challenging environment which panel operators face as well as the limitations of the current DCS console. Understanding the current status of a process plant accurately and comprehensively is an arduous task for operators (Mumaw et al., 2000).

Fig. 1.1 A screenshot of a Trends display. Source: Honeywell Inc.

Fig. 1.2 A screenshot of an Alarm Summary display. Source: Honeywell Inc.

Since alarms are both vital and problematic the design of the alarm system has become a popular area for human factors and display interface research. research in alarms and failure management has been fueled by the severity of many critical accidents, including the Three Mile Island incident (U.S. Nuclear Regulatory Commission's Fact Sheet on Three Mile Island, 2008), and more recently the BP Texas City explosion (BP America, 2005). Although such devastating incidents are extremely rare, their consequences are threatening and provide a justification for extensive research in alarms during failure management. Errington et al. (2005) showed how a novel Human-centered interface better supported controllers in abnormal situation detection, diagnosis and response than the traditional interfaces. Reising and Montgomery (2005) analyzed 37 unique operator consoles to see if current alarm system performance guidelines were being achieved, especially during upsets and alarm flood situations. In-depth investigations have also been conducted to understand the mental workload, attention, and performance capabilities that controllers face during failures (Woods, 1995; Wickens & Hollands, 2000; Roth & Woods, 1998 to name a few).

While alarms are effective in *failure management*, they may be more of a bane than a boom for operators in *routine supervision* of plants. Shaw (1993) pointed out how controllers may choose to “**control by exception**”, where they do not continuously monitor the process but instead respond to alarms as and when they occur. As long as no alarms occur in any particular area, the controller would not monitor or devote active attention to that area. This would otherwise result in limited time for controllers to take remedial action when abnormal situations do occur. It is only a matter of minutes before conditions in process plants change from unhealthy to critical, and within this time horizon the controllers have to detect the problems, analyze the cause, and apply the appropriate correction. Such a **reactive approach** towards monitoring can be taxing for controllers, and a more efficient way would be to improve routine supervision would be to initiate a more **proactive approach**.

1.3 PROACTIVE MONITORING – REDUCING CRITICAL SITUATIONS

Considering the limitations of alarm systems for routine supervision, expert process controllers may opt to engage in proactive monitoring during normal plant operations (Mumaw et al., 2000; Meyer & Bitan, 2002; Shaw, 1993; Spengelink, 1990). Meyer & Bitan (2002) noted how “better operators receive worse warnings”, and that the diagnostic value of the alarm system decreases for expert operators, as they take preemptive actions to reduce the probability of abnormal situations. These controllers tend to detect anomalies by monitoring various displays for unusual trends, unexpected output positions, and other comparative differences. This allows them to detect changes prior to the automated alarm system as well as critical abnormal deviations that the alarm system might have overlooked.

Such early intervention can minimize economic losses, maintain plant production state, and reduce the occurrences of critical situations (Figure 1.4). Burns (2006) elaborated that proactive monitoring has three distinct phases: deviation detection, problem prediction, and performing compensatory actions. These phases reflect Rasmussen’s skill-, rule-, and knowledge-based behaviors (for more details on the SRK Model, see Rasmussen, 1983). Rather than analyzing the problem after it has happened, controllers who engage in proactive monitoring anticipate the future state of process variables and predict how they affect the plant’s overall performance. In this way, they are able to take action prior to any occurrence of critical problems.

Fig. 1.4 Early intervention can prevent catastrophic losses (Burns, 2006)

Proactive monitoring is not without its challenges. As plants become more complex and involve more automated systems, operators face increasingly complicated monitoring tasks which may lead them further “out of the loop” in terms of experience in managing new-age plants (Burns, 2006). In order for proactive monitoring to be effective, expert controllers need to make accurate **predictions** of future process state. However prediction requires much mental resources, which is partly why people tend to be more proactive in task management when workload is modest, and more reactive when workload becomes high (Hart & Wickens, 1990). Some operators may also simplify their tasks in general and reduce their workload, and thus avoid maintaining elaborate and optimal planning strategies for task management (Liao & Moray, 1993; Raby & Wickens, 1994).

1.4 PREDICTIVE DISPLAY TO AID PROACTIVE MONITORING

In light of controllers' need to predict the future, explicit **predictive displays** can be introduced to aid controllers in proactive monitoring. The ability to predict is particularly necessary but also challenging in controlling systems with high inertia and long lags (Wickens, Gempler & Morpew, 2000). This is further compounded by the fact that humans do not predict very precisely (Wickens & Hollands, 2000). Without a preview representation of the future, humans have to analyze the current state of the dynamic system and process the data through a mental model of the system, before visualizing a predicted scenario which he is to act upon. The presence of lagging and dynamic external variables further increases the mental workload in prediction. Predictive displays seek to reduce this workload and provide assistance for users so that they can make fair estimations regarding the future state of the system. Predictive displays are currently used in many applications and have shown significant advantages, such as in air traffic control (Wickens, Gempler & Morpew, 2000; Endsley et al., 1999) and supertankers (van Breda, 1999). Currently no explicit predictive displays exist for process control.

1.5 RESEARCH OBJECTIVES

As predictive displays facilitate proactive monitoring, the main goal of this dissertation project is to **develop a predictive display for process control so as to improve console operator's performance**. This goal is attained through the following objectives:

1. review current knowledge on human prediction;
2. investigate existing predictive displays from other domains;
3. examine current predictive applications in process control;
4. analyze some cognitive contributors as well as current tools that control operators rely on for proactive monitoring; and
5. explore the viability of a predictive display for process control to support proactive monitoring and anticipatory control.

1.6 OUTLINE OF THIS REPORT

- Chapter 2 reviews the psychology of prediction, and reports on the project's first qualitative investigation to explore process control operators' mental models and their roles in proactive monitoring.
- Chapter 3 covers issues that can be used in the engineering of prediction. It covers various predictive display technologies. Findings are then presented from a qualitative study of two operators' use of the trend displays in DCS systems.
- Chapter 4 explores current predictive applications in process control, and discusses the pros and cons of using different methods to calculate predictions.
- Chapter 5 compiles two qualitative investigations that looked at operators' cognitive processes as well as visual tools that support proactive monitoring.
- Chapter 6 reports on a Simulator Study One which was used to validate the operators' control performance benefits of integrating explicit predictive cues with trend displays.
- Chapter 7 documents Simulator Study Two, a laboratory experiment using an actual Honeywell ExperionTM system schematic display similar to those found in industries, and incorporated various rate-of-change visualizations that were operated using a single-variate rate-of-change algorithm.
- Chapter 8 concludes the project report by comparing overall objectives and research accomplishments.

1.7 REFERENCES

- BP America. *Isomerization Unit Explosion Final Report*. December 9th, 2005.
http://www.bp.com/liveassets/bp_internet/us/bp_us_english/STAGING/local_assets/downloads/t/final_report.pdf (May 26, 2008).
- Burns, C. M. (2006). Towards proactive monitoring in the petrochemical industry. *Safety Science*, 44, 27-36.
- Burns, C. M., Skraaning Jr., G., Jamieson, G., Lau, N., Kwok, J., Welch, R., Andresen, G. (2008). Evaluation of ecological interface design for nuclear process control: Situation awareness effects. *Human Factors*, 50, 663-679.
- Carvalho, P. V. R., dos Santos, I. L., Vidal, M. C. R. (2006). Safety implications of cultural and cognitive issues in nuclear power plant operation. *Applied Ergonomics*, 37, 211-223.
- Crossman, E. R. F. W. & Cooke, J. E. (1974). Manual control of slow-response systems. In E. Edwards & F. Lees (Eds.), *The Human Operator in Process Control*, London: Taylor & Francis.
- Endsley, M. R., Sollenberger, R., Stein, E. (1999). The use of predictive displays for aiding controller situation awareness. In *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting* (pp. 51-56). HFES: Santa Monica, CA.
- Errington, J. & Bullemer, P. (1998). Designing for Abnormal Situation Management, In *AIChE 1998 Ethylene Producers Conference*, New Orleans, LA.
- Errington, J., Reising, D. V., Bullemer, P., DeMaere, T., Coppard, D., Doe, K., Bloom, C. (2005) Establishing human performance improvements and economic benefit for a human-centered operator interface: An industrial evaluation. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*. HFES: Santa Monica, CA.
- Gaines, B. R. (1969). Linear and nonlinear models of the human operator. *International Journal of Man-machine Studies*, 1, 333-360.
- Hajdukiewicz, J. & Wu, P (2006). Beyond trends: A framework for mapping time-based requirements and display formats for process operations. Human Factors and Ergonomics Society Annual Meeting Proceedings 2007” Perception and Performance, 1885-1889.
- Hart, S. G. & Wickens, C. D. (1990). Workload assessment and prediction. In H. R. Boohar (Ed.), *MANPRINT: An emerging technology. Advanced concepts for integrating people, machines and organizations*. New York: Van Nostrand Reinhold.
- Jamieson, G. (2007). Ecological interface design for petrochemical process control: An empirical assessment. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 37, 906-920.

- Lees, F. P. & Sayers, B. (1976). The behavior of process monitors under emergency conditions. In T. Sheridan & G. Johanssen (Eds.), *Monitoring behavior and supervisory control*. New York: Plenum.
- Liao, J. & Moray, N. (1993). A simulation study of human performance deterioration and mental workload. *Le Travail humain*, 56, 321-344.
- Meyer, J. & Bitan, Y. (2002). Why better operators receive worse warnings. *Human Factors*, 44, 343-353.
- Mumaw, R. J., Roth, E. M., Vincente, K. J., Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors*, 42, 36-55.
- Nachreiner, F., Nickel, P., Meyer, I. (2006). Human factors in process control systems: The design of human-machine interfaces. *Safety Science*, 44, 5-26.
- Raby, M. & Wickens, C. D. (1994). Strategic workload management and decision biases in aviation. *The international Journal of Aviation Psychology*, 4, 211-240.
- Rasmussen, J., (1983). Skills, rules, knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, 13, 257-266.
- Reising, D.V. & Montgomery, T. (2005) Achieving effective alarm system performance: Results of ASM Consortium benchmarking against the EEMUA Guide for Alarm Systems. In *Proceedings of the 20th Annual CCPS International Conference*, Atlanta, GA.
- Roth, E. M., & Woods, D. D. (1988). Aiding human performance: I. Cognitive analysis. *Le Travail Humain*, 51, 39-64.
- Shaw, J. A. (1993). Distributed control systems: cause or cure of operator errors. *Reliability Engineering and System Safety*, 39, 263-271.
- Spengelink, G. P. J. (1990). Aiding the operator's anticipatory behavior: The design of process state information. *Applied Ergonomics*, 21, 199-206.
- Stanton, N. A. (1994). *Human Factors in Alarm Design*. London: Taylor & Francis.
- U. S. Nuclear Regulatory Commission. *Fact Sheet on the Three Mile Island Incident*. March 1994. <http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html> (May 26, 2008)
- van Breda, L. (1999). Anticipatory behaviour in supervisory control. Delft, The Netherlands: Delft University Press.
- Vicente, K. J., Roth, E. M., Mumaw, R. J. (2001). How do operators monitor a complex, dynamic work domain? The Impact of control room technology. *International Journal of Human-Computer Studies*, 54, 831-856.
- Wickens, C. D., Gempler, K., Morphew, M. E. (2000) Workload and reliability of predictor displays in aircraft traffic avoidance. *Transportation Human Factors*, 2, 99-126.

Wickens, C. D. & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd Ed.). Upper Saddle River, NJ: Prentice Hall

Woods, D. D. (1994). Cognitive demands and activities in dynamic fault management. In *Human Factors in Alarm Design*, pp. 63-92. London: Taylor & Francis.

Woods, D. D. (1995). The alarm problem and directed attention in dynamic fault management. *Ergonomics*, 38, 2371-2393.

Young, L. R. & Meiry, J. L. (1965). Bang bang aspects of manual control in high-order systems. *IEEE Transactions on Automatic Control*, 10, 336-341.

Chapter Two

Basic Theories of Cue-based Mental Prediction

2.1 OVERVIEW

Proactive monitoring involves, in part, the ability to predict and anticipate future events. The prediction that is dealt with here can be depicted as “bottom-up”, or **cue-based prediction**. This is different from “top-down” prediction, which depends mainly on past memories and regular patterns (e.g.: I know the Old Faithful Geyser will erupt about 90 minutes since its last eruption.). Bottom-up prediction involves the need to perceive and process cues before deriving a prediction. Such prediction could be associated with proactive monitoring, where operators have to make a prediction of the future plant state based on the available cues. Hence to facilitate proactive monitoring we need to understand the theories behind cue-based mental prediction. Four topic areas relevant to prediction are reviewed in this chapter: The Psychology of Prediction, Situation awareness, Mental Models and Mental Simulation, as well as Expertise.

Over many years, research on the **psychology of prediction** has revealed many characteristics about us humans performing predictive tasks. People rely on their ability to predict in order to perform many tasks, from household chores to professional decision-making and prognosis. Despite being a frequent activity, mentally simulating the future still remains effortful and is oftentimes inaccurate and ambiguous for many people. Cognitive trends and tendencies as well as factors affecting the performance of mental prediction are also discussed.

Situation awareness studies have been used as a basis for formalizing the process of *perceiving*, *understanding* and *projecting* dynamic situations that are experienced. Situation awareness can informally be described as “knowing what is currently happening”, and having sufficient situation awareness often contributes to good task performance. Of more significance to proactive monitoring is *Level 3 SA*, described as the mental projection of a status into the near future.

To predict the future, some researchers describe the process of performing *mental simulation*, where a pseudo-system similar to the physical, dynamic system can be operated mentally to derive future outputs. This **mental model**, along with situation awareness, would be essential for driving a **mental simulation** and acquiring an accurate prediction. Many different definitions of mental model exist, and they are briefly summarized in this section.

It is often said that experts are people who can anticipate future events and are able to react to these events in a timely manner. These experts are known to possess skilled intuition and are thus able to exhibit proactive behavior when performing their tasks. They seem to know where to look for cues, and are able to quickly interpret these cues and take decisive actions. They are fluent in both bottom up and top down prediction. The topic of **expertise** should reveal certain cognitive aspects of effective prediction.

2.2 THE PSYCHOLOGY OF PREDICTION

Prediction is a key element in many everyday activities, from figuring out how long it will take to use the toaster, to forecasting the weather (Doswell, 2004). Professionally, doctors provide prognoses of their patients' future health and human resource managers decide which applicant would be best for the job. People make predictions often, yet the success in accurate predictions is limited (Sherden, 1998; Kahneman & Lovallo, 1993; Dawes, Faust, Meehl, 1989; Wickens, 1986; Meehl, 1954). Humans in general are fairly good at estimating mean values of data sets, as well as making dichotomous decisions (Wickens & Hollands, 2000). It is also fairly easy to predict the future linear trajectory of an object in motion, such as an aircraft in stable flight. Prediction becomes harder when higher-order derivatives are needed to be factored in, such as when the aircraft is turning during flight. Without visual representation of the future flight path, it is difficult to extrapolate a non-linear trajectory of an aircraft. Similarly, humans are not good at extrapolating non-linear trend lines; typically they opt for a more linear estimate (Wickens & Hollands, 2000; Wagenaar & Sagaria, 1975; see Figure 2.1, although notably if we were not presented

with historical data, we would actually have a harder time noticing the increasing rate of growth until much later, a phenomenon closely-related to psychophysics' discrimination threshold, Fechner, 1860; Weber, 1846).

Fig 2.1 People tend to extrapolate linearly and fail in predicting exponential growth.

The basic strategy of prediction often involves building and processing a mental scenario of the current situation. Much research in cognitive and social psychology indicates that people tend to anticipate a future event by constructing suggestive mental scenarios. The simpler and the easier the scenario is constructed, and the more plausible the event, the more likely people will believe it will happen (Atance & O'Neil, 2001; Dougherty, Gettys, Ogden, 1999; Kahneman & Lovallo, 1993; Kahneman & Tversky, 1982). Research on planning fallacy done by Buehler, Griffin, Ross (1994) revealed that most students, when tasked to complete a project for class, would focus on creating scenarios of how the project would be accomplished—either by elaborating on the plans they had for getting the task done or by focusing on the obstacles that would lay in their way. Essentially, the act of prediction involves mental simulation—taking current conditions and building causal chains of events to extrapolate a possible outcome (Kahneman & Tversky, 1982; Lagnado & Sloman, 2004).

However, while scenario building is a viable strategy for deriving predictions, people tend to be less thorough or complete in the scenarios that they build. People often base their scenarios on a few abstract, higher-order features of the events rather than the concrete minute details (Rottenstreich & Tversky, 1997; Kruger & Evans, 2004; Tversky & Koehler, 2004). Jorgensen (2004) described how teams who considered all the subtasks required in completing a software development project were more realistic and accurate in their time-to-complete prediction than teams who based their judgments on key features such as number of user screens and interfaces. People commonly fail to thoroughly consider all cues and details when predicting which results in less-than-perfect predictions.

The temporal distance also affects the *resolution* in people's calculations, wherein participants reported more abstract descriptions and thought more simply about events in the future as compared to immediate events (Liberman & Trope, 1998). This tendency depends partly on the variability and dynamic characteristics of the anticipated event and when the further it is scheduled. Nussbaum et al. (2006) noted that people's confidence drop when relying solely on low-level information for predicting distant events. In their study participants were asked to predict a trivia quiz score which these students would take either an hour or a month later. The quiz would feature either of two question types: relatively difficult open-ended question or multiple-choice questions. When told that the questions were open-ended, students' confidence of success dropped if they were taking the quiz one hour later, but not if the quiz was administered one month later. Participants believed that if given sufficient time in the future, performance capabilities may change and thus mitigate the effects of low-level factors such as question format. This drop in confidence is less evident in experts such as pilots (Sulis, Wickens, Chui, 2011), who may instead exhibit signs of overconfidence. High temporal span-of-prediction increases the influence of high-level information (e.g.: fundamental rules and theories, or in the previous example, trivia knowledge) and decreases the impact of low-level information (e.g.: noise, question format) on prediction (cf. Wickens, 1986, page 33, on useful span of prediction for various dynamic systems such as a small plane or an oil tanker).

During mental simulations, people often focus on one event and neglect alternative outcomes (Redelmeier, Koehler, Liberman, Tversky, 1995; Snyder & Swann, 1978) if they were prompted to consider alternatives their predictions would be more conservative and realistic. Dunning & Parpal (1989) polled students on the impact of prepared class notes on their potential course grades, and students responded that the notes would help significantly. However when students were asked whether the lack of these notes would hurt their grade, thereby asked to simulate the alternative of not having notes, students stated that the notes would not make much of a difference. Furthermore, Griffin & Tversky (1992) highlighted how people, while gathering evidence to support scenario development, also tend to over-emphasize the *strength* of the evidence (how much the evidence suggested a particular outcome over another) and under-emphasize on its *weight* (whether the evidence was valid and reliable).

Dunning (2007) asserted that people could make more accurate predictions if they paid more attention to a data-driven, bottom-up approach of prediction, also known as the *outside view* (versus *inside view*) of prediction. Taking an outside view would mean recognizing that a particular situation is an example of a category of similar past events, which can be additionally surveyed to develop outcome predictions (Lagnado & Sloman, 2004; Jorgensen, 2004). Research has shown that adopting the outside view allowed for more accurate predictions. Buehler et al. (1994) demonstrated how participants, when asked to recall similar past assignments that they had completed, were less likely to underestimate the time needed to complete a comparable assignment than participants who did not perform the recall. Individuals would also achieve higher prediction accuracy and avoid overconfidence when they use base-rate information if available (Dunning & Story, 1991).

2.3 INDUCTIVE & DEDUCTIVE REASONING

Given the focus on bottom-up prediction, a review on how inductive and deductive reasoning may shed some insights on how predictions (which are similar to conclusive reasoning) are derived. Fundamentally, people have to approaches to reasoning (Rips, 2001): Deductive reasoning works from the more general to the more specific through having a perceived general theory first, then observe for cues and

signs before confirming or rejecting this perceived theory; Inductive reasoning works in the opposite order, where cues and signs are observed for patterns before coming up with a general theory of the situation. Adapting to the context of prediction, *deductive prediction* first begins with a pre-conceived possibility or hypothesis of the future state, and in a hypothesis-testing fashion seek out information to either support or reject this prediction (Romeyn, 2004). This pre-conception may be due to a foundational understanding of factual knowledge, formal rules, or *mental models* similar to deductive reasoning (Johnson-Laird, 1999). In the case of process control prediction, it might involve careful study of the dynamics of the plant from operational manuals, noting where long time constants occur. Conversely, *inductive prediction* utilizes available information to derive trends and patterns so as to construct a prediction of the future (Rescher, 1999; Coffa, 1968). Here the operator may simply learn through experience (induce) which variables provide the best prediction of future states. Despite the process differences in deriving prediction, both approaches may still be subject to error.

Common cognitive constructs can be identified in both inductive and deductive predictions. Deductive predictions need a fundamental *mental model* of the situation in order to derive the initial guesstimate of the future. Bits and pieces of information are pieced together based on this mental model to allow for *mental simulation* to support the initial prediction. In the context of bottom-up, cue-based predictions, both types of predictions share the common need to rely on perceived cues, although both processes utilize the cues for different purposes (Induction perceives patterns to infer a prediction, deduction confirms initial hypothesized prediction). Cue perception reflects the importance of *situation awareness*, in particular Level 1 Situation Awareness: Perception. Going deeper into these constructs should reveal components and relationships within cue-based predictions.

2.4 SITUATION AWARENESS

Endsley (1995) defined situation awareness (SA) as the “perception of the elements in the environment within a volume of time and space (Level 1), the comprehension of their meaning (Level 2), and the projection of their status in the near future (Level 3)”. In the context of process control, the operator may first notice that the DCS indicates a high product temperature as the product flows into the distillation tower (Level 1). He understands that this means a failure in the cooling fans upstream (Level 2), and anticipates that the temperature in the distillation tower may soon reach a critical and unsafe level if this situation carries on (Level 3). Hogg et al. (1995) noted that understanding and projecting the future system state (Level 3 SA) is the most critical aspect of operator SA in a nuclear power plant operation. Figure 2.2 shows Endsley’s model of situation awareness. SA often applies in dynamic situations in which operators are monitoring or controlling, and is different from declarative or procedural knowledge about the situation or system, the latter characterizing long-term memory (Endsley, 1995; Adams & Pew, 1990; Durso, Rawson, Giroto, 2007; Durso & Sethumadhavan, 2008).

Fig 2.2 Model of situation awareness in dynamic decision making (Endsley, 1995)

According to Figure 2.2, situation awareness progresses in a linear fashion, as illustrated in Table 2.1. That is, a failure in Level 1 SA would probably mean a failure in Level 2 and Level 3 SA to the extent that prediction is bottom up. In many domains the leading cause of SA-related errors is also due to the failure of Level 1 SA (Jones & Endsley, 1996; Endsley & Rodgers, 1998; Langham, Hole, Edwards, & O'Neill, 2002). Failure of Level 1 SA may stem from poor salience or legibility of critical signals, as well as perception difficulty due to individual or environmental factors (e.g.: not perceiving the red light while driving towards the junction). Comparatively, Level 2 SA errors are often a result of improper integration or recognition of perceived data, or a wrong selection of mental model used for generating the situation assessment (e.g.: successive symptoms point to a new diagnosis, but the doctor ignores or interprets them inappropriately), while Level 3 SA errors are more prevalent when highly developed mental models, attention, and working memory capacity are lacking. Wickens (2008) noted that a breakdown in each would require different solutions in addressing them. A breakdown in Level 1 SA would prompt the design of better alerts, whereas a breakdown of Level 3 SA might mean the incorporation of predictive displays.

Table 2.1 Presence of SA levels while achieving perception, comprehension and projection.

To Achieve	Presence of SA Levels		
	Level 1	Level 2	Level 3
Perception	Yes	<i>no</i>	<i>no</i>
Comprehension	Yes	Yes	<i>no</i>
Projection	Yes	Yes	Yes

Through understanding the nature of SA errors, it can be seen that while Level 2 SA would typically be present during Level 3 SA, comprehending the situation is not always a key requirement in producing a mental projection of future state. A plant operator may know that pressure is increasing and hence the future assessment is for a likely explosion, but he may not know the source (i.e.: Level 2 SA) of the pressure blockage. Indeed this is why current procedures in nuclear power dictate assuring safety of the plant (e.g.: preventing a future catastrophe) before diagnosing the root cause.

As a cognitive concept, situation awareness also has relations to various other cognitive elements involved in human performance. Critical cognitive mechanisms that support SA include attention, working memory, and long-term memory (LTM) (Endsley, 1995; Lundberg, 1999; Wickens & Hollands, 2000), while some scientists highlight the importance of long-term working memory (LTWM) as well (Durso & Gronlund, 1999; Wickens, 2000; Ericsson & Kintch, 1995). Unlike working memory, which is processed in the order of seconds or minutes, LTM operates temporally in the order of hours, days and years. As such, LTWN is significant for SA as it allows the ability to rapidly store information in LTM and thus reduce working memory load. Given the relationship between working memory and SA, mental workload would thus also have an impact on a person's ability to maintain accurate situation awareness. Wickens (2002) noted that as people become busier with task management, the increase in mental workload due to factors like information load and time pressure would suppress the ability to maintain effective situation awareness. Naturally, domain experts known to be more capable at managing tasks were also found to have stronger situation awareness (Endsley, 2006; Randel, Pugh, Reed, 1996; Jodlowski, Doane, Sohn, 2002).

A main role-player of SA inside the LTM is the *mental model* of the state or system. The mental model is essential in achieving higher levels of SA (comprehension and projection) through its integration with recognized critical features in the environment to generate what Rouse & Morris (1985) would describe as “descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future states”. The combination of cue perception (Level 1 SA) and mental model would allow for a *mental simulation* to create two possible products: a *situational assessment* (Fracker, 1988, also known as a *situation model*) which facilitates comprehension of current situation (Durso, Rawson, Giroto, 2007); as well as a *prediction* or *projection of future state* over a “look-ahead” time span (Wickens, Gempler, Morphew, 2000; Endsley, 2006).

2.5 MENTAL MODEL

As featured in the previous two sections, the operation of a mental simulation requires fundamentally a mental model. The concept of mental models has many definitions (Wilson & Rutherford, 1989). Toffler (1970) defined mental model simply as a “subjective representation of external reality”, while Rasmussen (1979) elaborated on two forms of mental model: first a spatial, cognitive mapping of components, second a formal understanding of variables and relationships used to process data. Rasmussen (1986) also believed that it is not necessary to have detailed models of the actual processes, just higher-level structural models sufficient to engage mental activities, while Bainbridge (1988) further described these models as the “background knowledge” that which users often refer back to during cognitive processes. Various domain-specific references to mental models can be found in the literature: users’ conceptual models (Moran, 1981), conceptualizations (Baggett & Ehrenfeucht, 1988) and device models (Kieras & Bovair, 1984) to mention a few. . Applications of mental models often relate the notion of a schemata (Mayer, 1983), which is a memory-based knowledge structure or cognitive representation of a particular domain. The association between mental model and schemata is echoed by Johnson-Laird (1983), Jones (1987) and Rumelhart (1984). It is generally understood that in complex human-machine systems, operators possess and maintain a mental representation of the systems they are managing in order to perform cognitive activities and control tasks.

However, given how mental models are acquired, to describe mental models as various forms of memory-based visualizations may sound over-simplistic. Particularly in the context of process control, research has shown that having a general understanding through rote learning of a system model do not facilitate the operator’s ability to control the system (Crossman & Cooke, 1962; Kragt & Lamdeweerd, 1974; Landeweerd, Seegers, Praageman, 1981). Instead, there is a need to understand the dynamic relationships and underlying processes within the system rather than just knowing the system’s structure of causal sequencing (Attwood, 1970; Wickens & Hollands, 2000), which is thus best acquired through actual interaction with the system and experiences of incidents (De Keyser, 1988). Notably, this method of understanding allowed operators to act more proactively and anticipate events in the system

(Brigham & Liaos, 1975). This is a key consideration for this project as it will affect how experiment participants, namely undergraduates, should be trained. As such, an investigation is needed to better understand how process control operators' mental models are derived, updated, and applied during the course of their work.

While the absolute definition of mental model is debatable, the functions of the mental model are more certain. Refined mental models help guide the attention of pilots during visual scanning (Bellenkes, Wickens, Kramer, 1996). Prior understanding of the device model (how the device works in terms of its internal structure and processes) would facilitate learning, retention and execution of operation procedures (Kieras & Bovair, 1984). Holland, Holyoak, Nisbett and Thagard (1986) reported how "default values" from the mental model can be used in place of unknown current values of the operating system, thus allowing people to operate effectively even when provided with limited information. From a spatial context, people try to orientate themselves in a foreign environment (or in the case of space, foreign visual perspective of the same environment) by establishing landmarks and reference points in the environment, and in essence perform "dead-reckoning" using their mental models in their heads (Oman, Shebilske, Richards, Tubre, Beall, Natapoff, 2000; Vidal, Amorim, , Berthoz, 2004; Tversky, 1993). The mental model, as a stable representation of the system, supports troubleshooting and problem-solving, even if the problems encountered are novel. The mental model acts as a basic frame for pattern recognition, or at the very least serves as a guide for identifying critical cues to monitor

Mental models also allow for the development of expectations, which help drive the deployment of attention and expedite predictions (Moray, Lootsteen, Pajak, 1986; Moray, 1997). Operators are constantly monitoring the environment for cues to determine the state of the system. Given the vast amount of information and the limited supply of attention, operators use mental models to direct their attention toward critical cues. Based on the mental model, an air traffic controller can *expect* the kinds of activity that occurs in his low-altitude sectors, such as knowing where to expect receiving aircrafts from other controllers and when aircrafts appear to be deviating from flight plans (Durso & Dattel, 2006). It can be said that *Level 1 SA: Perception* is achieved through using a well-developed mental model for dynamic

direction of attention to critical cues (Endsley, 1995). While mental models generate expectations that tell operators what (or what not) to look for, they are also an important component that strategizes what plans of action should be taken. Although industrial process control are almost always a slow, closed-loop process, highly-skilled operators tend to use a discrete open-loop management strategy (Crossman & Cooke, 1962; McLeod, 1976). It is this mental model that allows these operators to mentally simulate and derive an appropriate plan of action. Without this mental model which expert operators evidently possesses, the operator would end up engaging in slow, inefficient closed-loop control of initiating an input, wait to see what happens, and then respond again (Wickens & Hollands, 2000).

People often use mental models of a dynamic system for *mental simulation* to predict the future state and understand how it will change from the current state (Klein & Crandall, 1995). Through this cognitive process, users can develop expectancies of how the system should be behaving. Klein (1999, Chapter 5) described how experts would frequently perform mental simulation so as to visualize and understand what was currently happening or would be happening in the near future. On a similar note, Johnson-Laird (1983) stated that mental models “enable individuals to make inferences and predictions, to understand phenomena, to decide what action to take and to control its execution, and above all to *experience events by proxy*”. The concept of *experiencing events by proxy*, or more aptly *mental simulation*, is repeated in many studies regarding decision-making and prediction (e.g.: Toffler, 1970, Rasmussen, 1986, Kieras & Bovair, 1984). This cognitive process can be used to explain a current situation (by visualizing how the past state arrived at the present state), to predict what is going to happen and anticipate for it (by visualizing along with the current cues how the present state would play out into the future), as well as to evaluate a potential course of action to find out if it has any flaws.

2.6 MENTAL SIMULATION

Like mental models, the idea of experiencing events by proxy has been well recognized in the literature. Kahneman & Tversky (1982) described a “simulation heuristic” that involved running through alternatives of a situation so as to determine how to react. De Groot (1965), in his research on chess players, noted a phenomenon he called “progressive deepening” where chess players mentally searched a decision tree deeply rather than broadly in order to derive options and countermoves. In a study of jurors, Pennington & Hastie (1993) described jurors as trying to build a story that would best match with the evidence and how they expected people behaved. This “story model” involved deriving a simulation that “best matched” the evidence given so as to conclude what actually happened. Regardless of the various definitions and contextual applications, in general mental simulations can be deemed as imaginative cognitive constructions and reconstructions of events, either to understand the past, assess the present, or predict the future (Sanna, Stocker, Clarke, 2003; Taylor, Pham, Rivkin, Armor, 1998).

In prediction and forecasting, mental simulation has been described in terms of “upward” and “downward” deviation from actual (Carroll & Shepperd, 2009; Roese, 1994; Markman, Gavanski, Sherman, McMullen, 1993). The structure of mental simulation typically follows that of conditional propositions, involving both an antecedent and a consequence. An example of a conditional proposition from a pilot’s context may be “if I fly through that thick, dark cloud, there might be strong updrafts and downdrafts within the cloud circulation, and we might experience turbulence”. Given the same example, a downward simulation would change the antecedent to arrive at a worse outcome than expected, such as “but if I fly around the cloud, I may not know where it ends and it might expend my fuel”. Conversely, an upward simulation would change the antecedent to arrive at a better outcome than expected, like “if I fly up and over the cloud, I should be safe”. Typically, upward simulations are most useful for future preparation and deciding best action (Sanna, Chang, Meier, 2001; Carroll & Shepperd, 2009), although both upward and downward simulations are not mutually exclusive during each instance of mental processing.

More importantly, mental simulation supports the forming of expectations that enable people to prepare for the future. Mental simulations allow people to “travel back in time” to past episodes and then project these episodes forward in time to simulate and prepare for possible future situations (Roberts, 2002; Gollwitzer & Kinney, 1989). From a reverse fashion, given the current situation state people can compare what the future may be, and set them up towards improving their present position in order to achieve their future goal (Oettingen, Pak, Schnetter, 2001). In a very hypothetical situation, Wendy envisioned that her future lifestyle with her current boyfriend Jake may not be comfortable and pleasant, and thus decided to ditch him now. With downward mental simulations, expectations of the future are more realistic (perhaps less optimistic), such as college seniors who are soon graduating and reduce their expected salary as compared to college sophomores (Shepperd, Ouellette, Fernandez, 1996). Mental simulations represent the mechanism, both in terms of upward as well as downward simulations, by which people generate future outlooks as well as revisions of these outlooks (Carroll & Shepperd, 2009).

Of greater interest in this project is the use of mental simulation in the context of dynamic, naturalistic decision-making. Notably, Klein & Crandall (1992) explored the role of mental simulation within the Recognition-primed Decision-making Model (RPD). This model asserts that people use situation assessment to decide the best course of action and use mental simulation to evaluate this course of action. Klein and Crandall’s definition of mental simulation differed from the rest, since they considered the rich and diverse domain-specific associations that people use when making decisions beyond just “running the model” in the head.

There are four primary functions of mental simulation in naturalistic decision-making:

1. To *create a planned course of action*, allowing the decision maker to anticipate the “look and feel” of ensuing events and to adequately prepare for them.
2. To allow decision makers to *evaluate the potential course of action*, through experiencing it by proxy to answer questions like “will it work?” or “what could go wrong?”
3. To help the decision maker to *understand the situation* through reexamining details back in time.
4. To *deepen and expand* the decision maker’s comprehension of situations, to offer the ability to create and operate hypothetical models, to mentally “observe” these models in action.

However, Klein & Crandall (1995) also noted that mental simulation is vulnerable to time pressure and expertise. Great time pressure can limit the extent of mental simulation, thus reducing the ability to perform all of its functions. Aside from time constraints, the individual’s experience level also plays a role in effective mental simulation. Sufficient task experience as well as domain knowledge are required to form the building blocks with which to assemble an adequate mental simulation. Domain experts seem to have no difficulties specifying key parameters and constructing action sequences to generate reliable predictions. Expertise appears to be an important factor in deriving predictions.

2.7 EXPERTISE

Experts are often known for their comprehensive “mental model” of the system as well as their efficiency in task performance, troubleshooting and making predictions. Experts may identify and use large, meaningful patterns, utilize effective strategies for problem-solving and decision-making, handle adversities better, or be simply more skilled, quicker and possess superior memory as compared to non-experts (Shanteau, 1992; Glaser & Chi, 1988). In order to type fast, expert typists rely on the ability to look ahead in the text so as to identify, process and *anticipate* letters to be typed

(Salthouse, 1984). Skilled tennis players with much tournament experience are able to *anticipate* where an opponent's shots will land, even before the opponent's racquet has contacted the ball (Williams, Ward, Knowles, Smeeton, 2002). Novice players focus their attention mainly on the opponent's tennis racket, but experts are able to pick up subtle, early movement cues of the opponent. . The expertise to intuitively know where to look (bottom up processing) for informative cues as well as employ effective strategies are important to most skilled actions, including: playing soccer (Williams & Davids, 1998), car driving (Endsley, 2006), and aviation pilots (Bellenkes, Wickens, Kramer, 1997). Naturally, experts also tend to report more perceivable cues than novices, and thus more likely to derive appropriate strategies in response to the current situation (Fowlkes, Salas, Baker, Cannon-Bowers, Stout, 2000).

Aside from utilizing effective strategies, experts may also possess vast knowledge of situational patterns and past experiences (i.e.: knowledge for top-down prediction). In chess, acquired patterns rather than innate abilities account for the skill differences between novice and master players (Chase & Simon, 1973). Research in naturalistic decision-making has showed how experts learn from identifying patterns and background knowledge to swiftly interpret not just informational cues, but also cue configurations and structural relationships in the dynamic environment. Serfaty et al. (1997) studied how expert battle commanders would come up with an initial plan by first generating a mental model of the current situation and recognizing potential patterns and solutions. They would then interact with the situation to garner more information and come up with an "improved" mental model, which they will then use to visualize more effective plans. Some experts are able to recognize the scenario and make sound decisions quickly, and Klein (1989) described these experts as possessing the ability to perform *Recognition-Primed Decisions* (RPD).

Experts appear to make quick and effective decisions based on their highly-skilled "intuition". Simon (1992) characterized intuition as "nothing more and nothing less than recognition" of cues in the environment. According to Kahneman & Klein (2009), to develop skilled intuition, the environment must first possess adequately valid cues to the nature of the situation. These cues should have sufficient regularity, and that the rules of interpreting these cues should remain as consistent as possible.

With these factors in place, the individual would then need ample time to learn and practice the rules of interpreting these cues. Frequent practice leads to automated behavior, as highlighted by Rasmussen's work (1983) in which rule-based behavior involves quick and effortless cognitive processes to match situations with pre-defined rules and patterns.

Conversely, failed or imperfect intuition can often be attributed to the lack of valid cue perception, inadequate knowledge of the rules, or the negative effects of heuristics and biases. Even experts make mistakes, such as bushfire-fighters who, when left alone to their own devices to predict the spread of fire, failed to consider the wind direction and slope angle present in the physical environment. Lewandowsky et al (1997) conducted this study to find that only the group of firefighters who were given a visual model about the environment did correctly integrate the wind direction and slope angle. Even if firefighters possessed a well-developed mental model, appropriate cue perception must be perceived in order to achieve reliable prediction.

Thus there are times when computers would predict better than humans. Flawed intuitive judgments are often attributed to the effects of heuristics and biases, such as when applying overly-simplified heuristics onto complicated situations (Kahneman & Frederick, 2002). The environment plays a role in hindering the performance of intuitive judgments too. The environment can be insufficiently predictable, in which available cues are weak and uncertain, the rules of interpreting them are inconsistent, and opportunities to learn these rules are lacking. In such situations, algorithm-based predictions are more advantageous than manual predictions, because at least algorithms are more consistent than humans. Statistical analysis is more likely to identify and consistently use weakly valid cues, and thus is better than humans at sustaining above-chance accuracy (Karelaia & Hogarth, 2008). Even in highly predictable environments, algorithms perform better than humans as algorithms do not suffer from occasional lapses in attention.

2.8 GRAPHICAL SUMMARY OF LITERATURE REVIEW

Reviewing these topics of Psychology of Prediction, Situation awareness, Mental Models and Mental Simulation, as well as Expertise, it comes as no surprise how these topics are inter-related. For an expert to develop skilled intuition, the individual would have to be in an environment where valid cues are sufficiently salient (Level 1 SA), and with consistent rules of cue interpretation the individual would thus be able to learn cue recognition over time. Experts also run a mental simulation to generate a predicted future state, and through their skills as well as experiences allow for faster simulation performance. To perform mental simulation, a mental model of the system would be required, with information perceived from the environment to construct scenarios that which may be incomplete. Naturally, Level 1 SA (perceiving information) is essential for achieving both Level 2 (comprehending the situation) and more importantly Level 3 SA (predicting future state), although, as noted before, levels 2 and 3 are not necessarily sequentially linked. All these are summarized and represented in Figure 2.3, as described further below.

Fig 2.3 Graphical representation of the psychology of cue-based prediction.

Given our understanding of SA, cue processing is necessary in order to achieve higher SA levels. Cue perception is partly driven by the user's mental model of the system, which directs the user's attention onto crucial information found in the environment, as well as past or similar experiences that the user may have. Through the interaction between the mental model and perceived cues, the user is able to generate a mental simulation to derive a comprehension of the current state. In order to come up with a prediction, both the cues and mental model in the simulation would have to be processed along with a temporal factor, the amount of "look-ahead time". The further this time span is into the future, the higher the effort required (Sulis, Wickens, Chui, 2011), the higher the tendency to rely on abstract information, and the lower is our confidence and accuracy in deriving an accurate and detailed prediction. This process of deriving situation comprehension and prediction through the reliance of the user's inherent mental model parallels knowledge-based behavior in novel situations illustrated by Rasmussen's SRK Model (1983). With expertise, it is also possible to bypass this simulation process and understand the current situation / predict future state through *Recognition-Primed Decision-making* or RPD (which is to say that the mental simulation process is really implicit and automatic). Lastly, results from mental comprehension or prediction can help drive additional cue searches, as in the example of a doctor looking for confirming cues for an earlier diagnosis.

The model is able to systematically explain the success of cue-based predictions. In the context of air traffic control, a controller monitoring a radar display pays particular attention to a pair of converging aircrafts based on his mental understanding of aviation as well as past experiences. He notices a pair, and with an inherent understanding of how commercial aircrafts function, he looks at additional relevant cues such as airspeed, altitude changes, travel intent etc. Mentally working these information together in his head allows him to comprehend the current states of the two aircrafts, but he is more concerned about whether they will eventually be in close proximity of each other. He mentally visualizes the flight paths of both aircrafts over a period of time and *predicts* that a possible proximity breach. The controller then initiates appropriate actions, and monitors the environment for feedback. Experts have the added advantage of being "recognition-primed", and that they are quick to associate perceived cues with certain diagnoses or predictions (rule-based behavior).

The model also maps the various factors that cause prediction failures. As predicted by situation awareness models, a failure in perceiving critical cues would lead to subsequent failures in situation comprehension and projection, as is the case of pilots who failed to notice the disengaged autopilot and wrongly assumed the aircraft's mode of control. While novices at times are able to perceive cues as efficiently as experts, they may not have a well-developed mental model to understand the situation or predict what will happen next, and naturally their mental simulation would also be flawed. The interaction between the perceived cues and the user's mental model may drive the user to search for additional cues in the wrong areas of interest, eventually deriving an erroneous conclusion. As noted by Doane et al. (2004), although both novice and expert pilots have easy access to visual cues and well-developed mental models of aircraft performances, novices tend to fair worse in deriving situation comprehension and predictions. A failure in the feedback loop can be seen in examples when cue perception is hindered by applying inappropriate experiences or situation comprehension, such as a doctor wrongfully searching for allergy symptoms in a case of an asthma attack. Lastly, although experts are capable of performing recognition-primed decision-making, they may occasionally apply an improper recognition rule to the situation.

Synthesizing various academic sources, accurate cue-based predictions require an awareness of the relevant cues and a reliable mental model of the process or situation. Deriving a prediction may involve mental simulation, an implicit process that takes effort to be made consciously explicit through methods such as think-aloud or cognitive mapping. Experts can predict swiftly through bypassing mental simulation via means of recognition-primed decision-making. The graphical model summarizes these key aspects of prediction as adapted from various literature in human factors and psychology, ranging from psychological studies in prediction and human intuition, to situation awareness, mental models, and expertise. It serves as a guide to indicate the factors that influence effective prediction, thus providing the direction for subsequent approaches in this project. Given the importance of perceiving cues, the next chapter looks at the various predictive tools found in other domains.

This work focused mainly on “bottom-up” cue-based predictions and to a lesser extent “top-down” expectancy-driven predictions. The nature of process control tasks requires operators to make judgments and predictions based on the information that’s being presented. Undeniably, people do make predictions based on their own expectations of the situation too. Kahneman & Tversky (1973) described such predictions as “intuitive”, and noted that these predictions oftentimes ignore statistical logic and reliable evidence, even when these information ran against their intuitive expectations. Many of the popular decision-making heuristics proposed by the duo (Tversky & Kahneman, 1974; Kahneman, Slovic, Tversky, 1982), such as overconfidence bias and representativeness heuristics, can be used to describe human’s behavior in top-down predictions. However the approaches to improving top-down and bottom-up predictions can be quite different. De-biasing techniques are usually educated to users through training, whereas we see more potential engineering technological solutions that facilitate cue-detection for bottom-up prediction.

2.9 REFERENCES

- Adams, M. J., & Pew, R. W. (1990). Situational awareness in the commercial aircraft cockpit: A cognitive perspective. In *Proceedings of the 9th Digital Avionics Systems Conference* (pp. 519-524). Virginia Beach, VA: IEEE.
- Atance, C. M. & O'Neill, D. K. (2001). Episodic future thinking. *Trends in Cognitive Science*, 5, 533-539.
- Attwood, D. D. (1970). The interaction between human and automatic control. In F. Bolam (Ed.), *Paper making systems and their control*. London: British Paper and Board Makers Association.
- Baggett, P. & Ehrenfeucht, A. (1988). Conceptualizing in assembly tasks. *Human Factors*, 30, 269-284.
- Bainbridge, L. (1988). Types of representation. In L. P. Goodstein, H. P. Andersen, H. E. Olsen (Eds.), *Tasks, Errors, and Mental Models*. London: Taylor & Francis.
- Bellenkes, A., Wickens, C.D., & Kramer, A. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space and Environmental Medicine*, 68, 569-579.
- Brigham, F. R. & Liaos, L. (1975). Operator performance in the control of a laboratory process plant. *Ergonomics*, 29, 181-201.
- Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67, 366-381.
- Carroll, P. J. & Shepperd, J. A. (2009). Preparedness, mental simulations, & future outlooks. In K. Markman, W. Klein & J. Suhr (Eds.), *The Handbook of Imagination and Mental Simulation*, 429-445. New York: Psychology Press: Taylor & Francis.
- Chase, W. G. & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual Information Processing*. New York: Academic Press.
- Coffa, J. A. (1968). Deductive predictions. *Philosophy of Science*, 35, 279-283.
- Crossman, E. R. F. W. & Cooke, J. E. (1962). *Manual Control of Slow Response Systems*. Presented at the International Congress on Human Factors in Electronics, Long Beach, CA.
- Dawes, R. M., Faust, D., Meehl, P. E. (1989). Clinical versus actuarial judgement. *Science*, 243, 1668-1674.
- de Groot, A. D. (1965). *Thought and choice in chess*. New York: Mouton.
- de Keyser, V. (1988). How can computer-based visual displays aid operators? In E. Hollnagel, G. Mancini, D. D. Woods (Eds.), *Cognitive Engineering in Complex Dynamic Worlds*. London: Academic Press.
- Doswell, Charles A., 2004: Weather Forecasting by Humans—Heuristics and Decision Making. *Weather Forecasting*, 19, 1115–1126.

- Dougherty, M. R. P., Gettys, C. F., Odgen, E. E. (1999). MINERVA-DM: A memory processes model of judgments of likelihood. *Psychological Review*, 106, 180-209.
- Dunning, D. (2007). Prediction: The inside view. In A. Kruglanski & E. Higgins (Eds.), *Social Psychology: Handbook of Basic Principles*, 2nd ed., New York: Guilford.
- Dunning, D. & Story, A. L. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? *Journal of Personality and Social Psychology*, 61, 521-532.
- Durso, F. T. & Dattel, A. R. (2006). Expertise in transportation. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, Cambridge University Press.
- Durso, F. T., & Gronlund, S. D. (1999). Situation Awareness. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay & M. T. H. Chi (Eds.), *Handbook of Applied Cognition* (pp. 283-314). Chichester: John Wiley.
- Durso, F., Rawson, K., Giroto, S. (2007). Comprehension and situation awareness. In F. Durso, R. Nickerson, S. Dumais, S. Lewandowsky, & T. Perfect (eds.), *Handbook of Applied Cognition*, 2nd Ed. Hoboken, NJ: Wiley.
- Durso, F. T., & Sethumadhavan, A. (2008). Situation awareness: Understanding dynamic environments. *Human Factors*, 50, 442-448
- Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37, 32-64.
- Endsley, M. R. (2000). *Situation models: an avenue to the modeling of mental models*. Paper presented at the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society, Santa Monica, CA.
- Endsley, M. R. (2006). Expertise and situation awareness. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance*. Cambridge: Cambridge University Press.
- Endsley, M. R. (2006). Situation awareness. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (3rd ed., pp. 528-542). Hoboken, NJ: John Wiley & Sons, Inc.
- Endsley, M. R. & Rodgers, M. D. (1998). Distribution of attention, situation awareness and workload in a passive air traffic control task: Implications for operational errors and automation. *Air Traffic Control Quarterly*, 6, 21-44.
- Ericsson, A., & Kintch, W. (1995). Long term working memory. *Psychological Review*, 102, 211-245.
- Fechner, G.T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Härtel, 2, 559
- Fowlkes, J. E., Salas, E., Baker, D. P., Cannon-Bowers, J. A., Stout, R. J. (2000). The utility of event-based knowledge elicitation. *Human Factors*, 42, 24-35.
- Fracker, M. L. (1988). A theory of situation awareness: Implications for measuring situation awareness. In *Proceedings of the Human Factors Society 32nd Annual Meeting*. Santa Monica, CA: Human Factors Society.

- Glaser, R. & Chi, M. (1988). Overview. In M. Chi, R. Glaser, M. Farr (Eds.), *The Nature of Expertise*. Hillsdale, NJ: Erlbaum.
- Gollwitzer, P. M. & Kinney, R. F. (1989). Effects of deliberative and implemental mind-sets on illusions of control. *Journal of Personality and Social Psychology*, 64, 552-560.
- Hogg, D. N., Folleso, K., Strand-Volden, F., & Torralba, B. (1995). Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. *Ergonomics*, 38(11), 2394-2413.
- Holland, J. H., Holyoak, K. F., Nisbett, R. E., Thagard, P. R. (1986). *Induction: Processes of Inference, Learning and Discovery*. Cambridge: MIT Press.
- Jodlowski, M. T., Doane, S. M., Sohn, Y W. (2002) Mental models, situation models, and expertise in flight situation awareness. In *Proceedings of the Human Factors Society 46th Annual Meeting*. Santa Monica, CA: Human Factors Society.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual Review of Psychology*, 50, 109-135.
- Jones, D. G. & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmental Medicine*, 67(6), 507-512.
- Jorgensen, M. (2004). Top-down and bottom-up expert estimation of software development effort. *Information and Software Technology*, 46, 3-16.
- Kahneman, D. & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*. New York: Cambridge University Press.
- Kahneman, D & Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree. *The American Psychologist*, 64, 515-26.
- Kahneman, D. & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39 17-31.
- Kahneman, D., Paul S., Amos T. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychology Review*, 80, 237-251.
- Kahneman, D. & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- Kieras, D. & Bovair, S. (1984). The role of mental models in learning to operate a device. *Cognitive Science*, 8, 255-273.
- Karelaia, N & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, 134, 404-426.

- Klein, G. (1989). Recognition-primed decisions. *Advances in Man-Machine Research*, 5, 47-92.
- Klein, G. A. (1999). *Sources of Power: How People Make Decisions*. Cambridge, MA: MIT Press
- Klein, G. A. & Crandall, B. W. (1995). The role of mental simulation in naturalistic decision making. In P. Hancock, J. Flach, J. Caird, and K. Vicente (eds.), *Local Applications of the Ecological Approach to Human Machine Systems (vol. 2)*. Hillsdale, NJ: Erlbaum.
- Kragt, H. & Landeweerd, J. A. (1974). Mental skills in process control. In E. Edwards & F. P. Lees (Eds.), *The Human Operator in Process Control*. London: Taylor & Francis.
- Kruger, J. & Evans, M. (2004). If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Personality and Social Psychology*, 40, 586-594.
- Lagnado, D. & Sloman, S. (2004). Inside and outside probability judgment. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making*. Boston: Blackwell.
- Landerweerd, J. A., Seegers, J. J., Praageman, J. (1981). Effects of instruction, visual imagery, and educational background on process control performance, *Ergonomics*, 24, 133-141.
- Langham, M., Hole, G., Edwards, J., O'Neill, C. (2002). An analysis of 'looked but failed to see' accidents involving parked police vehicles. *Ergonomics*, 45, 167-185.
- Lewandowsky, S., Dunn, J.C., Kirsner, K., Randell, M. (1997). Expertise in the management of bushfires: Training and decision support. *Australian Psychologist*, 32, 171-177.
- Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, 75, 5-18.
- Lundberg, J. (1999). *Distributed Situation Awareness for the Civil Cockpit: Theory and Operationalization* (No. LIU-KOGVET-D-0013-SE). Linköping: Linköping Universitet.
- Markman, K. D., Gavanski, I., Sherman, S. J., McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology*, 29, 87-109.
- Moran, T. P. (1981). The command language grammar: A representation for the user interface of interactive computer systems. *International Journal of Man-Machine Studies*, 15, 3-50.
- Moray, N. (1997). Human factors in process control. In G. Salvendy (Ed.), *Handbook of Ergonomics and Human Factors*. New York: Wiley.
- Moray, N., Lootsteen, P., Pajak, J. (1986). Acquisition of process control skills. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-16, 497-504.
- Nussbaum, S., Liberman, N., & Trope, Y. (2006). Predicting the near and distant future. *Journal of Experimental Psychology: General*, 135, 152-161 .

- Oettingen, G., Pak, H., Schnetter, K. (2001). Self-regulation and goal-setting: Turning free fantasies about the future into binding goals. *Journal of Personality and Social Psychology*, 80, 736-753.
- Oman, C. M., Shebilske, W. L., Richards, J. T., Tubre, T. C., Beall, A. C., Natapoff, A. (2000) Three dimensional spatial memory and learning in real and virtual environments. *Spatial Cognition and Computation*, 2, 355-372.
- Pennington, N. & Hastie, R. (1993). A theory of explanation-based decision making. In G. Klein, J. Orasanu, R. Calderwood, C. E. Zsombok (Eds.), *Decision Making in Action: Models and Methods*, 188-201. Norwood, NJ: Ablex.
- Randel, J. M., Pugh, H. L., Reed, S. K. (1996). Differences in expert and novice situation awareness in naturalistic decision making. *International Journal of Human-Computer Studies*, 45, 579-597.
- Rasmussen, J. (1979). On the Structure of Knowledge—A Morphology of Mental Models in a Man-Machine Context (Riso Report M-2192). Roskilde, Denmark: Riso National Laboratory.
- Rasmussen, J. (1983). Skill 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. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. Amsterdam, The Netherlands: North-Holland.
- Redelmeier, D., Koehler, D. J., Liberman, V., & Tversky, A. (1995). Probability judgment in medicine: Discounting unspecified alternatives. *Medical Decision Making*, 15, 227-230.
- Rescher, N (1999). *Predicting the Future: An Introduction to the Theory of Forecasting*. Albany: State University of New York.
- Rips, L. J. (2001). Two kinds of reasoning. *Psychological Science*, 12, 129-134.
- Roberts, W. A. (2002). Are animals stuck in time? *Psychology Bulletin*, 128, 473-489.
- Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, 66, 805-818.
- Romeyn, J. W. (2004). Hypotheses and inductive predictions: including examples on crash data. *Synthese*, 141, 333-364.
- Rottenstreich, Y. & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review*, 104, 406-415.
- Rouse, W. B., & Morris, N. M. (1985). *On looking into the black box: Prospects and limits in the search for mental models* (DTIC #AD-A159080). Atlanta, GA: Center for Man-Machine Systems Research, Georgia Institute of Technology.
- Rumelhart, D. E. (1984). Understand understanding. In J. Flood (Ed.), *Understanding reading comprehension*. Newark, NY: International Reading Association.
- Salthouse, T. A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General*, 113, 345-371.

- Sanna, L. J., Chang, E. C., Meier, S. (2001). Counterfactual thinking and self-motives. *Personality and Social Psychology Bulletin*, 27, 1023-1034.
- Sanna, L. J., Stocker, S. L., Clarke, J. A. (2003). Rumination, imagination, and personality: Specters of the past and future in the present. In E. C. Chang & L. J. Sanna (Eds.), *Virtue, Vice, and Personality: The Complexity of Behavior*, 105-124. Washington DC: American Psychological Association.
- Serfaty, D., Macmillan, J., Entin, E. E., & Entin, E. B. (1997). The decision making expertise of battle commanders. In C. E. Zsombok & G. A. Klein (Eds.), *Naturalistic Decision Making*. Mahwah, NJ: Lawrence Earlbaum.
- Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes*, 53, 252-266.
- Shepperd, J. A., Ouellette, J. A., Fernandez, J. K. (1996). Abandoning unrealistic optimism: Performance estimates and the temporal proximity of self-relevant feedback. *Journal of Personality and Social Psychology*, 70, 844-855.
- Sherden, W. A. (1998). *The Fortune Sellers: The Big Business of Buying and Selling Predictions*
- Simon, H. A. (1992). What is an explanation of behavior? *Psychological Science*, 3, 150-161.
- Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis testing processes in social interaction. *Journal of Personality and Social Psychology*, 36, 1202-1212
- Sulis, K., Wickens, C. D., Chui, Y. P. (2011). Prediction in situation awareness: confidence bias and underlying cognitive abilities. *The International Journal of Aviation Aviation Psychology*, 21, 153-174.
- Taylor, S. E., Pham, L. B., Rivkin, I. D., Armor, D. A. (1998). Harnessing the imagination: Mental simulation, self-regulation, and coping. *American Psychologist*, 53, 429-439.
- Toffler, A. (1970). *Future Shock*. London: Bodley Head.
- Tversky, B. (1993) Cognitive maps, cognitive collages, and spatial mental models. In Frank A. U. and Campari, I. (Eds.) *Spatial Information Theory: A Theoretical Basis for GIS, Proceedings COSIT '93*. Lecture Notes in Computer Science, 716, 14-24. Berlin: Springer.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Tversky, A., & Koehler, D. J. (2004). Support theory: A nonextensional representation of subjective probability. In E. Shafir (Ed.), *Preference, belief, and similarity: Selected writings by Amos Tversky*. MIT Press.
- Vidal, M., Amorim, M.-A., Berthoz, A. (2004). Navigating in a virtual three-dimensional maze: How do egocentric and allocentric reference frames interact? *Cognitive Brain Research*, 19, 244-258.
- Wagenaar, W. A. & Sagaria, S. D. (1975). Misperception of exponential growth. *Perception & Psychophysics*, 18, 416-422.

Weber, E.H. (1846). Tastsinn und Gemeingefühl. In R. Wagner (Ed) *Handwörterbuch der Physiologie*. Vol. III, 481-588. Brunswick: Vieweg.

Wickens, C. D. (1986). The effects of control dynamics on performance. In K. R. Boff, L. Kaufman & J. P. Thomas (eds.), *Handbook of Perception and Human Performance: Volume 2*. New York: Wiley.

Wickens, C. D. (2000). The trade-off of design for routine and unexpected performance: Implications of situation awareness. In D. J. Garland & M. R. Endsley (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Erlbaum.

Wickens, C. D. (2002). Situation Awareness and Workload in Aviation. *Current Directions in Psychological Science*, 11, 128-133.

Wickens, C.D. (2008). Situation Awareness: Review of Mica Endsley's 1995 Articles on Situation Awareness Theory and Measurement. *Human Factors*, 50, 397-403.

Wickens, C. D. & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall

Wickens, C. D., Gempler, K., Morpew, M. E. (2000). Workload and reliability of predictor displays in aircraft traffic avoidance. *Transportation Human Factors*, 2, 99-126.

Williams, A.M. & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. *Research Quarterly for Exercise and Sport*, 69, 111-128.

Williams, A. M., Ward, P., Knowles, J. M., Smeeton, N. J. (2002). Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology: Applied*, 8, 259-270.

Wilson, J. R. & Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors*, 31, 617-634.

Chapter Three

The Engineering of Prediction: Predictive Aids

3.1 OVERVIEW

Aside from reliable mental models, deriving accurate predictions also requires perceiving the right cues. Oftentimes cues can be made more salient through technological solutions. In fact, with the advancement of computing power, it has even become possible for machines to replace manual predictions, and outperform humans in deriving timely and accurate forecasts. Today's computers can now run simulations through programmed simulator models and data feed in order to analyze situations and generate predictions. Yet the *engineering of prediction* shares similar limitations as humans too, as the topics reviewed in this chapter will show.

There currently exist various **types of predictive aids** used in many domains and industries. The functions of these predictive solutions range from improving cue salience such as alerting the user of an impending undesirable state or visualizing the future trajectory of moving objects, to advanced computerized simulations which generate detailed predictions. These in turn facilitate users in decision-making and anticipatory actions to mitigate problems so as to keep operations within healthy boundaries.

Predictive aids are not without their **performance limitations**. Oftentimes these tools rely on a computational model of the system as well as its span of prediction. Overly-simplified models would not generate accurate predictions, whereas models which are too complicated may require too much computation resources. The further the automation can predict, the more useful it becomes for most users. Too short a span of prediction, particularly for sluggish systems, would not provide any significant benefit for users to anticipate and initiate control procedures promptly. Yet the further this span of prediction, the more complex the process becomes in deriving accurate, specific predictions. Such are some of the performance limitations that predictive aids face.

With automated predictive aids facing limitations in calculating predictions, their resultant performances may not always be flawless. **Imperfect automation** in predictive technology is a concern as most predictive aids may not always be 100% accurate all the time. It is thus critical to consider the effects of (possibly unexpected) imperfect predictive aids on their users.

3.2 TYPES OF PREDICTIVE AIDS

Humans are weak in making predictions. While people tend to make predictions often, they are often not accurate (e.g.: Kahneman & Lovallo, 1993). Yet most of us are required to make predictions throughout a variety of tasks. We predict whether the weather will stay sunny as we hang our laundry out to dry. We thus turn to the weatherman who makes a forecast on where showers will likely occur in the country. Fire fighters anticipate how bushfires will spread so as to come up with the best resolution strategy (Lewandowsky et al, 1997). The pilot envisions where his fast-moving aircraft will be in the near-future and makes preemptive control inputs so as to position his aircraft ideally on, and not overshooting, the glide slope for landing. Since human prediction is essential but limited, technological solutions have been developed and improved the performance of these predictive tasks (Jensen, 1981; Palmer, Jago, Baty, O'Connor, 1980; Wickens & Morphew, 1997).

Predictive aids or displays fundamentally support users by making information about the near-future more obvious so as to reduce users' effort required in making mental predictions. As seen in the previous chapter's literature review, a lot of mental activities go on in deriving an estimate of the future. Having a computer to *automate* this process and summarizing the result in a visual output will certainly make inferences about the future more straight-forward. This is especially so in systems with sluggish, higher-order dynamics (Wickens, Haskell, Harte, 1989; Lintern, Roscoe, Sivier, 1990; Morphew & Wickens, 1998), as mental simulations can be hard to perform particularly under time pressure, and when those systems have complex and sometimes non-intuitive dynamics such as often is the case with process control (Roth & Woods, 198x). However, predictive aids do not always provide benefits, since

predictive aids especially for complex systems may not be 100% reliable. Like humans, the computer makes some assumptions about the future forces acting on the system in order to infer about the future. At times not every critical force is accounted for by the automation, particularly when there tends to be a tradeoff between complex computation and processing speed. The notion of imperfect automation will be covered in greater detail later. In general, predictive aids can be found in three kinds of operations: **selection of action**, **manual control**. And process control, including both air traffic control, and manufacturing processes.

Selection of action

The future is unknown, yet many action-based decisions are made in present time based on what people predict the likely future outcome would be. People invest in stocks and shares based on (ideally) a calculated prediction that these investments would reap financial returns in the future. Weather forecasts would influence whether one should make a trip to the beach or stay indoors. Without any technological aid, making these decisions would involve the human inferring from cues derived from the current environment (a weakening U.S. dollar; thick, dark cumulonimbus clouds in the far distant). Stock brokers for example rely on current economic news as well as market performance trend plots. These cues present bits and pieces of relevant information on the current situation or past states, and the human has to mentally combine these information together and perform computation in order to extrapolate the future. Such mental processes are known to be effortful (Johannesen, Moray, Pew, Rasmussen, Sanders, Wickens, 1979), and the resulting predictions may still be untimely or inaccurate.

Rather than relying on mental computations, a statistical approach involving the processing of algorithms has proven to generate improved predictions. A meta-analysis study by Grove et al (2000) revealed a superiority in performance by algorithms over manual clinical predictions in various analyses such as length of psychiatric hospitalization, college academic performance, and job turnover. Statistical algorithms operate in a fashion similar to humans, wherein relevant cues are picked up and processed, but omit the computational inconsistency often found in humans such

as psychological influences and biases. Algorithms thus prompted the development of technological interventions which aided decision-makers who require a prediction about the future (Wickens, Mavor, Parasuraman, McGee, 1998; Gallaher, Hunt, Willeges, 1977). To reduce the occurrences of people defaulting loans, some banks now provide loan officers with software that aid in the evaluation and approval of personal loans. These software consider demographic and personal data rather than the officers' subjective impressions to determine the customer's reliability in financing the loan in the future.

Dichotomous decision-support aids similar to the personal loan software (to loan or not to loan) can also be found in meteorological systems. Tsunami warning systems aim to relay possible impending rogue waves to affected shorelines in hopes to warn and advise people to evacuate to higher ground. Tsunamis are series of waves which occur because of an underwater earthquake in the sea or large lake. The earthquake causes a large displacement of water flowing away from the epicenter, which typically goes unnoticed in the deep waters but swell to become huge waves as they approach the shore. Without any efficient means of early visual detection, a combination of modeling, seismic readings and knowledge of tsunamis is used to detect and predict tsunamis (Titov, 2009). Despite the advancement in tsunami prediction technology, false alarms are still prone to occur during seismic activities. Tsunami warning systems are also unable to provide timely warnings during sudden tsunami incidents. Nonetheless such systems serve to aid humans in deciding if tsunamis will likely occur in the near future, and whether evacuations should be conducted.

Predictive aids can be continuous too rather than discrete. Weather forecasters often use spatial displays to illustrate the predicted path that storms might take (Figure 3.1). These storm tracker predictors demarcate the possible range using a confidence interval. In another words, the middle track is the most likely path Hurricane Frances would take, while the edge of the potential track area shows areas with low track path probability. By considering various environmental factors and with a certain level of confidence, computers can predict the range of potential tracks the hurricane will take, and government organizations such as the Federal Emergency Management Agency

(FEMA) in the United States can decide who should be evacuated. Conversely, we can safely say from the prediction that Cuba (located south of Hurricane Frances) would not be affected. The weather forecast allows us to confidently conclude the areas where the hurricane will *not* reach.

Fig. 3.1 A storm prediction depicting Hurricane Frances' possible track into the future.

Source: National Oceanic and Atmospheric Administration, USA.

While computerized predictors may be more reliable than humans in some cases, they do share certain limitations found in human judgments too. Like us, predictive algorithms depend on the cues which they are programmed to infer, but when these available cues are persistently weak and uncertain, the extrapolated predictions will still be ineffectual. Predictive software for financial investments remain much to be desired, as visible cues are limited and other influential factors are dynamic and not easy to detect (Clements & Hendry, 2001). The algorithm's span of prediction and its subsequent prediction confidence are also akin to human characteristics. In the hurricane forecast, the ambiguity of Hurricane Frances' track increases as the prediction extends further into the future. For the same confidence level, the hurricane prediction is less accurate and reliable if it is made far into the

future, just as we are typically more uncertain about the future as the number of possible developments increases over time (Dunning, 2007). The notion on span of prediction will be elaborated further later in this chapter.

Manual control

Whereas the *selection of action* is a discrete process, *manual control* involves a continuous process of selecting actions and deciding what the next input should be in order to maintain the system within a desired state over time. Many of today's manual-control systems can be described as a continuous output generated by a continuous input, and the human operator is a part of this closed-loop process by continuously observing the situation, initiate an appropriate input, so as to achieve a desired output. One drives by constantly observing the current traffic situation and adjusting the car through the steering wheel, accelerator and brake pedals. A process plant is continually monitored by a controller who manages various process parameters so as to derive the optimum production process. Certain manual control tasks can be described as *tracking*, in which the operator's goal is to either stabilize the system around a reference despite facing disturbances or pursue an evasive target over time. This tracking loop is generally represented in Figure 3.2, which is adapted from models in Control Theory (Goodwin, 2001; see also Wickens, 1986) and illustrates the elements influencing a simplified manual control task. Based on a reference, the operator makes an action on the controls which feed inputs into the system. The system's performance is reflected back onto the display and the cycle repeats. Throughout the control process, disturbances may act on the system, and the operator would have to adjust the controls in order to compensate for the disturbance inputs. Examples of such disturbances include keeping the car on a straight path on the highway while experiencing consistent wind gusts coming from the side. Notably, manual closed-loop control can be more complex, featuring aspects such as feed-forward (versus feed-back) predictive controls (Crossman & Cooke, 1974) as well as adaptive and learning control loops (Moray, Lootsteen, Pajak, 1986; Moray, 2004).

Fig 3.2 Representation of a closed-loop tracking operation

Wickens & Hollands (2000) noted that manual control tasks can be difficult when operators have to deal with system lags and disturbance inputs. Lags or any form of time delay are generally detrimental to any human performance, be it manual control or selection of action. Manually-controlled system lags can occur between the input controls and the system where the system is sluggish in responding to the control inputs, as well as between the system and the display in which a time delay exists in transmitting the system's performance status. As such the operator has to anticipate these lags when making control inputs by predicting where the future output would be. A classic example would be in space teleoperation, where rovers on the moon and Mars are being controlled remotely from Earth. On the moon, visual data takes 2.5s to travel to Earth, and control inputs take the same amount of time to travel back to the machine, effectively creating a lag time of 5s and resulting in less-than-ideal control performance by the operator (Krotkov, Simmons, Cozman, Koenig, 1996). Controlling a Mars rover increases this challenge, as it would involve an even longer lag of potentially up to 45 minutes (Krotkov & Simmons, 1996).

Regardless of task objectives, any system lags can be mitigated using predictive aids that allow at least some form of preview of forthcoming command inputs or disturbances (Jensen, 1981; Tomizuka & Fujimura, 1979). The control task becomes more challenging when dynamic disturbance inputs are present, which would act on the system beyond an operator's controls. Oftentimes disturbance inputs cannot be anticipated, and operators do not know the presence of these disturbance inputs until they have already acted on the system. A car driver is unable to "see" oncoming wind gusts before they have affected the vehicle. Automated predictive aids would

greatly mitigate the effects of lags and disturbances, minimizing the mental efforts required of operators and allowing them to preview the system's future output.

Aircraft controls have benefited much from predictive aids, and predictive displays can be seen in today's glass cockpit concepts. As aircrafts are moving in very high speeds, current spatial information becomes obsolete very quickly and pilots instead constantly seek out information that tells them more about future situations so that they can "stay ahead of the plane" by anticipating future aircraft state and make early control inputs. Pilots can now refer to the Cockpit Display of Traffic Information (CDTI), which visualizes other aircrafts flying within the vicinity (Boeing, 1977). When linear predictors are added to the CDTI, pilots are able to better anticipate airspace conflicts and their workload is significantly reduced (Morphew & Wickens, 1998; Wickens, Gempfer & Morphew, 2000). Each predictor line or "noodle" (Figure 3.3) provides information on the heading as well as future location of each aircraft, assuming no change in the current control inputs and flight performance data. Predictor lines that curved according to the aircrafts' turn rate information further benefitted the pilots in maintaining aircraft separation (Palmer, Jago, Baty, O'Connor, 1980; Hart & Loomis, 1980). These predictors served as perceptual cues that aid pilots in predicting future situations and potential conflicts, a task which pilots would otherwise have to mentally compute and increase cognitive workload.

Fig 3.3 An illustrated navigation display found in aircraft cockpits, showing the “noodles” of the own aircraft as well as other traffic in the vicinity (Morphew & Wickens, 1998)

The Traffic Alert and Collision Avoidance System (TCAS) is another successful predictive aid currently implemented in commercial aircrafts (Kuchar & Drumm, 2007). Unlike the CDTI, the TCAS is primarily a discrete event predictor, alerting pilots of potential mid-air conflicts with other aircrafts. When a collision is predicted to occur within the next 20 to 48 seconds, the TCAS sounds out a spoken message “traffic, traffic” to the pilot. If the situation deteriorates further whereby the conflict will occur in the next 15 to 35 seconds, the TCAS would either announce “climb, climb” or “descent, descent”, prompting the pilot to take the appropriate control action immediately. Through calculating the flight data of the two aircrafts in conflict, the TCAS determines which aircraft is advised to climb or descent. The TCAS also selects the maneuvers that require the smallest change in order to achieve the required separation. However, the TCAS does not explicitly indicate where or when the collision would likely to happen, only that it is present in the near future. Other similar aviation displays that provide information on future hazards include the Ground Proximity Warning System (GPWS) as well as the Synthetic Vision-Primary Flight Display (Prinzel & Wickens, 2009), both of which provide terrain data and warnings.

Jensen (1981) described the advantage in pilot performance when showing a display with present as well as predicted position of the aircraft during a landing approach. Doherty & Wickens (2001) further explored the implementation of preview and prediction elements in 3D immersed perspective flight path displays. Adapted from this research, Figure 3.4 shows the preview element—the tunnel-in-the-sky indicating where the aircraft should be as time progresses. The predictor element—the white aircraft symbol beyond the aircraft’s current location (black bore sight) indicates the aircraft’s projected future position given no further input (from pilot or turbulences). These two elements can also be described using the terms “command state” and “predictive state”; the preview element illustrates the command state (ideal path information) where the aircraft should be, and the predictor element presents the predictive state (state of aircraft in the near future) which the pilot is trying to control. Doherty & Wickens (2001) found that while predictor elements in general were beneficial, pilots benefitted more if known command state information (tunnel-in-the-

sky) were presented. In actual flight situations, attention towards such displays may come at the cost of attention diverted off other tasks, such as visual scanning outside the cockpit or monitoring the flight instruments, thus posing a source of danger. This could also be a concern in process control, as any displayed information would distract operators from considering other non-displayed information.

Fig. 3.4 A screenshot of a tunnel-in-the-sky display featuring both the predictor and preview elements (Doherty & Wickens, 2001)

The tunnel-in-the-sky display has since been developed into what is known today as the Synthetic Vision System. The National Aeronautics and Space Administration (NASA) in the United States has shown statistical performance benefits in pilots when negotiating through difficult approach routes in low-visibility conditions (Prinzel et al., 2002; Kramer et al., 2009). In this study, the pilots flew NASA's Boeing 757-200 aircraft which was experimentally retrofitted. A post-hoc interview revealed all the pilots' increased confidence in flying airspace procedures with conditions of low visibility and many high terrain obstacles when using the Synthetic Vision System. They commented that the Synthetic Vision System provided more situational awareness, and allowed them to predict the flight path more easily. It should be noted, however, that overly-increasing realism in visual displays can cause visual clutter and the reduction of critical information salience, as well as the phenomenon known as naïve realism (Smallman & Cook, 2011; Smallman & John, 2005) wherein users possess "misplaced, blanket faith" on synthetic displays which actually hinder performance.

Air traffic control

Air traffic control (ATC), despite vastly different from manual-based, active vehicular control, shares the same human-in-the-loop concept. In ATC, controllers need to devote much cognitive resources in computing potential conflicts and collisions, and then communicate process changes to live, moving targets in hopes that these targets would comply accordingly. Managing air traffic is complex: controllers assign instructions to aircrafts, then monitoring for pilots' response and actions, while all the time maintaining safe separation between all other aircrafts within the airspace. Long total transmission lag of around 20 to 25 seconds make this process additionally difficult (Wickens, Rice, Keller, Hutchins, Hughes, Clayton, 2009). Adding to the dynamic nature of air traffic, pilots may request for alternative actions. They may also react unexpectedly faster or slower to the commands. All these factors make it more difficult for ATC to predict the future state.

Current ATC Radar displays have velocity vectors (predictor lines similar to the “noodle” found in CDTIs) for each aircraft depicted on screen, showing the location where each aircraft would be after a pre-determined amount of time later given no further change in speed and trajectory (Figure 3.5). In the past ATC controllers would rely on the history trails of aircrafts to aid them in extrapolating their future locations. Most of today's systems include velocity vectors which are simply straight lines showing where aircrafts will be in the future given their current velocities, and these vectors remain straight even if the aircraft turns. They also aid controllers in estimating how fast an aircraft is going (length of the line), what direction the aircraft is heading, and whether two or more converging aircrafts are in a potential airspace separation conflict. Newer radar displays may feature velocity vectors which curve and thus provide the turn-rate information of turning aircrafts. If the aircraft had filed a flight plan with the ATC, certain radar displays can also map out these flight plans, further providing controllers with expected positions the aircraft will be in the future. Oinonen et al (2009) revealed limited benefits in straight-line vectors for multiple aircraft tracking, but which showed trends of facilitating performance than with no velocity vectors at all. Figure 3.5 also illustrates a variant of the ATC display which

integrates weather information, using colored patches to denote precipitation. While all these information may aid operators in managing the aircrafts, they create “display clutter” which may be detrimental to performance (Wickens, Kroft & Yeh, 2000; Nunes et al., 2006).

Fig 3.5 A screenshot of an ATC display, showing predictive lines of each aircraft and weather information to aid controllers’ decision-making

Maritime

Beyond aviation, maritime vessels benefit from predictor displays as well. As boats are operated on low-friction water surfaces, control inputs tend to experience certain amounts of time-delay reaction. Operators of quick-moving vessels face the challenge of sluggish controls in a fast-changing environment. Some automated controllers on high-speed passenger-carrying vessels now have the ability to “look ahead” from 30 to 120 seconds, thus giving an early indication of the vessel’s future state and whether it was approaching danger (Kallstrom, Bjore, Bystedt, 1996; Kallstrom & Bjore, 1997). Larger, heavier ships are described to have even higher inertia, as the lags experienced by these vessels are even more significant. It can take minutes before the heading of a heavy ship reacts to an angled rudder. Novice helmsmen therefore have a tendency to oversteer as they fail to anticipate this input-output lag. The seriousness of this problem becomes more apparent when large ships have to travel through confined waterways amidst heavy traffic. An evasive maneuver

may be initiated, but may not always result in prompt changes by the ship. Hence, sailors of large ships like supertankers have benefit from the use of predictive displays (van Breda, 1990).

Sullivan et al (2006) explored the benefits of having a quickening display for large ships. A ghost image of a simulated 40-foot, 16-ton vessel was projected ten seconds into the future based on a linear regression model using the ship's current and previous rudder angle, course over ground, and heading (Figure 3.6). As the ghost ship turned according to the current control input, the participants were essentially steering the predictor image of the vessel rather than that of the current ship. Results showed that experienced and especially novice participants performed better when using the predictor to maneuver the vessel along a fixed track. Interest in improving the helmsman's control over large vessels has lead to the development of automated predictive prototypes which support real-time prediction of the vessel's position and heading (Transport Canada, 1999).

Fig 3.6 A screenshot of the vessel navigation simulator featuring the predictor used by Sullivan et al (2006)

Categorization of predictive aids

Across various operations, predictive aids can be classified according to Table 3.1. Predictive aids may provide either *discrete* (e.g.: presence versus absence of certain states) or *continuous* forms of information regarding the future. These information may either *explicitly* indicate the future expected forecast of the event, or would otherwise be *implicit* and require the user to interpret and come up with a prediction. Notably, most predictive aids aim to provide Level 3 SA (Endsley, 1995) information and omitting Level 2 SA details. In time-critical situations like aviation, pilots probably do not need to thoroughly comprehend the details of the situation like what might the conflicting aircraft's current state or intention, but they require sufficient awareness of the impending collision based on their converging flight paths. This is also reflected in the model illustration (Figure 2.3) described in Chapter Two whereby predictions can be derived separately from situation comprehension.

Table 3.1 Categorization of predictive aids

	Discrete	Continuous
Implicit	TCAS	Aviation velocity vectors
Explicit	Tsunami warning system	Aircraft flight plans; Hurricane forecasts

An implicit-discrete aid alerts the user the presence of a potential event, serving primarily to draw the user's attention and awareness towards it. While the aid is able to calculate the potential of an oncoming event, it does not state where and when this event will occur. If the exact detail is required, the user has to mentally process this along with other relevant information so as to derive a prediction. The TCAS is one such example wherein it alerts the pilot of a potential mid-air collision, but does not indicate to the pilot where and when this collision will take place.

An implicit-continuous aid provides constant updates about the future state and give cues of a looming collision, but still requires the user to manually derive where and when this collision will occur. The straight-line velocity vectors found in aviation displays such as CDTIs and ATC radar screens project the locations of surrounding aircrafts in the near future given their current velocity and heading, but users have to determine on their own whether two converging aircrafts are potentially in conflict, and visualize the conflict's time and location in the airspace. Even more "implicit" displays would include ATC radar displays that feature the historical tracks of aircrafts, which allow controllers to observe their recent behaviors and infer their future intentions. Pilots too are better able at "staying ahead of the plane" and anticipate maneuvers using tools like Vertical Situation Displays (Prevot & Palmer, 2000) and glideslope descent-rate cuing (Lintern, Kaul, Collyer, 1984). These tools do not explicitly predict future states, but present information which is useful for prediction.

Explicit-discrete and explicit-continuous aids share similar descriptions as the previously-mentioned, except these aids now clearly define a prediction of where the targets will be. The tsunami warning system is an explicit-discrete aid which alerts the community of when a specific shoreline will be struck by tsunamis during a certain period of time. Examples of explicit-continuous aids would be hurricane forecast displays as well as ATC radar displays which feature flight plans of various aircrafts. Both tools point out where the targets will probably be at various moments in the future, along with potential conflicts like populated areas to be evacuated (hurricanes) or mid-air collision (aviation).

3.3 PERFORMANCE LIMITATIONS IN AUTOMATED PREDICTIONS

To derive a prediction, automated aids rely on two key factors: a model of the system to be predicted, and a time element denoting how far into the future the prediction is meant for. Similar to how humans make bottom-up predictions (prediction based on inputs from the environment), predictive aids typically process current data through a series of modeling algorithms, and simulate this model to project what would happen over a period of time given the current settings. The

“mental models” of these predictive aids can vary in complexity depending on the systems which they are trying to emulate.

A fairly simple model would be aircrafts. The basic behavior of fixed-wing aircrafts, with limited degrees of freedom and few control inputs, is fairly easy to simulate. Airplanes don't move backwards in flight, and their rate of turn can be derived via their current bank angle. Hence a curved velocity vector can be calculated for CDTI displays representing the aircraft's turning flight path (Hart & Loomis, 1980). However this model, which relies solely on the primary control data of the aircraft for input, may not factor in wind effects, nor can it predict pilot intentions, the latter requiring the comprehension of the aircraft's flight plan. A simple model, while easy to compute and simulate, will have limited predictive capabilities. More complicated models like the artificial neural network models used in Transport Canada's Ship Predictor System (1999) or those used in tsunami warning simulations (Titov, 2009) provide higher predictive “resolution”, but would require more processing resources and thus requiring more time before coming up with accurate predictions. At times these lags limit predictive aids' abilities to provide timely predictions for sudden events, and may thus be unsuitable for fast-changing environments.

The amount of “look-ahead” time also dictates the quality of predictions made by these automated aids. Naturally the further the prediction is able to cover, the higher the value of this prediction becomes. However, automation shares the same problem humans do regarding the span of predictions, in which the accuracy of prediction decreases as the span of prediction increases further into the future, even as this accuracy loss is mitigated somewhat by a long time constant (sluggishness) of the system. That is, for example a 1 minute look-ahead time for a super tanker will be a more accurate predictor than a 1 minute look-ahead time for a light aircraft. But within each dynamic system, accuracy will degrade with longer LAT. People progressively lose more confidence when predicting more distant events using low-level information pieces (Nussbaum, Liberman, Trope, 2006). These bits of information become more uncertain further into the future, since they would have more time to vary dynamically. The uncertainty can be described using a normal distribution, wherein the statistical odds of a variable's change magnitude follow a Gaussian pattern. For example, a

passenger airplane will most likely be flying straight and remain in a straight-line location instead of making a turn. Should it turn, it is more likely to make a gradual than a sharp turn, even though both maneuvers are equally possible to be executed by the pilot. The probability of the aircraft's future location therefore follows a normal distribution. This exponential growth of Gaussian-pattern variability in a system's future state can be deemed as *Gaussian perturbation* (Figure 3.7, see also Wickens, Mavor, Parasuraman, McGee, 1998 on deviation within displayed reliability, as well as Gempeler & Wickens, 1998 on the “spray angle” of reliability estimate generation). Considerations towards the Gaussian perturbation can be seen in hurricane predictions, in which the range of potential track directions increases as the forecast spans further into the future. A predicted trajectory may thus also appear to be asymmetric, but still abide by the Gaussian pattern of probability, such as when predicting a plane flying towards a mountain. The predicted norm would be for the plane to react by ascending over the obstacle, and less likely to fly straight or descent into the mountain.

Fig 3.7 The Gaussian perturbation describes the growth in uncertainty as the span of prediction increases, in which probability of each possible directional change follows a normal distribution.

Wickens (1986) noted that the usefulness in the amount of “look-ahead” time varies from system to system, and depends on the bandwidth and magnitude of disturbance inputs, as well as the inertia of the system's behavior. The two extremes of such systems can be illustrated using a heavy ship and a light propeller aircraft. A

heavy ship such as an oil tanker has high inertia, in which it takes the ship a relatively long time before it reacts to a control input. As a result, the ship is more resistant to higher bandwidth and magnitude of disturbance inputs, and inherently allows for a longer prediction span of its behavior. Conversely a small, general aviation propeller aircraft has low inertia, so although it responds nimbly to control inputs, it is also more susceptible to disturbance inputs. Longer prediction spans for small aircrafts are less accurate in the presence of strong, stochastic disturbance inputs, and therefore may not be as useful as shorter prediction spans. Jensen (1981) revealed that a predictor interval of 8 seconds was considerably beneficial for light aircrafts. Nonetheless, in general all systems benefit from relatively longer span of prediction given weak disturbance inputs, low-inertia systems appreciating it more than high-inertia systems.

3.4 IMPERFECT AUTOMATION

Given the limitations of automation in deriving predictions, predictive aids are oftentimes never perfect. A speed-accuracy trade-off exists in most automated predictive aids, where given a limited amount of computational capabilities a swiftly-generated prediction may be too generalized and incomprehensive, but an accurate prediction may require longer processing time and therefore become too slow to be of value. An unreliable predictive aid may not just be ineffective (Metzger & Parasuraman, 2005), but possibly become detrimental to task performance (Wickens & Dixon, 2007; Levinthal & Wickens, 2006; Bliss & Acton, 2003; Parasuraman & Riley, 1997; Meyer, 2004), although oftentimes the fundamental solution is to calibrate the user's trust to the actual reliability of the automation (Muir, 1987; Muir & Moray, 1996). It is thus important to understand the impacts of imperfect automation before considering its implementation into any system. Two factors influence the reliability of predictions: Variability of the future and discrete error.

The predicted future state may have high levels of variability, the Gaussian perturbation shown previously in Figure 3.7 being an example of this uncertainty. This variability can also be described using confidence intervals, wherein the predictive aid is certain of a future state, or range of states, given a specific level of confidence. An example in the hurricane forecast domain would be when the predictor produces a

large range of possible trajectories, which would not benefit the user in pinpointing where the hurricane might be in the near future. The variability of the future may change continuously over time, and the predictor's confidence intervals may vary during different stages of the system's process.

Another form of unreliability comes in the form of discrete error, in which the tracked target eventually behaved in a manner totally different from what the predictive aid indicated. Discrete error is, as the name implies, a discrete event: the predictor is either predicting correctly or produces a discrete error. It can occur in conjunction with confidence levels, such as when a range of possible trajectories has been predicted given a certain confidence level, but the hurricane eventually deviates out of this indicated range. In dichotomous predictive aids, the confidence levels of the predictor would result in either more false-alarms or misses (Wickens, Rice, Keller, Hutchins, Hughes, Clayton, 2009), and that prolonged exposure to these conditions can affect performance like attention distraction (Wickens, Dixon, Goh, Hammer, 2005) and the lack of compliance to alarms or the "cry-wolf" effect (Meyer, 2004; Wickens et al, 2009). Gempler & Wickens noted that providing a fixed "wedge predictor" to show the other aircrafts' possible heading directions did not significantly help pilots during moments of discrete errors.

Nevertheless, automated predictions do not need to be perfect in order to be beneficial, so long it provides more help than bother to its users. Wickens, Gempler and Morpew (2000) noted in a study on imperfect CDTI display of air traffic avoidance that pilots always do use the predictive aid despite knowing that it is unreliable, as it is right most of the time, and when it is right it is very helpful for conflict avoidance. Wickens & Dixon (2007) did a meta-analysis of studies that examined imperfect dichotomous predictors revealed a strong positive linear function between automation reliability and the generated prediction's costs and benefits. The higher the automation's reliability, the more beneficial it would be for the user. A predictive aid's reliability level of 71% was deemed the cut-off point, going below this threshold and the user would perform better without being presented the automated analysis at all. Even when the automation is below this reliability threshold, the negative effects of unreliable automation can be mitigated. Wickens et al (2009)

studied an ATC traffic conflict alerting system with only 50% reliability on average and thus generating false alarms fairly often. The controllers understood the need for high sensitivity in detecting potential conflicts due to the serious nature of mid-air collisions. The inconvenience of false alarms was a small price to pay as compared to a conflict that missed detection. Controllers also had access to the raw air traffic information, and could distinguish true conflict alerts and false alarms manually and effectively.

3.5 SUMMARY

Technology has come a long way to aid humans in making predictions. The automation of prediction shares similar processes as how humans do it: predictive cues are picked up, processed through an algorithmic model, and run a simulation to derive a forecast of the future system state. How far the prediction has to predict into the future also affects accuracy of prediction, as the further the span of prediction, the more dynamic the unknown future input variables would be. Predictive aids can either be explicit or implicit, and their predictions can either be discrete or continuous. More often than not, these automated aids are imperfect, their confidence intervals fluctuating over time or occasionally producing discrete errors. Nonetheless, given sufficient reliability levels, imperfect predictive aids can still be beneficial for users.

After reviewing the psychology (Chapter Two) and engineering (this Chapter) aspects of prediction, we now focus deeper towards the context of process control. Much technological advancement in terms of intelligent and automated controls exist in this industry, yet automated predictive aids meant at supporting operators' decision-making appear to be limited as compared to other domains. In the next chapter, the general algorithms used in today's process control technology are reviewed to see if any of which can be used to power a prototype predictive display, as well as find out what makes process control unique and challenging towards developing predictive tools.

3.6 REFERENCES

- Bliss, J. P. & Acton, S. A. (2003). Alarm mistrust in automobiles: How collision alarm reliability affects driving. *Applied Ergonomics*, 34, 499–509.
- Boeing Commercial Airplane Company. (1977). *Cockpit displayed traffic information study* (Report No. D6-42968). Seattle, WA: Author.
- Clements, M.P., Hendry, D.F. (2001). An historical *perspective* on forecast errors. *National Institute Economics Review*, 177, 100-112.
- Crossman, E. R. F. W. & Cooke, J. E. (1974). Manual control of slow-response systems. In E. Edwards & F. Lees (Eds.), *The Human Operator in Process Control*, London: Taylor & Francis.
- Doherty, S. M. & Wickens, C. D. (2001). Effects of preview, prediction, frame of reference, and display gain in tunnel-in-the-sky displays. *Proceedings of the 11th International Symposium on Aviation Psychology*. Columbus, OH: Dept. of Aerospace Engineering, Applied Mechanics, and Aviation, Ohio State University.
- Dunning, D. (2007). Prediction: The inside view. In A. Kruglanski & E. Higgins (Eds.), *Social Psychology: Handbook of Basic Principles*, 2nd ed., New York: Guilford.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37, 32-64.
- Gallagher, P. D., Hunt, R. A., Williges, R. C. A. (1977). A regression approach to generate aircraft predictor information. *Human Factors*, 19, 549-556.
- Gempler, K. S. & Wickens, C. D. (1998). *Display of predictor reliability on a cockpit display of traffic information* (Final Tech. Rep. No. ARL-98-6/ROCKWELL-98-1). Savoy, IL: University of Illinois Institute of Aviation, Aviation Research Lab.
- Goodwin, G. (2001). *Control System Design*. Upper Saddle River, NJ: Prentice Hall
- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A metaanalysis. *Psychological Assessment*, 12, 19-30.
- Hart, S., & Loomis, L. L. (1980). Evaluation of the *potential format* and content of a cockpit display of traffic information. *Human Factors*, 22, 591–604.
- Jensen, R. S. (1981). Prediction and quickening in perspective flight displays for curved landing approaches. *Human Factors*, 23, 355-363.
- Johannsen, G., Moray, N., Pew, R. W., Rasmussen, J., Sanders, A. F., Wickens, C. D. (1979). Final report of the experimental group. In N. Moray (Ed.), *Mental Workload: Its Theory and Measurement*. New York: Plenum.

- Kahneman, D. & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39, 17-31.
- Kallstrom, C.G., & Bjore, A. (1997). A track autopilot predictor system for ships. Proceedings of the 11th ship control systems symposium. Southampton, England, vol. 2, 23-37.
- Kallstrom, C.G., Bjore, A., Bystedt, S. (1996). Track autopilot predictor system for STENA HSS 1500. *Proceedings of the International Conference on High Speed Marine Craft, Safe Design and Safe Operation*. Bergen, Norway.
- Kramer, L. J., Bailey, R. E., Prinzel III, L. J. (2009) Commercial flight crew decision making during low-visibility approach operations using fused synthetic and enhanced vision systems. *International Journal of Aviation Psychology*, 19, 131-157.
- Krotkov, E. & Simmons, R. (1996). Perception, planning and control for autonomous walking with the Ambler Planetary Rover. *International Journal of Robotics Research*, 15, 155-180.
- Krotkov, E., Simmons, R., Cozman, F., Koenig, S., (1996). Safeguarding teleoperation for lunar rovers: From human factors to field trials. *Proceedings of the IEEE Workshop on Planetary Rover Technology and Systems*, Minneapolis MN.
- Kuchar, J. K., Drumm, A. C., (2007). The Traffic Alert and Collision Avoidance System (TCAS). *MIT Lincoln Laboratory Journal*, Vol. 16, Number 2.
- Lewandowsky, S., Dunn, J.C., Kirsner, K., Randell, M. (1997). Expertise in the management of bushfires: Training and decision support. *Australian Psychologist*, 32, 171-177.
- Levinthal, B. R. & Wickens, C. D. (2006). Management of multiple UAVs with imperfect automation. In *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors & Ergonomics Society.
- Lintern, G., Roscoe, S. N., Sivier, J. L. (1990). Display principles, control dynamics, and environmental factors in pilot training and transfer. *Human Factors*, 32, 299-318.
- Metzger, U., & Parasuraman, R. (2005). Automation in future air traffic management: Effects of decision aid reliability on controller performance and mental workload. *Human Factors*, 47, 35-49
- Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors*, 46, 196-204.
- Moray, N., Looftstein, P., Pajak, J. (1986). Acquisition of process control skills. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-16, 497-504.

- Moray, N. (2004). *Ergonomics: Major Writings. Volume 4—Manual Control, Industrial Processes, and Automation*. Boca Raton, FL: Taylor & Francis.
- Morphew, E. M., & Wickens, C. D. (1998). Pilot performance and workload using traffic displays to support free flight. In *Proceedings of the 42nd Annual Meeting of the Human Factors & Ergonomics Society*. Santa Monica, CA: Human Factors Society.
- Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man–Machine Studies*, 27, 527–539.
- Muir, B.M. & Moray, N. (1996). Trust in Automation: Part II. Experimental Studies of Trust and Human Intervention in a Process Control Simulation. *Ergonomics*, (39)3, 429-460.
- Nunes, A. & Matthews, M. L. (2002). The effect of predictive aid usage on controller strategies & mental demand under direct routing. In *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors Society.
- Nussbaum, S., Liberman, N., & Trope, Y. (2006). Predicting the near and distant future. *Journal of Experimental Psychology: General*, 135, 152-161 .
- Oinonen, K., Oksama, L., Rantanen, E., Hyona, J. (2009) Do velocity vectors support multiple object tracking? In *Proceedings of the 53rd Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors Society.
- Palmer, E. A., Jago, S. J., Baty, D. L., O'Connor, S. L. (1980). Perception of horizontal aircraft separation on a cockpit display of traffic information. *Human Factors*, 22, 605-620.
- Parasuraman, R. & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39, 230-253.
- Prevot, T. & Palmer, E. A. (2000) Staying ahead of the automation: A vertical situation display can help (SAE Technical Paper 2000-01-5614). Warrendale, PA: **SAE** International.
- Prinzel, L.J., Kramer, L.J., Comstock, J.R., Bailey, R.E., Hughes, M.F., & Parrish, R.V. (2002). NASA synthetic vision EGE flight test. In *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics*. Santa Monica, CA: Human Factors Society.
- Smallman, H. S. & Cook, M. B. (2011). Naïve realism: Folk fallacies in the design and use of visual displays. *Topics in Cognitive Science*, 3, 579-608.

- Smallman, H. S. & John, M. S. (2005). Naive realism: Misplaced faith in realistic displays. *Ergonomics in Design: The Quarterly of Human Factors Applications*, 13, 6-13
- Sullivan, B.M., Ware, C., Plumlee, M.D., 2006, "Predictive Displays for Survey Vessels", *Human Factors and Ergonomic Studies (HFES)*, San Francisco, CA, USA, 16 - 20 October, pp. 1 - 5. *Conference Proceeding*.
- Titov, V.V. (2009). Tsunami forecasting. In E. N. Bernard & A. R. Robinson (Eds.), *The Sea, Volume 15: Tsunamis*. Cambridge, MA: Harvard University Press
- Tomizuka, M. and Fujimura, M. (1979). Extended signal quickening for manual control. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-9, 669-676.
- Transport Canada (1999). Development and field testing of a neural network ship predictor system (SPS) (Technical Report TP 13368E). Ottawa, ON: Transport Canada
- van Breda, L. (1999). *Anticipatory behaviour in supervisory control*. Delft, The Netherlands: Delft University Press.
- Wickens, C. D. (1986). The Effects of control dynamics on performance. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Volume II) (pp. 39.1-39.60). New York: John Wiley & Sons.
- Wickens, C. D. & Dixon, S. R. (2007). The benefits of imperfect diagnostic automation: a synthesis of the literature. *Theoretical Issues in Ergonomics Science*, 8, 201-212.
- Wickens, C. D., Dixon, S. R., Goh, J., Hammer, B. (2005). Pilot dependence on imperfect diagnostic automation in simulated UAV flights: An attentional visual scanning analysis. In *Proceedings of the 13th Annual International Symposium of Aviation Psychology*. Dayton, Ohio.
- Wickens, C. D., Gempler, K., Morpew, M. E. (2000) Workload and reliability of predictor displays in aircraft traffic avoidance. *Transportation Human Factors*, 2, 99-126.
- Wickens, C. D. & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall
- Wickens, C. D., Kroft, P., Yeh, M. (2000). Database overlay in electronic map design: Testing a computational model. *Proceedings of the IEA 2000 / HFES 2000 Congress*. Santa Monica, CA: Human Factors & Ergonomics Society.
- Wickens, C. D., Mavor, A. S., Parasuraman, R., McGee, P. (1998). Airspace system integration: The concept of free flight. *The Future of Air Traffic Control*. Washington DC: National Academy Press.

Wickens, C. D., & McCarley, J. (2008). *Applied attention theory*. Boca-Raton, FL: Taylor & Francis.

Wickens, C. D. & Morpew, E. (1997). Predictive features of a cockpit traffic display: A workload assessment (Technical Report ARL-97-6/NASA-97-3). Savoy, IL: University of Illinois, Institute of Aviation Research Lab.

Wickens, C.D., Rice, S., Keller, D., Hutchins, S., Hughes, J. & Clayton, K. (2009). False alerts in the air traffic control traffic conflict alerting system: Is there a cry wolf effect? *Human Factors*, 51(4), 446-462.

Chapter Four

Predictive Applications in Process Control

4.1 INTRODUCTION

In order to develop a predictive trends display for process control, a means of calculating the prediction over a given span of time is required. This chapter reviews the existing research as well as process control technology that might be adapted for this project. Roth and Woods (1988) developed a simple two-element predictive trend display which projected the future location of the trend line given no change to the current control input (Figure 4.1). The predictive calculations were based on the degree of feedwater flow in or out of the reservoir to indicate the level of water. It functioned as a “straight-line” predictor akin to those found in air traffic control displays, where the predictor line merely indicates the trajectory and location of an aircraft, and leaves no clue as to what future flight path the aircraft may have. Peacock, Schlegel & Brace (1985) proposed the use of a process simulation to run parallel to live operations, but at a faster, “quickened” speed so as to provide information about the future. Similar to the previous example, this would allow the operators to anticipate what’s to come and react proactively.

Fig. 4.1 Predictive trend display developed by Roth and Woods (1988).

However oftentimes one process variable is influenced by many other variables in the system, which thus can drastically increase the computational load. Aside from the interactions between multiple components, the process behavior itself is often dynamic and *nonlinear*. Cooper (2006) gave a good example using gravity-drained tanks (basically two barrels, each with a hole at the bottom, stacked one on top of the other). Whenever the flow of fluid into the top barrel increased, the increased water volume forced more water into the bottom barrel. Assume that we can only control the opening size of the orifice that fed water into the top barrel, each equal increment at the top controller led to *increasing* increments of flow rate into the lower barrel (Figure 4.2). This is hence a nonlinear process, which is typical for process control.

Fig. 4.2 An equal increment in one variable causes an increasing increment in another (Cooper 2006).

Due to the complexity of process control, “intelligent” predictions that project contextual, non-linear trends are more beneficial than “straight-line” predictors similar to those in air traffic control. In order to explore the potential of non-linear trends in intelligent process control with predictive displays, suitable algorithm theories were identified through reviewing the chemical engineering literature as well as consulting process engineers. Understanding these algorithms is a daunting experience for people

without knowledge in chemical engineering background. Hence this chapter presents a general overview of the predictive algorithms currently available in process control.

While these algorithms appear to be viable means for calculating predictive information, the nature of industrial process control poses significant challenges that would undermine their effective application. The remainder of this chapter gives an account of process control characteristics, and discusses the limitations of developing intelligent predictive displays in this domain.

In current process plants, predictive technology has been employed to automate production operations. **Advanced process control** allows the use of automation to control plant systems so as to manage and optimize the production process. Process optimization requires finely-controlled inputs in a stabilized system, and the predictive capabilities in-built in the automation allow more accurate automated process control. Such automation requires reliable predictions of control variables over a future time horizon so as to derive suitable actions to be taken. With increasing technological capabilities, calculations that were once considered too complex can now be computed. Many process control laboratories and companies have thus developed variants of control algorithms, some of which are trade secrets. Methods for calculating process control predictions are generalized below using two concepts: Model Predictive Control and Qualitative Trend Analysis.

4.2 MODEL PREDICTIVE CONTROL

Model Predictive Control (MPC) utilizes an explicit dynamic model of a plant to predict the future plant state. The prediction is then used to manipulate input variables to achieve the output target (Garcia et al., 1989). The MPC program can identify an input sequence over the control horizon which minimizes costs while considering system constraints, and at the same time maintain the output as close to the target level as possible.

In Figure 4.3, at the present time “ k ”, the behavior of the process over the horizon “ p ” is considered. Using a model of the plant, the process output “ y ” is predicted based on the changes in the manipulated input variables. Whenever a new input is made through the manipulated variables, the program calculates and predicts the new output over a period of time. As the process moves along in time, it constantly updates itself with a computation that uses the current plant state to derive a new prediction.

Fig. 4.3 Optimizing the predicted output through MPC (Garcia et al., 1989).

Most models assume process linearity, commonly referred to as **linear MPC**, which can oversimplify the actual process, and may therefore be inadequate for large, industrial operations. This leads to the development of **nonlinear MPCs** where a more accurate, nonlinear process model is used for prediction. Such a model is very complex and cannot be based on the original linear models. Rather it is derived from either **first-principles** or an **empirical** model, or a combination of both (Hansen, 1998).

A first-principles model includes the application of mass, energy, and momentum balances, together with rate and equilibrium equations to define the process behavior. For a dynamic model these need to be transient or time differential equations. An empirical model is developed based on actual dynamic plant data. Compared to empirical models, first-principles models require much less process data for development. As long as the underlying assumptions in the model hold, a first-principles model can calculate and extrapolate the process predictions, even if the process experiences unexpected stochastic disturbances. However, large first-principles models are hard to develop, maintain, and require extensive computation, and thus it can be challenging to use them for industrial processes. Empirical models, on the other hand, do not require detailed process understanding for their development. They rely on historical plant data, lab simulations or plant tests to derive the model and generate process predictions. Relative to first-principles method, data-driven modeling induces less computation workload. On the other hand, prediction reliability decreases during unexpected events, which are not captured in historical data or offline simulations.

Variants of MPC algorithms have been developed for many industrial applications, and their usage is rapidly growing (Qin & Badgwell, 2003). Given the condition where the plant is within its normal operating envelope or performing within calculated expectations, operators are satisfied with linear-modeled MPC controllers in managing the process. Coming up with adequate nonlinear MPC programs is still a challenge, and a plant process that is extremely non-linear is thus unable to benefit from current MPC methods.

Due to the computational complexity of nonlinear MPC algorithms, successful nonlinear MPC applications are generally smaller in size and scope as compared to linear MPC applications (Martin & Johnston, 1998). This raises the issue of “speed-accuracy trade-off”, where the version for control by linear or nonlinear algorithms is either too inaccurate for industrial process control, or takes too long to compute to be used for predictive aids. Most successful applications are achieved through balancing this trade-off, where the reduced reliability of the automation, given a tolerable computational speed, is still beneficial enough to justify its use. Yet effective

utilization may still be restricted to specific plants and operating conditions which the models are designed for, such as a stable plant operation and a preprogrammed, anticipated process events.

MPC models have to tolerate a certain amount of unmeasured disturbances and modeling errors. While it is easy for MPC to perform using known disturbances to the system, MPC techniques are difficult to implement when the disturbances are unknown or not considered in their computations. Grimm et al (2004) demonstrated this lack of robustness when the model does not factor process perturbations which resulted in the process being manipulated within an oscillating control range but nowhere near its optimum state. MPC Models usually expect these disturbances to remain constant over the prediction horizon. Yet under many conditions the assumptions don't hold true.

Batina et al. (2001) accounted from literature three basic approaches in which MPC deals with disturbances. The first approach assumes that the disturbance is known, and is either zero or constant throughout the process interval. This is deemed too "optimistic" and unrealistic as it ignores the effect of what the disturbance may have on system performance. In an example where raw water is being treated via a dynamic cleaning process (e.g.: varying amounts of chemical treatments depending on water quality), the first approach assumes that the quality of raw water always remains in a typical expected range, and that the amount of treatments would not increase when the water quality deteriorates beyond this expected range. The second approach assumes the unknown disturbance to belong to a class of signals, and the MPC calculations are based on a *min-max* approach: the input variables are computed at their minimum values, while the disturbance variables are computed at their maximum values. This approach tries to identify the worst possible disturbance realization, and thus is generally considered too "pessimistic" to be cost-efficient. In the case of the water treatment example, when the process' MPC model predicts the future raw water quality it anticipates for the worst possible example, and thus allows more chemical treatments to be used albeit uneconomically. The third approach factors in the probability of violating constraints, thereby investigating if the control system can be improved by a considerable degree at the expense of a small risk. If the current

application allows that risk, the potential performance benefit despite using such imperfect predictors would still be significant (Batina, 2004). That is to say, when the water treatment process' MPC model predicts the future raw water quality, it factors in the probability of poorer quality and its impact on the final product before controlling the amount of chemicals to be used. Solutions on tackling stochastic disturbances have been proposed, such as incorporating linear properties on the nonlinear model, thereby sacrificing prediction accuracy.

Essentially, the success of MPC programs relies on the use of a basic dynamic model for predictions. With many interacting components, designing an accurate nonlinear model of a large-scale plant is very difficult. Successful implementations would also require constant updating and configuring, which requires much manpower.

4.3 QUALITATIVE TREND ANALYSIS

Qualitative features in process data are often used in intelligent control systems for process analysis and prediction. One means of identifying a process situation is by analyzing past trend data. This may allow a compact representation of the actual trend through the detection of significant process “signs”. In most cases, process events leave a distinct trend or marker in the monitored sensors (Venkatasubramanian et al., 2003). These distinct markers can be utilized in identifying the status of the process or even any underlying abnormality. Through these signs, a signature of the particular event can be formed. A database of these signatures would facilitate computer analysis as it scans the past trend data, matches it with the database, and identifies a current and possible future state of process variables. The faster the analysis and classification, the earlier a potential fault may be detected and diagnosed, and the quicker the implementation of corrective actions (Cooper & Lalonde, 1990; Uraikul et al., 2007). This technique of extracting useful trend features to form coded event patterns is commonly known as Qualitative Trend Analysis (QTA) in the process industries.

Many types and techniques of QTA exist. Cheung and Stephanopoulos (1990a) introduced the method of **triangulation** to represent trends, where each segment of a trend is represented by its initial slope, final slope (or critical point of a trend), and a

line segment connecting the two critical points. A series of such triangles would form a process trend episode (Figure 4.4), and the triangles within the trend can be considered as geometric primitives, showing the “signature” of this episode via a series of different basic components (Figure 4.5). A combination of episodes would thus form a trend which completely describes the qualitative state of the system (Janusz & Venkatasubramanian, 1991). As such, the triangles in Figure 4.5 reflect the boundaries for the actual trend, illustrating the maximum error in the trend representation. While providing qualitative data interpretation, triangulation of trend data can also be fed into process modeling to derive quantitative analysis. The triangulation method of representing trends allows sufficient inferences for both measured and unmeasured trends.

Fig. 4.4 Representing a trend data using triangulation (Cheung and Stephanopoulos, 1990a).

Fig. 4.5 Geometrical basic triangular components (Cheung and Stephanopoulos, 1990a).

A similar representation language was proposed by Janusz and Venkatasubramanian (1991), where seven fundamental shapes known as **primitives** are used to represent any trend (Figure 4.6). Each primitive consists of the sign of the first derivative, and the sign of the second derivative (or zero). Therefore, each primitive informs whether the function is increasing, decreasing, or not changing, as well as the concavity. Using the method of backward finite difference (technical details see Janusz & Venkatasubramanian, 1991), the derivative values of each primitive at a determined time point are calculated. To extract the qualitative features of a trend, the primitives are first assigned to each time interval, combined to form episodes, and the sequencing of the episodes generate the trend. Using this method on Figure 4.7, a trend signature can be developed and described as “(D3) (A1) (G4) (E4) (A1) (B4)”. Each parenthesis indicates the primitive which the portion of the trend is identified as, as well as the duration of the primitive that occurred. Trends are further evaluated by running it through a trend classification tree of increasing complexity as well as pattern matching. The model in the study by Janusz and Venkatasubramanian (1991) was able to convert the final analysis into a descriptive complete sentence, such as “*the temperature is oscillating with a decreasing amplitude to a higher steady state*”.

Fig. 4.6. Primitives identified by Januz & Venkatasubramanian

Fig. 4.7. Assigning the primitives, forming the episodes, and eventually developing the trend signature as “(D3) (A1) (G4) (E4) (A1) (B4)” (Janusz & Venkatasubramanian, 1991).

QTA relies on the fact that process signals can be represented at different levels of details, and that similar events result in qualitatively similar trends and vice versa (Maurya et al., 2007). Process events can be a combination of activities happening at different time-scales. Generally the trend of a variable may be generated from different processes, as it is usually a function of other process variables (Cheng & Stephanopoulos, 1990b; see Figure 4.8). A deviation may occur slowly over a long period of time, or quickly and thus forming sharp peaks or drops. If the signal detection window is designed to capture the smallest process dynamics, a slow, long-duration deviation would not register as an activity and hence resulting in incorrect diagnosis. As such, the window should be able to adapt accordingly to the frequency content of the trend in order to pick up the right information at the most appropriate resolution. The window size should be small enough to facilitate unique identification of primitives, and yet large enough so as not to be greatly affected by noise. Vedam and Venkatasubramanian (1997) came up with an adaptive trend identification algorithm which adjusts the window size according to the frequency content of the data. Their adaptive window was thus able to identify primitives fairly accurately in the presence of significant levels of noise and when the trend data evolved over a long period of time.

Fig. 4.8. Formation of a trend variable by considering how a primitive trend pattern interacts with minute deviations and noise (Cheung & Stephanopoulos, 1990b).

To be robust, QTA methodologies have to incorporate some form of noise filtering on the trend data. Most often, trend data do not appear as a smooth-flowing line, but rather jagged and edgy due to the presence of noise (e.g.: sensor noise, process disturbance, see Figure 5.8). Hence smoothing mechanisms are needed to filter out the noise and derive the underlying qualitative structure of the trend data. A filter can be linear or non-linear, time-invariant or time-variant, causal or non-causal. Existing filters include moving averages, exponential smoothing, orthogonal smoothing, Savitzky-Golay Filtering etc. For QTA, in particular, filters need to possess two significant characteristics (Venkatasubramanian et al., 2003). First, due to the nature of process control and the functional relationships between various components, it should be able to operate with derivatives of varying orders (first, second, or higher). Second, as changes in trends occur at different scales, optimal detection depends on the use of filters with appropriate scale factors. One such filter used in QTA is the Gaussian filter (Figure 4.9, technical explanation can be found in Cheung & Stephanopoulos, 1990b), which begins smoothing the data at a small scale, and gradually increasing the filter size to encompass the whole trend.

Fig. 4.9 Illustration of Gaussian smoothing at successive scales (Cheung & Stephanopoulos, 1990b).

QTA seems to avoid many problems of imperfect modeling in MPC models. It offers a cleaner, less computationally-intensive way of deriving a process prediction. Using QTA to generate predictive information is not without its difficulties. Most QTA methods rely partly on a database of patterns which aid the program in matching and identifying the current qualitative data analysis. The database may be limited to known events that have already occurred, and new events would not be recognized. While similar events should carry somewhat similar signatures, these patterns may vary depending on the differences in components and production methods between process plants.

Given the exact same task, trend patterns may appear differently due to some components being broken, under repair, or not functioning properly. Conversely, the process may be dealing with a new event, but the program wrongly identified it as a different, older event. The effectiveness of QTA's application would be dependent on the comprehensiveness of the database. Because of the need to abstract information from varying scales of past trend data, it is still unclear how far back into the past the data collection needs to go before deriving useful predictions. Depending on the required QTA resolution, the program may, for example, require 20 minutes of fresh historical data in order to predict one minute into the future. For an event to be

accurately identified, enough amount of the event's trend pattern needs to be captured before the program can understand and predict the current situation. This may result in a lag or delay in the predictive information, which may harm the operator's performance. The predictive information may show that, based on past history, the process variable is increasing at 5 degrees per minute, when in actual fact the increase is 10 degrees per minute, but there is insufficient past data to detect the increase.

4.4 CHALLENGES FOR PROCESS CONTROL PREDICTIVE DISPLAYS

Due to the nature of process control, prediction of future parameters and trend behavior of process variables can be difficult to automate. In order to be useful for operators, automated predictive aids must be sufficiently reliable (Wickens, Gempier & Morpew 2000). While the accuracy of the prediction need not be spot on, it should not deviate too much from the actual behavior of the variable and lead operators to the wrong conclusions. The frequency by which the prediction falls below the minimum accuracy threshold is an indication of the predictive aid's reliability. Intelligent predictive displays that rely on control algorithms to generate forecasts face many challenges, most of which funnels down to a common problem of unreliability.

Steady-state Algorithms

The key issue in utilizing control algorithms is that they are designed for automatic management and optimization of the production process. Advanced process control programs are often designed for operation within the normal operating conditions and a moderate range of process disturbances. This is akin to the auto-pilot function in commercial aircrafts: once the plane is in straight and level flight, the automation takes over and provides smoother, more efficient inputs than manual control. In both cases, the automation works around anticipated disturbances. Operators need only to monitor the program, and respond only when problems occur, in which the automation would be shut off.

Predictive information is of most value to operators during abnormal situations. If abnormal situations are detected early, they allow operators to perform preemptive

intervention and prevent losses. Yet control algorithms are most unreliable during this period, and thus cannot provide the aid that operators seek. With reliability constantly fluctuating during the dynamic production process, presenting a trend predictor may be more of a clutter than convenience. Much work and consistent upkeep will be needed for such a predictive aid, and benefits will only be reaped long after implementation, when the trend predictor has been tuned and is now stabilized.

Different Lab and Field Performances

Most predictive algorithms and the software that use them seem to have a high level of reliability based on lab tests and simulations (Qin & Badgwell, 2003). The reliability level can drop significantly when implemented in live operating plants. During normal operations, components can go down due to faulty or maintenance work, but the plant can still function safely due to redundancy systems in place. Developing a complete model that mimics the entire plant is effortful and challenging; live calculations will also take up too much computing resources. Developers place great efforts in fitting all anticipated process events and disturbances into the models or pattern-recognition databases, but even then all possible variations of disturbances will still not be captured fully.

Maintenance and Upkeep

Considering the algorithms' steady-state applications as well as dips in reliability during online operations, predictive aids require consistent attention from engineers in maintaining and updating the program. Engineers will need to capture models and databases during the moments in which unexpected events and disturbances occur. While such resource commitments are still reasonable for process optimization programs, it would not be economical to use predictive displays that require constant care from engineers in order for them to be useful for operators. This is further compounded by the problem of capturing the uniqueness of each process plant. Process plants are rarely identical in terms of production process, hardware and technological tools. While the fundamental prediction concept may be the same, the predictive software would need more maintenance in a more complex plant.

External variables

Algorithms may not be able to capture random disturbances that originated beyond the production process, such as weather and atmospheric conditions. Operators often mention some process chances which were due to weather phenomena, such as the expansion of tanks due to high temperatures, or reduction in temperature within distillation towers due to storms. While human operators are quick to point out these problems and consider them in their control strategy, it may be difficult to program these factors so as to reflect them accurately in automated prediction.

4.5 SUMMARY

Process control operators apparently benefit from presenting future behaviors and trajectories of plant parameters. However, while it is intuitive to develop similar, “intelligent” predictive displays found in other domains, current process control technology is unable to support such an initiative without remaining economical. Algorithms from Model Predictive Control and Qualitative Trend Analysis face issues of reliability, mainly due to the inherent complexity of industrial process control. These algorithms were designed for optimizing stable operations, and would fail quickly during occurrences of abnormal situations. Furthermore, the changing state of each unique process plant means that updating the predictive model or the trend pattern database would require constant, extensive manpower resources. Until these issues which cause unreliable predictive automation are resolved, utilizing these algorithms for predictive trend display development would create more problems than solutions. Perhaps the benefits of a process control predictive aid should be explored first, before establishing the need to come up with a realistic predictive algorithm: one that does not rely on real-time processing of tens to hundreds dynamic process variables, that is cost-effective to maintain in terms of labor and cost, that is simple to understand for process control operators, and most importantly, resilient to abnormal situations.

4.6 REFERENCES

- Batina, I., Stoorvogel, A. A., Weiland, S. (2001). Stochastic disturbance rejection in model predictive control by randomized algorithms. *Proceedings of the American Control Conference*. Arlington, VA.
- Batina, I. (2004). *Model Predictive Control for Stochastic Systems by Randomized Algorithms*. PhD dissertation, Technische Universiteit Eindhoven.
- Cheung, J. T.-Y. & Stephanopoulos, G. (1990a). Representation of process trends—Part I. A formal representation framework. *Computers & Chemical Engineering*, 14, 495-510.
- Cheung, J. T.-Y. & Stephanopoulos, G. (1990b). Representation of process trends—Part II. The problem of scale and qualitative scaling. *Computers & Chemical Engineering*, 14, 511-539.
- Cooper, D. J. (2006) *Design and Tuning Recipe Must Consider Nonlinear Process Behavior*. ControlGuru.com. Available at <http://www.controlguru.com/wp/p61.html>
- Cooper, D. J. & Lalonde, A. M. (1990). Process behavior diagnostics and adaptive process control. *Computers & Chemical Engineering*, 14, 541-549.
- Garcia, C. E., Prett, D. M., Morari, M. (1989). Model predictive control: Theory and practice—A survey. *Automatica*, 25, 335-348.
- Hansen, M. A. (1998). Nonlinear model predictive control: current status and future directions. *Computers and Chemical Engineering*, 23, 187-202
- Janusz, M. E. & Venkatasubramanian, V. (1991). Automatic generation of qualitative descriptions of process trends for fault detection and diagnosis. *Engineering Applications of Artificial Intelligence*, 4, 329-339.
- Martin, G. & Johnston, D. (1998) Continuous model-based optimization. In *Hydrocarbon processing's process optimization conference*, Houston, TX.
- Maurya, M. R., Rengaswamy, R., Venkatasubramanian, V. (2007). Fault diagnosis using dynamic trend analysis: A review and recent developments. *Engineering Applications of Artificial Intelligence*, 20, 133-146.
- Peacock, B., Schlegel, R. E. & Brace, T. (1985). Polar coordinate process control display. *Ergonomics International 85: Proceedings of the 9th Congress of the International Ergonomics Association*, Bournemouth, England.
- Qin, S. J. & Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11, 733-764
- Roth, E. M. & Woods, D. D. (1988). Aiding human performance: I. Cognitive analysis. *Le Travail Humain*, 51, 39-64.

Uraikul, V., Chan, C. W., Tontiwachwuthikul, P. (2007). Artificial intelligence for monitoring and supervisory control of process systems. *Engineering Applications of Artificial Intelligence*, 20, 115-131.

Vedam, H. & Venkatasubramanian, V. (1997). A wavelet theory-based adaptive trend analysis system for process monitoring and diagnosis. *Proceedings of the American Control Conference*. Albuquerque, NM.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., Yin, K. (2003). A review of process fault detection and diagnosis Part III: Process history based methods. *Computers and Chemical Engineering*, 27, 327-346.

Wickens, C. D., Gempler, K., Morphew, M. E. (2000) Workload and reliability of predictor displays in aircraft traffic avoidance. *Transportation Human Factors*, 2, 99-126.

Chapter Five

Field Studies

5.1 OVERVIEW

Before work can commence on developing a prototype process control predictive aid, certain knowledge gaps need to be filled first. According to the conceptual model (Fig. 2.3) established in the literature review, the two fundamental ingredients for prediction are “mental model” and “cue perception”. While both topics were addressed through references and applications in other domains, details specific to the context of process control were limited which therefore fueled the need to conduct these two field studies. **Qualitative Investigation One** looked at console operators’ mental models, eliciting information on how their mental models were derived, updated, and applied during task operations. Results of this study would help establish training protocols for subsequent lab-based simulator experiments to ensure that test subjects attained the appropriate mental models first prior to participating. **Qualitative Investigation Two** examined the types of information visualization displays that operators rely on to stay proactive, answering the questions “What are the cues that proactive operators perceive, and how are these cues presented on the display console?”. These details would provide insights on what critical cues are needed for a prototype process control predictive aid, and how these cues should be visualized.

5.2 QUALITATIVE INVESTIGATION ONE: OPERATORS’ MENTAL MODELS

Previous literature review on mental models revealed the multiple different definitions of mental models. Bainbridge (1992) highlighted how these different definitions which were designed for different contexts and applications resulted in much confusion, particularly when describing mental models in the process control domain. In process control, the perspective that mental models are permanent mental representation of some part of the external world (Edwards & Lees, 1974; Gentner & Stevens, 1983) has been adopted and shall be used as the operational definition for this project. Such mental models may lack the fine and complex details of the actual

system that is being visualized, but sufficient enough to infer what is happening in the process which cannot be observed directly, explain or choose what actions to take, or predict what is going to happen next (Bainbridge, 1992).

The purpose of the operator's mental model has been reiterated across many process control studies (Goodstein, 1982; Lind, 1982; Rasmussen, 1981; Sheridan, 1976 etc.). Bainbridge (1986) noted that a good mental model should have a goal-directed structure of information which allows for *mental simulation* to take place. It is a knowledge platform that drives skill-based processing and controls rule-based activities, or otherwise provides the capacity to reason and predict future plant state in order to attain knowledge of the situation. Evidently, such strong mental models are often associated with expert plant operators, armed with years of on-job experience working with the plant and dealing with abnormal situations. It is thus interesting to find out more about how expert operators gain and interact with this crucial mental model.

5.3 METHOD: ETHNOGRAPHIC OBSERVATIONS

An ethnographic approach was adapted to study expert console operators in a petrochemical processing site. Ethnographic studies have been conducted to investigate expertise in practice, such as in the context of anesthesiology (Smith, Goodwin, Mort & Pope, 2003), teaching (Moallem, 1998), and rugby refereeing (Ollis, MacPherson & Collins, 2006). Ethnography research involves a range of methods to understand the meanings of people's actions and explanations in day-to-day conditions. Insights are generated through describing and interpreting these individuals' actions or behaviors. In this study, one researcher was attached to one operator to conduct observations and details garnered from an observation were documented on paper as no audio or visual recordings were permitted. Whenever convenient the researcher would ask the operator specific structured questions regarding mental models and how they contribute to proactive monitoring. Observed events were also probed for additional details and explanations. Protocols were established to prevent the researcher from affecting operations should abnormal or critical events occur.

A total of ten expert console operators, all of them male, were observed from a petrochemical processing site in South Africa. These operators have been working on this site for 13 to 30 years (mean=23.6), and have been recommended by the operations management to be highly knowledgeable and skilled amongst their peers. Operators may or may not manage the same processing unit (e.g.: alkylation unit, crude tower, fractionators etc.). For each operator, observation began from the start of his shift in the morning (6am) where the outgoing operator would hand over the control console. The observation would end when the subject has completely handed over the console controls to the operator from the night shift (6pm). Only weekday, daytime shifts were observed as most of the refinery activities are conducted during these times. A total of 120 hours of observations was yielded from this study.

Three key themes were the focus of this study: *deriving*, *updating*, and *applying* mental models. As with any system user, a mental model of that system would be developed to help the user in understanding the system better, possibly coming up with more effective interaction strategies. Content for each theme is thus compiled from commonly observed findings and comments.

5.4 RESULTS & DISCUSSION FROM ETHNOGRAPHIC STUDY

A Day in the Life of a Console Operator

As in most process control operations, the role of the console operator was to oversee and manage a production unit or set of units from within a control room through the distributed control system (DCS). From this position, the console operator was able to have an overview of the entire process and remotely control almost all the equipment, some of which may be located very far apart and controlling them manually out in the field usually do not provide a good sense of situation awareness. The console operator was hence part of a team, which included field operators and team leaders, providing a “bird’s-eye view” of the overall unit(s), monitoring the production process, and coordinating with the field operators. Particular to this site, the console operators consisted of employees which were once field operators for many

years before receiving this “promotion”, and therefore aside from the team leaders, these console operators were domain experts who would guide the field operators.

Table 5.1. Typical weekday dayshift routine

Coming on shift (6am – 7am)	Shift handover Review shift log Email updates Check system bypasses and alarms
Morning (7am – 10am)	Shift meeting “Virtual Rounds” Updates from production planning Request for and monitor lab results
Late morning (10am – 12noon)	Request for and monitor lab results Manage process Update shift log
Afternoon (12noon – 5pm)	Manage process Update shift log Request for and monitor lab results
Prior to End (5pm – 6pm)	Prepare and await shift handover

Table 5.1 summarizes the routine that these ten dayshift console operators would go through. Console operators worked 12-hour shifts which began either at 6am or 6pm. These daily events were not in any particular order and were not always pegged to specific time of the day, but they were consistently conducted during time periods as outlined in Table 5.1. The specific details within each event may differ depending on daily process changes, scheduled activities, as well as various process unit differences.

The incoming operator would start his day by understanding what had been happening in the previous shift. He would engage in a verbal discussion with the outgoing operator at the console, occasionally calling up and using console displays to provide additional clarity. After the outgoing operator stepped out, the operator would check his email inbox for operational matters and updates, as well as the console displays for critical system notifications such as alarms and bypasses.

After the daily shift meeting, in which the operator would be briefed by his supervisor on the day's expectations, the operator would return to the console and begin his "walk-around" of the plant. The goal of this "virtual round" is similar to the physical rounds conducted by field operators, except that this is done from the console perspective. Eight of the ten operators actually explicitly related this event to the field rounds, and described the task as "mentally visualizing what he would see if he was out in the field". The console operator would begin scrolling through the various schematic console displays on the DCS to observe the process setup and readings of various parameters throughout all the units under his command. This virtual round would typically last around 90 minutes each day, and gave the operators a good overview of the units' conditions.

During this time, operators were asked how they knew what the parameter values should be, and if their expected values were different from those presented on the DCS, which values they would trust. All ten operators gave similar responses, attributing the knowledge of values to their working experiences on the DCS as well as from their past experiences in the field where they knew what the values of key parameters should be during various operation modes. In an event of a conflict between mental and actual values, operators trusted the DCS more, but would take simple actions to verify that the DCS readings were correct, such as justifying the readings with the operating state of relevant components.

Periodically, operators would receive information regarding production quality of the processes they are managing. Operators would submit requests to have product samples undergo quality checks in the laboratories. Results from these tests would return after a few hours depending on the time of day, and this routine would be conducted at least 4 times each shift, more frequently for other high-production units. It was also a trend for operators in this site to request for laboratory tests nearing the end of their shifts, so that when the incoming operator took over, he or she would have test results early in the shift. Operators would occasionally receive calls from the production planning department on new product requirements, thereby requesting changes to the current processes.

Deriving Mental Models

Expert, senior operators are known to have a wealth of knowledge. These ten operators attributed much of their process knowledge to the many years of practical experience working as field operators for at least 8 years before being promoted and entering the control rooms as console operators. This was typical for this site, although plans for accelerated console operator training and reduced field exposure were being discussed. When asked explicitly, all the operators referred to their field experiences as a major source for deriving and developing their mental models required for their tasks. In various similar expressions, operators described the need to “know the plant outside in order to work inside”.

All ten operators also cited their past experiences in handling upsets and situations as contributing factors. Being exposed to more incidents, these experts were thus able to better understand the process components, such as each component’s characteristics and whether they tend to perform in a certain manner, as well as how their operations affect other inter-related components. Through these past experiences, these operators could quickly relate newly encountered scenarios back to the past in order to aid their decision-making. Such behaviors reflect the nature of recognition-primed decision-making that many experts in various domain exhibit (Klein, 1999), and bestow the quality of being proactive amongst these operators.

The heavy emphasis in practical training and process knowledge meant that operators were able to visualize and recall not just the physical layout and characteristics of production components, but also the procedures and actions required for specific process events. One operator described in detail what process knowledge meant to him:

“To have process knowledge is to know what is the line-up, the layout of the plant, how the process works internally, why the equipment is here, what is happening within each component, what are the normal operating values and limits, what are the operating procedures.”

In addition to the knowledge gained through prior practical training, operators were also trained onsite in thermodynamic fundamentals. All in all, the “process knowledge” that these operators possessed encompassed various information which allowed them to have a mentally-derived conceptual understanding of the process situation. We recognize these as characteristics of mental models essential for operator expertise and effective process control.

Updating Mental Models

Updating the mental model, a stable concept based on our implied definition, is different from updating the situational assessment or “situation model”. Once developed, mental models are difficult to modify (Durso & Gronlund, 1999), and when it does happen it is usually over a span of hours, days, or even months (Wickens, 2008). Conversely, deriving an updated comprehension of the current state occurs quickly in the matter of seconds or minutes through perceiving the relevant cues (Endsley, 1995). Updating the mental model is thus facilitated by knowing if there are any major changes to the fundamental operational flows and processes. Arguably, the most simplistic, functional form of the process system which is resistant to changes would be deemed as the mental model (e.g. Rasmussen 1986). Feedback from control operators revealed that even though significant process changes are infrequent, they have to keep these new modifications in mind whenever decisions have to be made or when they visualize the systems in their heads. The presence of new installations, the decommissioning of old components, new process routes or re-routes, all these details in relation to the entire operation will need to be established before an operator can work effectively for the rest of the shift.

More often, operators had to update their mental model “modes” (or situation model) during shift. Although the general mental model remains stable, the model may feature many operational modes that the operator may need to switch around depending on the situation, such as from steady-state production to initiating emergency repairs and operating on bypass. Console operators were informed early at

the start of their shift about any operational changes or updates before further events were initiated. As equipment conditions can change over a shift, operators returning back to work twelve hours after their previous shift may find that the process units in a different setup or operation mode. Hence operators always share information about equipment failures, faults or bypasses during shift handovers. Outgoing operators often highlight areas of concerns in which the incoming operators do not just have to monitor carefully, but to also keep it in mind and take the situation into account when performing process moves. Incoming shift operators would also verify this information by checking the shift log as well as DCS console for notices on bypasses, active alarms and deactivated alarms. Similar details may also be reported during the morning meetings. Naturally as they interact with the process units throughout their shift, their mental model modes would adapt and change too according to the day's events. Similarly operators may also pick up new information during virtual rounds as they orientate themselves to the operational performance status of various process components. It is noteworthy that most prediction failures (given our interest in cue-based prediction) appear to result from situation- rather than mental-model failures, and which is more prone amongst novices (Doane, Sohn, Jodlowski, 2004).

The ten operators were asked if maintaining accurate mental model modes was a cognitive challenge given their busy operations, and whether they encountered experiences in which they momentarily operated with a wrong mental model mode. Seven operators noted that the information regarding component operation changes and process deviations can easily be found on the console screens, and that during routine monitoring operators would come across and be reminded of these information fairly easily. Three operators related it to memory, and described how they would try not to recall past configuration modes and treated each day as new, novel situations. Nonetheless they have no problems recalling details of recent past operation modes, and that they also had the shift logs for referring to the past. All ten responded that they do not suffer from problems related to operating with the wrong mental model. However four operators cited past incidents whereby they understood the situation wrongly, and thus made wrongful process moves. These problems seem to stem from decision-making biases such as anchoring bias.

Applying Mental Models

Possessing an updated mental model of the unit, the operators can now apply it to various tasks as required of them. When asked about the differences between experts and novices, oftentimes the topic of monitoring strategies was raised. To quote a participant, an operator should “know where to look, what to look out for, what values to expect, and why the values are as such”. Possessing efficient scanning strategies is one of many attributes of an expert, as already commonly reported in literature on expertise. Operators were prompted to describe how they knew where to look. Many operators credited their “experiences and process knowledge” to tell them where to look. One participant gave a vivid account of his thought process:

“When I arrive (virtually through the DCS) at a certain area of the plant, I would visualize myself as being physically in the field and imagine what equipment and what reading to look out for.”

This participant relied on his years of past experience as a field operator, along with the understanding of the processes that he acquired then, to perform his duties while sitting behind a control console. As this scanning behavior became more practiced over time, it is implied that expert operators automate this mental visualization and hence figure out quickly what to look out for without having to engage much cognitive resources. Such behaviors also reflect the proactive attributes that these expert operators are known for on this site. Beyond just efficient scanning, these experts were also able to confidently explain the current data readings and anticipate future readings: “why is the pressure at this reading now, why will the reading increase five minutes later after a process move is made to this unit etc.” Operators were able to draw rough sketches of the connections between various components, and were able to easily explain causal relationships between these components that affect the production process.

As a console operator who has a good overview of the production process, he also had the responsibility of commanding and coordinating with the field operators. When field operators are inexperienced, it was up to the console operator (who after

all was once an experienced field operator) to provide procedural instructions. With so many procedures currently existing in each process unit, how is it that these operators know the correct steps to take? Quoting an explanation:

“Memorizing the procedures won’t work. You need to know how the process is behaving outside, and take steps according to how you understand the process.”

Instead of blindly following a sequence of instructions, the console operator actually visualized the current process flow, and initiate systematic actions while monitoring process’ reaction towards achieving the desired state. The console operator therefore must decompose complex models to their controllable inputs (e.g. the pressure in a vessel may be reduced by either opening a relief valve or by reducing the temperature), and then deciding on the most appropriate action to take / command to give. Of all the ten observation sessions, only one instance was documented of an expert console operator guiding a novice field operator over the radio.

5.5 QUALITATIVE INVESTIGATION ONE: SUMMARY

To better understand how process control operators make bottom-up predictions, a qualitative study was done to understand how expert console operators interacted with their mental models, a vital component of bottom-up predictions. We classified the interaction into three forms: deriving, updating and applying the mental model. Results from this study allowed us to understand expert operators’ utilization of mental models, and what novice operators lacked.

Console operators appear to require a strong sense of spatial mapping and causal relationships within the units being monitored. Such knowledge facilitated them in performing their tasks, such as monitoring the process readings, coordinating and giving instructions to the field operators etc. A strong mental model of the units being managed would to support these task performances, and the results suggest how a strong mental model can be developed, is updated, and would be applied by console operators.

Notably, console operators engage heavily in mental simulations, a task which otherwise not be possible without the presence of mental models. There were many instances whereby the operators had to visualize themselves as being out in the field, or how the process was behaving as the operators initiated process moves. To “experience by proxy” appeared to be a common phenomenon, one that was highly beneficial in the complex world of process control where the user is limited to information displays, buttons and radio communication to manage and oversee a large chemical process.

The operators we observed and interviewed underwent a long training process as a field operator before finally becoming a console operator. Being out in the field and directly working on the units has helped these operators to develop a strong mental model of these units which they had to now manage remotely behind four walls. It would be beneficial (and our interest) to observe other sites with different console operator training setups to uncover how their mental models and cognitive processes might differ, as well as how the expert-novice gap can be mitigated most effectively.

5.6 QUALITATIVE INVESTIGATION TWO: TREND DISPLAYS IN PROCESS CONTROL

Amongst all other displays in the distributed control system (DCS), the interest in trend displays was partly due to a popular but undocumented comment in the process control industry that proactive operators often monitor trend displays. This hinted that trend displays are technological aids which helped operators in making predictions and anticipatory actions. Trend displays, or commonly referred to as “*Trends*” by process control operators, chart out data readings over a period of time. Trends can bring up multiple process variables within one screen through overlapping and color-coding the various graphs. Besides deciding on what variables to be monitored on the trend display, operators can also adjust the time scale, ranging from a few minutes to a few days.

The benefits of presenting data plotted over time have been well-established. Specifically, line graphs possess characteristics that make them suitable for applications in process control. They are good for showing temporal information, such as changes in data properties and variables over time. In closed-loop control domains, current situations are largely determined by past system states, and temporal information is thus beneficial for understanding current system states, predicting future system states, choosing appropriate control inputs, and detecting possible system anomalies (Bennett et al., 2005). Line graphs are advantageous for tasks requiring comparison of two different points, and patterns in line graphs may also “pop out” in a way that would not be as visible in other representations (Meyer, Shinar & Leiser, 1997; Wickens & McCarley, 2008; Figure 5.1). Line graphs can also produce salient, high-level emergent features that allow for easy anomaly identification (Pomerantz, 1986) in addition to mental integration of information and identifying overall trends (Carswell & Wickens, 1996). Hajdukiewicz & Wu (2004) documented the benefits of Trends in process control, which included abilities like detecting a change in a variable, its direction and rate of change, as well as any time-based patterns.

Fig. 5.1. The line graph on the top right clearly shows a decreasing trend in both variables through their slopes, as well as the different rate-of-change between the two variables through their angles (Wickens & McCarley, 2008).

Given the limited literature available on Trends in process control, a qualitative study was conducted to understand how Trends were used toward proactive monitoring. A series of semi-structured interviews were conducted to elicit operator's task and information requirements when using trend displays. The operators were prompted to think of an incident where they were mentally challenged as they were using trend displays. Probes from the Critical Decision Method (CDM) were used to deepen the story and operators were encouraged to verbalize their thoughts (Klein et al, 1989). CDM probes were chosen for the interview as they had been successfully applied in the study of expert behavior in similar environments. Several different professional scenarios have been investigated including: fireground commanders (Klein, 1998), air traffic controllers (Seamster, Redding, Cannon, Ryder, & Purcell, 1993), and ambulance dispatch managers (Wong & Blandford, 2001).

5.7 METHOD: KNOWLEDGE ELICITATION INTERVIEWS

A series of semi-structured interviews were conducted to elicit operator's task and information requirements when using trend displays. The operators were prompted to think of an incident where they were mentally challenged as they were using trend displays. Probes from the Critical Decision Method (CDM) were used to deepen the

story and encourage the operators to verbalize their thoughts (Klein, 1989). CDM probes have been successfully applied in the study of expert behavior in similar environments where quick and critical decisions were made, including: fireground commanders (Klein, 1998), air traffic controllers (Seamster, Redding, Cannon, Ryder, & Purcell, 1993), and ambulance dispatch management (Blandford et al., 2002).

In this study 17 subjects were interviewed, 9 from *Shell Bukom* in Singapore, and 8 from *Petronas* in Malaysia. These subjects had work experiences ranging between 14 to 18 years (average 16.1). They were considered as experts because they were either promoted to shift supervisors or were otherwise identified by the plant management to be highly proficient with the DCS console. All the subjects were male, full-time employees with regular work tasks and structured shift rotations within their respective plant units. Each subject was interviewed individually in a private room shared with the interviewers only. Informed consent was sought from each participant. All subjects from both sites used the Honeywell TDC 3000 systems as their DCS, and were thus familiar with the DCS' built-in Trends display setup shown in Figure 5.2. Although each site has different variants and customizations in their DCS suited for their process plants, all the subjects were familiar with the DCS' general interface, functions and capabilities.

Fig. 5.2. A screenshot of the Trends displayed in the Honeywell TDC 3000 DCS.

5.8 RESULTS & DISCUSSION FROM INTERVIEWS

Different Trend Data Presentation

During a normal, healthy plant operation, trends are monitored in addition to graphics and alarm summary displays. At the sites visited, two of the four DCS console screens were usually devoted to trends. However operators have different opinions as to the optimum number of trends that should be displayed on one page at any given time. No guidelines regarding this issue currently exist, and the differing preferences suggests that research needs to be done at more sites beyond the boundaries of South-east Asia before an industry-generalized guideline can be established.

8 trend lines per page. Most DCS provide this option of displaying 8 trend information on one page (Figure 2), which allows operators to draw associations and comparisons between these trends. This gives the operators the flexibility to bring up whichever variable that is of interest and extract data values and contextual inferences. However,

these pages have a maximum limit of 8 trends per page, and users may need to switch between other trend pages to monitor beyond these 8 trends and understand the bigger picture. Process control typically requires operators to manage over a hundred critical process variables, and operators who preferred this display did not find it a hassle to periodically switch between pages using hotkeys on the consoles.

32 trend lines per page. One of the sites we visited had a customized display option, which summarizes 64 of the most important process variables within 2 pages, where each page is divided into 4 quadrants, and each quadrant contained 8 trends. This option allows operators to get an “overview” glimpse of the plant’s health status. The selection of these 64 trends were pre-determined by senior operators and instrument engineers, and other trend data cannot be added or exchanged in place with these 64 trends. Other potential challenges in using this display format include display clutter, as well as the difficulty in establishing relationships and interactions between the displayed parameters.

Using Historical Trends as Additional Aid

In addition to the DCS trend displays, operators also reflected the use of complementary systems which log trend information going back as long as three years. This separate trend display is known as **Historical Trends**, and is usually set up in a PC computer located close to the DCS console. The PC provides added support to the DCS, as the DCS stores trend information for a limited amount of time—up to 96 hours. After which the data will typically be compressed and lose its resolution. Hence trend data that originated a month ago may not have the required detail in the DCS that might be crucial to the operators. Hansen (1995) and Spenkelink (1990) noted in their researches that historical information in the form of trends hampers early detection of current abnormal events. However, as Historical Trends are presented on a separate display, we assume that such negative effects were mitigated as they do not interfere with the current trend displays, and only come into play when required. Operators we interviewed generally agreed that accessing Historical Trends greatly facilitated diagnosis and troubleshooting.

Framing the Proactive Monitoring Decision Process

Fig. 5.3. The Human Intervention Framework

The Human Intervention Framework (Figure 5.3) is a model proposed by the Chemical Manufacturers Association which is adapted from models describing supervisory control by Jens Rasmussen, Tom Sheridan, and David Woods (example see Rasmussen & Goodstein, 1987) The decision process in proactive trend-monitoring found in this study can be mapped using this Human Intervention Framework. This framework simplifies the number of stages used to describe cognitive behavior underlying intervention activities. Putting it in perspective:

- **Initiating Event** Trend deviation or abnormal fluctuation
- **Orienting** Detect abnormal data fluctuation
 If trend pattern or meaning is recognized, skip to *Acting*
- **Evaluating** Compare with other information: Alarms Summary, Graphics etc.
 Hypothesizing the problem, root cause
- **Acting** Take corrective actions
- **Accessing** Monitor the plant to ensure desired condition is met

Operator Skills

Expert operators were analyzed to have strong **process knowledge**. Having this process understanding allows operators to anticipate the components that are interlinked with one another during plant upset events. While the Graphics displays provide information for process mapping, expert operators were confident enough not to refer to the Graphics displays for spatial information. Furthermore, Graphics displays do not show the finer network of components and linkages that contribute to the product flow within the process plant.

Expert operators also possessed rich **experiences in abnormal situations**. They noted that much of their job skills were attained over time, where they were exposed to new situations while learning expert techniques and efficient problem diagnosis. Experiences of abnormal situations increase operators' skills in monitoring trends by:

1. fostering new habits in reading trends
2. creating memories of abnormal trend patterns and corrective procedures
3. understanding relationships between data variables and how they affect one another during upsets

These experts commented on the challenges that less-proficient operators would face, describing the lower skills levels that are affecting their performances—poor process knowledge, poor ability to draw relationships between various variables, etc. Oftentimes operators with weaker process knowledge would monitor the Graphics displays due to their poor spatial knowledge of components and weak mental model of the dynamic system. They tend to rely on the Alarms system to inform them of anomalies, and thus are less capable of monitoring proactively.

Trend Patterns

Operators rely substantially on trend patterns in their diagnosis. Hajdukiewicz and Wu (2004) described many forms of trend patterns found in such time-based displays. This study further identified two context-based trend patterns that operators relied on. These trend patterns are used mainly for reference and comparison between the present situation and the past:

1. **Repeated patterns** are re-occurring, common sequences during plant operation (perceptual distinctions: anomaly is identified when trend deviates from regular pattern sequence)
2. **Event-specific patterns** are documented trend information for a particular event or upset (cognitive distinctions: anomaly is recognized when the trend pattern tallies with previously-occurred anomaly of the same nature)

Expert operators acknowledged that they often compare trend patterns. They noted the usefulness of these patterns when made available to novices, and that novices' confidence in their diagnosis would decrease greatly if past trend comparisons were not made.

5.9 QUALITATIVE INVESTIGATION TWO: SUMMARY

Current process control lack explicit predictive displays, but operators can derive predictive cues through Trends. Trend displays provide emergent features which inform console operators of potential problems. Experienced operators also rely heavily on trend patterns for situation diagnoses, and the temporal nature of this information allowed the operators to compare the displayed details with historical trend data. It is also due to this temporal characteristic that trend information require time lapse for problematic cues to become perceptually evident. Operators need to remain vigilant and proactive in picking up these cues if they want to take early remedial actions.

5.10 OVERALL SUMMARY

The two field studies provided valuable insights on how the experimental studies should be approached. Through interacting with the process system, operators were able to gain a mental representation of how the plant components were laid out. Getting to know the plant unit physically out in the field supported their visualization and mental simulation of how the process was running and what actions should be taken. Handling upsets and situations such that they were tasked to bring the process back and maintain it in a stable state also facilitated this establishment of the mental model. The use of trend displays allowed rate-of-change information to be elicited, a crucial cue that aided operators in anticipating what might happen in the near future. Trend patterns also provide an overview glimpse of whether the production process was as expected or off-normal. However, it was also highlighted that novice operators may not pick up these trend patterns or off-normal cues easily, cues which expert operators were known to rely on to stay proactive.

Trend displays thus appears to be a viable display platform to begin the development of a prototype predictive aid. As rate-of-change cues would require a span of time to pass before they become more obvious, designing visual aids that explicitly present rate-of-change information should hypothetically allow for faster detection and reaction by the operator. The next chapter thus reports on a simulator study which explored having explicit predictive cues on trend displays.

5.11 REFERENCES

- Bainbridge, L. (1986). What should a 'good' model of the NPP operator contain? In *Proceedings of the International Topical Meeting on Advances in Human Factors in Nuclear Power Systems*, American Nuclear Society: Knoxville, Tennessee.
- Bainbridge, L. (1992). Mental models in cognitive skill: The example of industrial process operation. In Y. Rogers, A. Rutherford, P. Bibby (Eds.), *Models in the Mind*. London: Academic Press.
- Blandford, A., Wong, B. L. W., Connell, I., Green, T. (2002). Multiple viewpoints on computer supported team work: A case study on ambulance dispatch. *People and Computer XVI—Memorable Yet Invisible: Proceedings of HCI 2002*, 139-156. Springer.
- Carswell, C.M. & Wickens, C.D. (1996). Mixing and matching lower-level codes for object displays: Evidence for two sources of proximity compatibility. *Human Factors*, 38, 1-22.
- Edwards, E. & Lees, F.P., (1974). *The Human Operator in Process Control*. London : Taylor and Francis.
- Gentner, D. & Stevens, A.L., (1983). *Mental Models*. Hillsdale, NJ: Erlbaum.
- Goodstein, L.P., (1982). An integrated display set for process operators. In *Proceedings of the IFAC/ IFIP/ IFORS/ IEA Conference on Analysis, Design and Evaluation of Man & Machine Systems*, Baden-Baden, FR Germany.
- Hajdukiewicz, J. & Wu, P (2006). Beyond trends: A framework for mapping time-based requirements and display formats for process operations. In the *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*, 1885-1889.
- Klein, G. (1998). *Sources of Power: How people make decisions*. MIT Press: Cambridge, MA.
- Klein G.A., Calderwood, R., MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 462-472.
- Lind, M., (1982). *Multi-Flow Modelling for Process Plant Diagnosis and Control*, Riso Report Riso-M-2357 (Roskilde, Denmark: Riso National Laboratory).
- Meyer, J., Shinar, D., Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39, 268-286.
- Moallem, M. (1998). An expert teacher's thinking and teaching and instructional design models and principles: An ethnographic study. *Educational Technology Research & Development*, 46, 17-37.
- Ollis, S., Macpherson, A. and Collins, D. (2006) Expertise and talent development in rugby refereeing: An ethnographic enquiry. *Journal of Sports Science*, 24, 309-322
- Pomerantz, J. R. (1986). Visual form perception: An overview. In E. C. Schwab & H. C. Nusbaum (Eds.), *Visual perception: vol. 2. Pattern recognition by humans and machines*. New York, NY: Academic Press.
- Rasmussen, J. & Goodstein, L. P. (1987). Decision support and supervisory control of high-risk industrial systems. *Automatica*, 23, 663-671.

Seamster, T. L., Redding, R. E., Cannon, J. R., Ryder, J. M., Purcell, J. A. (1993). Cognitive Task Analysis of Expertise in Air Traffic Control. *International Journal of Aviation Psychology*, 3, 257-283.

Sheridan, T. (1976). Toward a general model of supervisory control. In T. Sheridan and G. Johanssen (Eds.), *Monitoring Behavior and Supervisory Control*, New York: Plenum Press.

Smith, A., Goodwin, D., Mort, M., & Pope, C. (2003). Expertise in practice: an ethnographic study exploring the acquisition and use of knowledge in anaesthesia. *British Journal of Anaesthesia*, 91, 319-328.

Wickens, C. D., & McCarley, J. (2008). *Applied attention theory*. Boca-Raton, FL: Taylor & Francis.

Wong, W. B.L. & Blandford, A. (2001). Situation awareness and its implications for human-systems interaction. In W. Smith, R. Thomas & M. Apperley (Eds.), *Proceedings of the Australian Conference on Computer-Human Interaction OzCHI 2001*, 181-186. Perth, Australia: CHISIG, Ergonomics Society of Australia.

Chapter Six

Simulator Study One: Predictive Cues on Trend Displays

6.1 OVERVIEW

As revealed from the previous chapter, deriving cue-based prediction requires a mental representation of the system as well as perceiving the right predictive cues. These cues may not be easily perceived in process control, given the complexity of the systems, hundreds of process variables, and the lack of predictive automation. It was also found that expert operators with tens of years of experience are able to anticipate process deviations and perform timely maneuvers to maintain the production within limits, this despite having to deal with the sluggish nature of the process. During process monitoring, operators are aided by Trends, trend displays found in the distributed control system (DCS) which plot out data over time. Trends reveal patterns that provide contextual information, and also have emergent features to indicate process deviations. Reports from case studies as well as from Hajdukiewicz & Vicente (2002) noted the value of rate-of-change information towards anticipatory control and proactive behavior, yet this information requires effort and attention through constant monitoring of Trends.

An experimental study was proposed to understand whether explicitly presenting some form of predictive cue in a process control display would improve operators' anticipatory performance. Popular existing statistical process control models such as Moving Averages, Cumulative Sum (CUSUM), and Batch Means charts are in fact utilizing rate-of-change (ROC) to derive predictions to drive automated control. Manually deriving ROC information can be challenging, given that it is not explicitly made available to the operator, and that the operator has to extrapolate ROC from noisy data readings. As Endsley (1995) have pointed out, perceptual cues require attention, which can be hindered by a lack of the cues' salience or discriminability. ROC interpretation can also be time-consuming, as it is not obvious during the start of a deviation and thus requires the operator to monitor over time before emergent features become more definitive. Based on the graphical summary (see Chapter Two),

it is believed that an improvement in cue perception (Level 1SA) should facilitate predictive performance, and ultimately support proactive operator behavior.

6.2 RATE-OF-CHANGE REPRESENTATION

Rate-of-change information has proven to be control in control-based domains, including process control. Lintern, Kaul & Collyer (1984) demonstrated a modified aircraft landing system display that provided descent-rate cues for pilots to land with less error on aircraft carriers. It was reported that the test pilots preferred the landing display that additionally presented descent-rate information. Bellenkes, Wickens & Kramer (1997) revealed that experienced pilots frequently scanned the vertical velocity indicator during incidents of simultaneous altitude and heading changes which improved their dynamic mental model and situational awareness. In process control, Hajdukiewicz & Wu (2006) noted the tendency for operators to seek out information that portrayed variable changes: Did a change occur? Which direction was the change? What was the rate of change? etc. Today's operators try to answer these questions by interpreting trend displays and paying particular attention to trend patterns that provide cues toward rate-of-change.

Given the absence of explicit predictive displays in process control industries, it would thus be interesting to see whether adding explicit graphic or numeric representations of ROC to Trends would significantly improve process control performance. The comparison between graphical versus numerical data representation towards decision-making has been explored in display designs as well as risk management, revealing a general advantage in graphic data representation over numeric. When two or more data points or sets are to be compared (e.g.: rate-of-change of a parameter), numeric displays require the user to look up several entries and make calculations and visualizations, a mental task made easier when viewing graphic displays (Meyer, Shinar, Leiser, 1997; Lohse, 1991). In an experiment on the benefits of a novel anesthetic monitoring system (Charabati, Bracco, Mathieu, Hemmerling, 2009), the numerical format did not provide as much performance benefits as its graphical variant (with the mixed numeric-and-graphic representation showing the best results). Graphic data representations also supported risk avoidance

when making decisions like whether to evacuate a city due to an oncoming hurricane (Schwartz & Howell, 1985), whether hazards like gas bubbles in gas and oil drilling might potentially occur (Peacock, Schlegel, Dorman, 1983), or whether to implement safer but more expensive tires (Stone, Yates, Parker, 1997; Schirillo & Stone, 2005). Chua, Yates and Shah (2006) noted that graphs ignite stronger negative associations with riskier options and outcomes, therefore facilitating risk avoidance behavior. Tufte (1983) highlighted the advantages of incorporating numeric displays for reading specific values, as well as the salience effect that graphic modes have on dynamic, changing data patterns. However, it would be valuable for display designers if numerical cues are beneficial too. Given the vast amount of information required in process control (and resultant clutter in the distributed control system displays), it might be easier to integrate numeric than graphic ROC information into the displays.

6.3 METHOD

The “Honey Mixer” simulator

A virtual process system was designed and used for a computer-based experiment, known here as the “Honey Mixer” (Figure 6.1). This interactive microworld aimed to be simple enough for easy understanding and quick training particularly for university students, but complex enough to reflect the difficult nature of process control monitoring tasks. It also provides manipulation of variables for varying complexity and scenario difficulty. Microworld simulations have been used in research involving the interaction between operators and complex systems (Reising & Sanderson, 2002; Pawlak & Vicente, 1996; Vicente & Rasmussen, 1990). Thus the design of the Honey Mixer system is adapted from the reviews of these micro-systems. The behavior of the system and its components also reflected the nature of the products (e.g.: thicker viscosity of honey), as well as the characteristics of process control (lags, sluggish controls, inter-related variables etc.).

Fig. 6.1 The schematic layout of Honey Mixer

The purpose of the Honey Mixer is to mix honey with water to create a mixture product, before sending it into a storage tank and onward to packaging. Cold Water is first sent through the Heater to be warmed up prior to entering the Mixer. The temperature in the Heater is governed by the amount of Fuel Gas used, which indirectly commands the temperature of the heated water coming out of the Heater. At the same time, Raw Honey goes through the Heat Exchanger to be pasteurized before being mixed in the Mixer. Similar to the Heater, the temperature in the Heat Exchanger is controlled through the flow of Hot Water Supply, thus manipulating the temperature of the pasteurized honey. As the fluids enter the Mixer, the operator will be able to monitor the level, density and temperature of the mixture. This mixture within the Mixer can either flow out via the Packaging Line, or into an auxiliary Tank managed by the operator. The operator has controls of how much mixture enters or leaves the Tank. Lastly, additional honey-water mixture can be added into the Mixer through the Recycle Line.

The subject will be presented with a stabilized system, and his goal is to ensure that all the variables in the system do not exceed pre-defined operating limits. In particular, he is to ensure that the mixture level, density, and temperature are within “on-specification” requirements. Periodically disturbances via the Packaging Line and Recycle Line (marked red in Figure 6.1) will disrupt the process equilibrium, and the operator has to adjust the control variables accordingly so as not to breach any limits. Throughout the process the operator will also be given verbal commands to achieve secondary goals using the auxiliary Tank, namely to either keep the mixture in the Tank at a specific level or to maintain a certain flow rate out of the Tank.

As with most process control systems, manipulating the controls involve much consideration regarding other related components. An example is a scenario where the mixture in the Mixer is being drained out faster than it can be replenished. The operator would thus have to increase the flow rate of the heated water and the pasteurized honey into the Mixer, within appropriate ratios to meet density requirements.

Simulator’s control interface

The operator will be interacting with the virtual system through a display that mainly shows trend data as well as other control information (Figure 6.2). In each trend display the operator can cycle through various trend data by selecting the appropriate tab (Label 1). For the key variables (Mixer temperature, Mixer density, Mixer level and Tank level), predetermined parameter limits are set in the system and indicated in the trend plot via a pair of parallel red lines (Label 2). When these variables are within the limits and thus are “on-spec”, their respective tabs would appear green. Conversely when the variables exceed limits, their tabs would turn red. The operator would then make the required adjustments in the Heater, Heat Exchanger, and Tank in order to meet on-spec requirements and stabilize the system (Label 3). The operator can also use a second trend display for additional monitoring of variables (Label 4).

Fig. 6.2 The experimental display console, which shows:

1. Tabs of currently selected (grey), on-spec (green) as well as off-spec (red) variables
2. Trend line of selected variable (in this case, the level in the Tank)
3. Control console for Heater, Heat Exchanger, and Tank
4. Second trend display for operator monitoring use

Besides the default display, operators will also be presented with variants that include current rate-of-change information of the key variables, either in the form of a graphical straight-line slope (Figure 6.3) or in numerical format (Figure 6.4). Both predictors provide cues on the variables' direction of change (inclining or declining slope / positive or negative number) as well as the degree of change (steepness of slope / magnitude of number). While the numbers represented in the numerical rate-of-change format are directly derived from calculations in the system, the exact meaning of the numerical value serves no purpose to the operator. Instead, operators would take advantage through the varying magnitudes of the values during comparisons: a 1.4 incline is far steeper than a 0.14 incline. The numerical format explores its viability as an alternative rate-of-change representation, one which requires less screen estate.

Fig. 6.3 Experimental display console showing graphical rate-of-change slope.

Fig 6.4 Experimental display console showing numerical rate-of-change slope.

Research questions

The increase in salience of the rate-of-change cue should induce better predictions and allow for more “proactive-like behavior” from the operator, such that the operator would make process maneuvers earlier and better able to keep the process within operating limits. However, Hart & Wickens (1990) noted that prediction requires substantial mental resources, and people tend to be more proactive when workload was modest. As such, workload in terms of scenario difficulty was also factored into the experimental design and validated by comparing the baseline conditions between the high- and the low-workload scenarios. Specific research hypotheses (as illustrated in Figure 6.5) include:

- 1) Operators perform better during low-workload than high-workload setting;
- 2) In general, ROC cues benefitted operators in the high-workload setting;
- 3) Within high-workload scenario conditions, the presence (versus absence) of graphical ROC cue will result in lesser duration outside the operating envelope (i.e.: more alarms);
- 4) Within high-workload scenario conditions, the presence (versus absence) of numerical ROC cue will result in lesser duration outside the operating envelope;
- 5) True to other graphical-versus-numerical studies, that graphical ROC cue is more advantageous than numerical ROC cue;

Fig 6.5 The presence of ROC indicators should reduce duration outside operating envelope, with graphical being more beneficial than numerical visualization, and these benefits should be more evident in hard than in easy scenarios.

Experimental design

This study utilized a 2-by-3 factorial mixed-plot design featuring two independent variables: difficulty levels and ROC cues (Figure 6.6). Operators were randomly divided between the two workload groups, and will randomly undergo three different display conditions: Trends display with no explicit ROC cues, Trends with numeric ROC, and Trends with graphic ROC.

		Cue Representation		
Difficulty		None (Baseline)	Numeric Cue	Graphic Cue
	Easy			
	Hard			

Fig. 6.6. The 3 x 2 factorial experiment design

Easy and hard scenarios were manipulated through the magnitude of change during each disruption by the Packaging and Recycle Lines, as well as through the narrowing of the control limits for hard scenario conditions. While the frequency for disturbances to occur will be the same in both settings, a disturbance in the hard scenario meant a more drastic change in the input and output rate by the Packaging and Recycle Lines. Easy scenarios, although featured less drastic changes, shared the same frequency of change as difficult scenarios. In addition, narrower control limit bands naturally increased the challenge in keeping the process under control. Three easy and three hard scenarios were designed, each lasting thirty minutes, and randomly assigned to accompany the respective display condition. A series of pilot-testing was conducted using just the baseline display format to test, modify, and eventually select the final six scenarios that were deemed similar in their respective difficulty levels.

Dependent variables

Operator performance would be quantified by the amount of time the three critical variables exceeded beyond the pre-defined limits. A good performance would entail low amounts of limit breach, while a bad performance would mean multiple

limit breaches occurring simultaneously over long periods of time. The upper and lower limits are kept reasonably close together, allowing only for minor fluctuations and requiring the parameter to be kept in the middle of the limits in case of sudden disruptions. As such, subjects need to constantly page through the different parameters and monitor for possible disturbances, while at the same time trying to “optimize” the process by staying as close to the middle between the limits as possible. Figure 6.7 illustrates the dependent variable (amount of limit breaches) and the spectrum of performance to be anticipated by the subjects.

Fig. 6.7 The dependent variable and the spectrum of performances expected from experimental subjects

A brief questionnaire survey was administered at the end of the whole experiment (Appendix A) which polled the participants’ feedback about the experimental tasks and the various display conditions. The survey provided participants’ subjective ratings of the different display and difficulty conditions, and can be used to compare between the two independent groups differentiated by difficulty levels. Aside from quantitative measurements, the survey was another form of validating the difference in difficulty levels between the two conditions. Within hard scenarios, we expected participants to appreciate the presence of ROC cues more than those who experienced easy scenarios.

Subjects

A total of 44 students (8 males, 36 females, mean age 21.6) from Nanyang Technological University, Singapore were recruited for this study.

Procedure

After informed consent and demographic information were collected, subjects were randomly assigned to either the easy or hard difficulty group, and underwent a 20-minute training and practice session. Thereafter, subjects performed the three display conditions in random order, and randomly featuring one of the three easy or hard scenario settings.

6.4 RESULTS

Figure 6.8 shows a data plot with error bars on the mean durations of limit breaches for each experimental condition. The graph for easy condition appeared much flatter than that of the difficult condition. The Mauchly's Test of Sphericity was performed on each of the two independent groups (difficulty level), revealing no violation of sphericity. A between-subjects Analysis-of-Variance was first conducted between the two baseline groups (easy versus hard difficulty), revealing significant differences, $F(1,42)=19.184$; $p<0.001$, thus indicating that hard scenarios were appropriately designed to be more challenging than easy scenarios. Subjects performed worse when faced with hard scenarios as compared to easy scenarios (H_1 validated).

Fig. 6.8. Mean duration of limit breach for each condition. Error bars indicate 1 standard error.

As an overall analysis, a repeated-measure ANOVA with difficulty condition as a between-subject factor was conducted on the entire data to examine for main and interaction effects. Results revealed significant within-subject main effect of display type, $F(2,42)=4.32$, $p<0.05$, as well as significant between-subject main effect of difficulty condition, $F(1,42)=70.27$, $p<0.01$. No interaction effect was found, $F(2,42)=1.73$, $p=0.18$. There were differences in performance due to the different display types as well as difficulty condition, and although no interaction effects were found, subsequent cell means will be compared to answer planned, specific hypotheses.

Two Repeated-measure ANOVAs were conducted which revealed distinctions in performance between the three display conditions for hard scenarios ($F(2,42)=3.387$; $p<0.05$), and no significant differences for easy scenarios ($F(2,42)=1.033$; $p=0.365$). Figure 4.8 illustrated the pattern of improving control performance for hard scenarios given the presence of ROC cues. As such, the ROC cues benefitted more for the participants who experienced the hard scenarios than for those in the easy scenarios (H_2 validated).

We were interested to study whether presenting either a linear or numerical ROC visualization would *improve* operator performance over not presenting anything at all. Two planned one-tail t-tests were conducted for the baseline-linear pair, $t(21)=2.903$, $p<0.01$, as well as the baseline-number pair, $t(21)=1.66$, $p=0.056$. The presence of both ROC visualization cues aided the operator's performance during hard scenarios, as compared to without having any ROC visualizations at all (H_3 & H_4 validated).

The performance effects between linear and numerical ROC forms were analyzed to see if they were statistically different. A planned t-test was performed between linear and numerical visualizations, $t(21)=0.927$, $p=0.364$. Results thus indicated that operators did not perform differently when using either the linear or numerical ROC visualization (H_5 rejected).

The non-parametric Mann-Whitney U Test was conducted on the survey results comparing between the two independent groups based on difficulty levels (Table 6.1). Using two-tailed exact p-values and α -level of 0.05, the following statements were found to generate different responses between participants from the easy and hard scenarios. Notably, Statements 3, 4 and 12 reinforced the differences between the easy and hard scenarios (H_1) while Statements 7 and 10 reiterated the value of ROC cues during hard scenarios (H_2).

#3: The process scenarios were difficult to manage

Easy – majority disagreed

Hard – responses were mixed

#4: There were too many things to look out for during the scenarios

Easy – majority mildly disagreed

Hard – majority mildly agreed

#7: It was difficult to spot process changes without any predictors

Easy – majority mildly agreed

Hard – majority strongly agreed

#10: I did not need the Predictors to perform well

Easy – majority mildly disagreed

Hard – majority strongly disagreed

#12: Right now I do not feel fatigued after completing all the scenarios

Easy – responses were mixed

Hard – majority spanned between “neutral” and “strongly disagree”

Table 6.1. Percentage of people who responded to each survey statement, compared between the two independent groups (difficulty level). Mann-Whitney U Test results with exact p-values lower than α -level 0.05 are bolded and in red ink.

Percentage of people who rated a score of _____			Strongly Disagree							Strongly Agree		Exact Sig. (2-tailed)
			1	2	3	4	5	6	7			
1	I found the production process easy to understand.	Easy	0.0%	0.0%	0.0%	13.6%	36.4%	40.9%	9.1%	.567		
		Hard	0.0%	0.0%	4.5%	9.1%	45.5%	36.4%	4.5%			
2	I understood how the interface worked.	Easy	0.0%	0.0%	0.0%	4.5%	22.7%	45.5%	27.3%	.273		
		Hard	0.0%	0.0%	9.1%	9.1%	9.1%	63.6%	9.1%			
3	The process scenarios were difficult to manage.	Easy	13.6%	45.5%	9.1%	13.6%	13.6%	4.5%	0.0%	.006		
		Hard	0.0%	22.7%	9.1%	27.3%	18.2%	13.6%	9.1%			
4	There were too many things to look out for during the scenarios.	Easy	9.1%	18.2%	36.4%	13.6%	13.6%	9.1%	0.0%	.004		
		Hard	0.0%	9.1%	9.1%	22.7%	36.4%	9.1%	13.6%			
5	I could have managed the process better.	Easy	0.0%	4.5%	4.5%	36.4%	27.3%	27.3%	0.0%	.069		
		Hard	0.0%	0.0%	13.6%	13.6%	13.6%	45.5%	13.6%			
6	The recycled line and mixer output changed too frequently.	Easy	9.1%	31.8%	31.8%	18.2%	9.1%	0.0%	0.0%	.789		
		Hard	13.6%	27.3%	22.7%	22.7%	4.5%	0.0%	9.1%			
7	It was difficult to spot process changes without any Predictors.	Easy	0.0%	18.2%	9.1%	9.1%	40.9%	4.5%	18.2%	.015		
		Hard	0.0%	0.0%	4.5%	18.2%	4.5%	40.9%	31.8%			
8	The Line Predictor was easy to understand	Easy	0.0%	0.0%	0.0%	0.0%	9.1%	27.3%	63.6%	.674		
		Hard	0.0%	0.0%	0.0%	9.1%	0.0%	18.2%	72.7%			
9	The Number Predictor was easy to understand	Easy	0.0%	4.5%	0.0%	0.0%	27.3%	40.9%	27.3%	.383		
		Hard	0.0%	0.0%	4.5%	9.1%	18.2%	59.1%	9.1%			
10	I did not need the Predictors to perform well	Easy	9.1%	27.3%	27.3%	18.2%	13.6%	4.5%	0.0%	.001		
		Hard	31.8%	45.5%	22.7%	0.0%	0.0%	0.0%	0.0%			
11	It was easy to anticipate future process values without Predictors.	Easy	13.6%	22.7%	40.9%	9.1%	9.1%	0.0%	4.5%	.066		
		Hard	27.3%	45.5%	9.1%	9.1%	4.5%	4.5%	0.0%			
12	Right now I do not feel fatigued after completing all the scenarios.	Easy	0.0%	9.1%	27.3%	4.5%	22.7%	27.3%	9.1%	.016		
		Hard	22.7%	13.6%	13.6%	27.3%	13.6%	4.5%	4.5%			
13	I enjoyed the entire experiment.	Easy	0.0%	4.5%	0.0%	22.7%	27.3%	31.8%	13.6%	.838		
		Hard	9.1%	0.0%	0.0%	9.1%	40.9%	36.4%	4.5%			

It is interesting to also point out that majority of respondents from both the easy and hard scenario groups agreed strongly (67%) to Statement 8: “The Line Predictor was easy to understand”, as opposed to only 38% who felt that way about the number predictor. Similar support for the line predictor was found when participants were forced to rank-order their preferences between having line, number or no predictor, as 77.3% of participants in the easy condition and 90.9% of participants in the hard condition ranked the line predictor on top.

6.5 DISCUSSION

This study explored whether operator performance will be improved when explicitly presenting rate-of-change, an information that is otherwise only implicitly derived in today’s process control interfaces. Overall, the presence of ROC cues in any form did improve operator performance during high workload. This advantage was evident even though trend displays were known to implicitly provide rate-of-change information. The benefits of graphically-illustrated ROC have been reiterated by various literature in Ecological Interface Design (Burns & Hajdukiewicz, 2004; Vicente & Rasmussen, 1990). Considering that process control operators have to face digital information from hundreds of parameters, presenting data in graphical format should aid operators in detecting deviations better. While our quantitative findings did not support this claim, responses from the survey hinted the participants’ preference for graphical versus numerical representation.

True to our expectations, proactive monitoring was more challenging to achieve during high workload, and technological aids are thus more effective during such situations. Prediction requires much mental resources which can be limited in supply during high workload (Hart & Wickens, 1990). The Honey Mixer was designed to instill workload through the high magnitude of changes and the narrow alarm limits, thereby requiring the operators to figure out remedial actions under time pressure, and in the process consider the ripple effects their control maneuvers may have on other related parameters. Although this study showed that the benefits of the ROC cues did not bring operator performance close to low-workload situations, the improvement in control performance versus not having any cues at all was still significant.

Yet despite the challenging hard scenarios, the participants did not have to monitor a variety of visual displays unlike industrial process control systems, and thus the effects between graphical and numerical representation may not be effectively differentiated. Ecologically-designed graphical interfaces tend to benefit most when operators have to scan multiple screens as well as multiple pages within each screen. Our one-screen setup, and which has very few digits, may not fully tease out the difference in effort for processing graphical versus numerical data representation.

We anticipated practice effects given the short training time during the course of the experiment, and mitigated them through the randomization of scenarios and display condition orders. The Honey Mixer process was so simple that with extensive practice, one should be able to easily manage the process with or without the ROC aid. The strong performance benefits despite the short training procedure suggested possible effects toward bridging the expert-novice gap. In general, participants who had lesser exposure to the simulator still performed better when presented with the linear-format display.

Evidently, explicit predictive cues incorporated into a simulated process control display did support proactive monitoring behavior and improved performance. The rate-of-change calculation appeared to be a viable algorithm to support the development of process control predictive displays. Surveying existing predictive applications in the process control industries should help us in coming up with ways to further improve our predictive algorithm.

6.6 REFERENCES

- Bellenkes, A., Wickens, C.D., & Kramer, A. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space and Environmental Medicine*, 68, 569-579.
- Charabati, S., Bracco, D., Mathieu, P.A., Hemmerling, T. M. (2009). Comparison of four different display designs of a novel anaesthetic monitoring system, the 'integrated monitor of anaesthesia (IMATM)'. *British Journal of Anaesthesia*, 103, 670-677.
- Chua, H.F., Yates, J.F., and Shah, P. (2006). Risk avoidance: Graphs versus numbers. *Memory and Cognition*, 34, 399-410.
- Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37, 32-64.
- Hajdukiewicz, J. & Wu, P (2006). Beyond trends: A framework for mapping time-based requirements and display formats for process operations. *Human Factors and Ergonomics Society Annual Meeting Proceedings 2007*, 1885-1889.
- Hart, S. G. & Wickens, C. D. (1990). Workload Assessment and Prediction. In HR Booher (Ed). *Manprint: An integrated approach to systems integration* (pp. 257-296). New-York: Van Nostrand.
- Lohse, J. (1991). A cognitive model for the perception and understanding of graphs. In *Human Factors in Computing Systems-Reaching Through Technology, CHI '91 Conference Proceedings*, 137-144. New York: Association for Computing Machinery
- Meyer, J., Shinar, D., Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39, 268-286.
- Pawlak, W. & Vicente, K. (1996). Inducing effective operator control through ecological interface design. *International Journal of Human-Computer Studies*, 44, 653-688.
- Peacock, B., Schlegel, R., Dorman, R. (1983). Dynamic drilling displays. In *Proceedings of the Human Factors and Ergonomics Society 27th Annual Meeting*. 445-448
- Reising, D. V. C. & Sanderson, P. M. (2002). Work domain analysis and sensors II: Pasteurizer II case study. *International Journal of Human Computer Studies*, 56, 597-637.
- Schirillo, J. & Stone, E. (2005). The greater ability of graphical versus numerical displays to increase risk avoidance involves a common mechanism. *Risk Analysis*, 25, 555-566.

Schwartz, D. R., & Howell, W. C. (1985). Optional stopping performance under graphic and numeric CRT formatting. *Human Factors*, 27, 433–444.

Stone, E.R., Yates, J.F., Parker, A.M. (1997). Effects of numerical and graphical displays on professed risk-taking behavior. *Journal of Experimental Psychology: Applied*, 3, 243-256.

Tufte, E.R. (1983). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press

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.

Chapter Seven

Simulator Study Two: Rate-of-Change Visualizations

7.1 OVERVIEW

After showing the proof-of-concept that increased predictive cue salience has a positive effect on predictive behavior and proactive-like performance, the interest is now in developing and testing a **rate-of-change algorithm** that would support proactive control performance, yet simple enough for industry acceptance and implementation (see Chapter Four). There is also value in exploring how best to visualize the predictive cue in a **realistic overview schematic display**, which unlike trend displays tend not to possess any predictive features. Enhancing the schematic display should aid novice operators (who tend to rely more on schematic displays due to their weaker mental models of the system) in eliciting proactive monitoring behavior as well as anticipatory control performance. An experimental test of **different fundamental visualization designs** would provide the foundation for future predictive display objects developed by the industries.

7.2 AN UPDATED RATE-OF-CHANGE ALGORITHM

Following the clues in Simulator Study One, the rate-of-change concept is adopted and refined in consideration of the industry concerns regarding predictive algorithms. The rate-of-change of a parameter can be derived through comparing the parameter's readings over a moving time window. For example, if the current reading is higher than the reading thirty seconds ago, it would appear to be a positive rate-of-change. As this method would be vulnerable to a noisy parameter (one which has minute fluctuations over a general trend), a filter will need to be applied to level out the noise and bring out the general trend. This can be done through factoring in a noise filter time period. This period may be longer than the moving average window used to calculate the rate-of-change, but will vary between different parameters depending on their inherent noise. Figure 7.1 illustrates how this filtered rate-of-change (FROC) is calculated.

Fig. 7.1. Detailed breakdown of the filtered rate-of-change calculation.

The FROC algorithm differs from the rate-of-change calculation used in Simulator Study One in that unlike the latter, which was based on multiple process variables, this algorithm relies solely on the historical data of just one variable to derive that particular parameter's rate-of-change. This allows for easy configuration, and would not be affected when other related process variables are modified or become "out-of-sync" as compared to a model-based multivariate computation. The drawback of this simple setup is the inherent lag in updating the filtered rate-of-change. As the algorithm depends on immediate historical data to reflect the change, the calculated rate-of-change will be close to but not exactly at the actual rate-of-change. This Simulator Study Two will evaluate whether this FROC algorithm, given its pros and cons, will be sufficient to elicit positive operator control performance.

7.3 METHOD

The "Honey Mixer II"

This study utilized a micro-world simulator that was similar to the previous study's Honey Mixer layout, but adopted a schematic display that was similar to Honeywell International's Experion™ distributed control system (DCS) interface. Figure 7.2 shows a screenshot of this schematic display, which gave the operator an overview of the entire Honey Mixer process. There were reasons for adopting a schematic display over Trends. We were interested to see whether Trends was a useful

rate-of-change indicator in itself (versus nothing at all). Furthermore, current schematic displays do not feature any explicit or implicit predictive cues. Based on our previous ethnographic work, we discovered that novice operators in process plants rely on schematic displays more often than experts partly because they needed the schematics to supplement their limited mental models. As the schematic display provided an overview of the whole process, it would be deemed as a more appropriate setup for a monitoring task.

Fig. 7.2. A screenshot of Honey Mixer II display.

The process design in Honey Mixer II remained the same as in the previous Honey Mixer study. Water and honey were first heated up separately before coming together in the central mixer. The ideal water-to-honey ratio within the mixer was 2:1. The mixed product was then transferred into a holding tank before eventually out of the system towards downstream packaging. Periodically, additional honey product would be pumped into the system via the “Re-blended Honey” and the “Recycled Honey” flow lines. However, the concentration in the “Re-blended Honey” flow line was higher than normal, and therefore would affect both the level in the mixer as well

as the density of the product being mixed. The Honey Mixer II also incorporated minute amount of noise into its process, mimicking actual industry processes.

System's control interface

Interacting with the Honey Mixer II schematic display was also different than in the first Honey Mixer. The operator would select the component within the schematic layout, and a faceplate would appear at the side which showed the current reading, the current output capacity (i.e.: size of the valve opening in percentage), as well as the high and low alarm limits (the operating envelope). The operator keyed in the preferred setpoint, and then monitored the parameter as its current value tried to meet this setpoint. In most cases the current values would fluctuate about the setpoints in consideration to the inherent noise within the system. Whenever a parameter's current value exceeded the operating limits, a red border would appear around the parameter's reading within the schematic display along with a red square icon.

As with the previous simulator study, the goal is to ensure that critical parameters do not exceed pre-defined operating limits. The three critical parameters were the product density in the mixer, as well as the volume level in both the mixer and the holding tank. Periodically the system would receive "disturbances" through product flow into the mixer and/or holding tank from the re-blend and recycle flow lines respectively. The operator would have to adjust the flow of water and honey into the mixer as well as the flow into the holding tank so as to maintain equilibrium.

Rate-of-change visualizations

Aside from the baseline condition with no rate-of-change visualizations, there were five other display variants (Table 7.1) designed based on a progressively increasing FROC "data precision": Mini-Trends, Direction-of-change, Low-resolution Rate-of-Change, High-resolution Rate-of-Change, and the Predictive Indicator. In each display condition, one of these five features was embedded beside critical parameters in the schematics as well as within the faceplates. All of these visualizations have the same refresh rates of around one second.

Table 7.1. Five types of FROC visualization, with progressively increasing “data precision”

	<p>The Mini-Trends plotted a chart based on 12 data-points over a historical span of 5 minutes from current. This visualization represented an implicit form of FROC, in which the operator had to manually decide the rate-of-change.</p>
	<p>The Direction-of-Change (DOC) presented either an up-arrow, down-arrow, or a flat line based on the filtered rate-of-change. The directional thresholds were uniquely configured according to each parameter’s characteristics (such as noise and range of FROC) to support optimum sensitivity.</p>
	<p>The Qualitative Arrows presented five categorical information: significant increase, moderate increase, no change, moderate decrease, and significant decrease. The threshold between moderate and significant change followed a rough 20-50% range of the FROC (i.e.: given max rate-of-increase as 100% on the range, a moderate increase means a rate-of-increase that’s between 20-50% of the full possible rate-of-increase.)</p>
	<p>The Range Indicator showed the full range of FROC in a vertical dial. The full range for each parameter was uniquely configured according to its rate-of-change characteristics.</p>
	<p>The Predictive Indicator provided an explicit representation of the current reading as well as a predicted value 2 minutes into the future, based on the current filtered rate-of-change. Between the current reading and the predicted value was a trend made of historical predictions from past rate-of-change rates.</p>

The Predictive Indicator provided not just a predicted future parameter reading based on current FROC, but also presented historical predictions derived from past recorded FROC (Figure 7.3). Like the other four FROC visualizations, the Predictive Indicator had a refresh rate of around 1 second. In addition to the latest prediction made using the current FROC rate, the Predictive Indicator also logged three past predictions in 30-second intervals (i.e.: these predictions were made using the FROC at that point in history). Using the first of the three data points (highlighted using the green pointer in Figure 7.3) as an example, based on the prediction made 1 minute and 30 seconds ago, the parameter should reach that value in the next 30 seconds. Of course, the parameter may or may not reach this predicted value in the next 30 seconds, given the dynamic changes that might have occurred since that prediction was made. Nonetheless, this trend line of historical predictions provided added information regarding the parameter's behavior. The experiment would explore if such detailed information representation would provide additional value to operators.

Fig. 7.3. The Predictive Indicator

The resolution of ROC representation increased progressively in each of the display design variants. This design factor was made in consideration to the limitation in computing power of distributed control system (DCS) consoles in industrial control rooms. Even though a complex, information-rich ROC visualization should, hypothetically, be of most beneficial to process control operators, such visualizations may require a fair amount of computing power from the DCS. Draining too much computing power to process additional software applications would thus affect the timeliness of data refresh rates in the DCS displays, which in turn would make process control more difficult. It would be interesting to see from the industry's perspective how low can the quality of ROC representation be in order to still provide significant performance benefits versus having none at all.

Research questions

Similar to Simulator Study One, this experiment sought to investigate the operators' process control and prediction performance with and without FROC predictive aids, and that generally speaking our FROC algorithm is effective enough at **improving operator control performance** (i.e.: less red-line limit breaches) as well as **improving operator prediction performance** (i.e.: more accurate and rapid predictions). Through the analyses, the most effective predictive aid(s) would also be determined as a result. Specific details of each performance measurements can be found subsequently in the "Dependent variables" subsection of this chapter.

Aside from performance measurements, it may be hypothesized that any predictive information, being based upon higher derivatives (e.g., rate of change), will increase control activity; yet while higher control activity may be productive (effective, rapid predictive response to disturbances), it may also be counter-productive, a consequence of over-control and instability, which can often happen in with response to rapid changes in a lagged system (Wickens, 1986; Jensen, 1979). One of the important things that can modulate how productive an increase in control activity (if caused by the presence of a predictive aid) might be is the degree of precision and reliability of the predictive information. Thus it may also be assumed that more precise (if reliable) predictive information will render the increased control activity more

effective in minimizing out-of-tolerance error. Conversely, increased control activity might mean that the user is over-controlling, thereby resulting in worse performance and mitigating any benefits from the presence of predictors. It will be interesting to analyze our data for relationships between the amounts of **control activity** and operator performance.

Subjects

A total of 50 students from Nanyang Technological University, Singapore were recruited for this study (21 male, 29 female, mean age=21.84, SD=1.8).

Experimental design

Given that this study had one baseline condition (no ROC visualizations) and five “predictor” conditions (different ROC visualizations), a 5-by-2 factorial mixed plot was used as seen in Figure 7.4. Two scenarios were designed and each scenario lasted 24 minutes, both involving periodic changes in re-blend and recycled product flow occurring between 2- to 5-minute intervals. Subjects were randomly assigned to do either the baseline or the predictor condition first, and the two scenarios were randomly paired to either of the two conditions. This is to mitigate any possible learning effects or carry-over effects, as the introduction of the predictor before or after the baseline scenario would most likely have an effect on how the subject would view and interact with the system (e.g.: lack of visualization now may create a sense of handicap). Within-subject analyses were used between each of the five baseline-versus-ROC-visualization pairs. A between-subject analysis was used to compare across all five ROC visualizations.

Fig. 7.4. The experimental design for Simulator Study Two

Dependent variables

Operator control performance was measured using two ways: the *Alarm Presence Score*, and the duration of scenario with no alarm limit breaches.

The Alarm Presence Score is the sum of all the documented alarm limit breaches for all three critical parameters (mixer level, mixer density, storage tank level) sampled at each five-second interval. In a 24-minute (288 five-second intervals) scenario example, a good performance may have no alarm limit breaches and thus an Alarm Presence Score of zero. A bad performance may have one critical parameter at alarm limits for 24 minutes, and thus resulting in a score of 288. The worst possible performance would entail all three critical parameters exceeding their alarm limits throughout the entire scenario, and therefore leading to a score of 864 (288 x 3). Given that there could be three alarms being triggered at any given time, this scoring system allowed for performance to be quantified based on the “number*duration” of the alarms triggered (versus comparing just how long the operator kept the process alarm-free). As such, even if two operators produced the same amount of scenario duration with no alarm limit breaches, one may still fair worse than the other if the Alarm Presence Score was higher.

Operator prediction performance was quantified by measuring the accuracy of their predictions, as well as their response time needed to give a prediction. In order to derive prediction accuracy and response time, six prediction probes were developed for each scenario. During the actual experiment at specific scenario periods (but unknown to subjects), subjects were told to predict the reading of either one of the three critical parameters one minute into the future. A 1-minute span of prediction was used because during simulation calibration, it was found that it took on average approximately 1 minute after the start of a process disturbance to trigger the first alarm (i.e.: a salient visual alert) given no action was taken. This 1-minute span was further validated statistically by correlating a 30-second (6 data points) series of FROC-calculated 1-minute prediction after the disturbance occurred, against a 30-second (6

data points) series of the parameter's actual value 1 minute since each prediction was made. Results yielded a significant positive correlation, $R=0.583$, $p<0.01$, showing that a 1-minute span was sufficient enough to derive a good sense of what a parameter's future value would be.

The operator's prediction quality was derived by comparing the absolute differences between the operator's prediction versus 1) the actual parameter reading one minute later; and 2) the FROC-calculated prediction at the time the operator's prediction was made. A greater absolute difference would mean a lower prediction accuracy.

The time between the prediction probe and the verbal response by the operator is manually documented (to measure time taken to make predictions), and the difference between baseline and predictor conditions for each operator will be compared to see if response times differed significantly.

Given that data are logged every 5 seconds, **operator control activity** was calculated by summing up the total number of times the setpoints of control parameters changed. The control parameters were the water and honey flows into the mixer as well as the product flow from the mixer into the holding tank. A change was counted as whenever the setpoint for one of these parameters became different.

Procedure

After informed consent and demographic information were collected, subjects first underwent basic training and hands-on practice with the Honey Mixer II model and schematic interface. A second training focusing on a specific predictor was performed prior to the predictor condition, either before or after the baseline condition depending on the assigned random order. During both trainings proficiency was defined as the ability to maintain no alarm limit breaches for a 5-minute duration within the training scenarios. All recruited subjects managed to pass the proficiency criteria. Subsequently, the presentation order of the scenarios, the display conditions, and the assigned one of five predictors were all random.

7.4 RESULTS

Operator control performance

The participants' control performance was quantified using the Alarm Presence Score as well as the amount in their scenario durations which had no alarm limit breaches (Percentage Duration with No Alarms). A Shapiro-Wilk Test of Normality performed within each subgroup revealed no significantly off-normal distributions.

Figure 7.5 depicts the overall Alarm Present Measurements as a joint function of display group and predictor presence (vs. baseline). In this 5-by-2 factorial ANOVA (with predictor presence being a repeated measure), a marginal main effect of display group ($F(4,45)=2.02$, $p=0.08$) as well as a significant main effect of predictor presence were found. No significant interaction effects were found for Alarm Presence Scores, $F(4,45)=1.09$, $p=0.37$.

Fig. 7.5. Graph showing Alarm Present Measurement scores. Error bars indicate 1 Std. Error.

A series of between-subjects F-tests within each scenario type (i.e.: *Baseline* versus presence of *Predictors*) was conducted. Within the Baseline scenarios, there were no significant differences between the five display groups, $F(4,45)=0.993$, $p=0.421$. During scenarios with the Predictors present, there were significant

differences between the five display groups for both Alarm Presence Scores, $F(4,45)=3.787$, $p<0.01$. All 50 participants generally performed the same when not provided with any predictive aids at all. Conversely, the presence of some of the predictors elicited greater control performance benefits than others.

Although the interaction between group and condition was not significant, it is possible that this non significance was the result of exceptionally high variance in one of the display conditions. On this basis, and the fact that one of the displays (Qualitative Arrows) shows clear predictor benefits, it was decided to perform a series of planned, one-tailed pair-wise T-tests on the Alarm Presence Scores for each of the five display groups (comparing baseline versus predictor scenarios). Results revealed significant performance benefits from Qualitative Arrows, $t(9)=3.577$, $p<0.01$ and Predictive Indicator, $t(9)=1.942$, $p<0.05$. Benefits from Mini-Trends ($t(9)=0.467$, $p=0.32$), Direction-of-Change ($t(9)=0.941$, $p=0.19$) and Range Indicator ($t(9)=3.577$, $p=0.12$) were not statistically significant. The presence of Qualitative Arrows and Predictive Indicator improved the operators' control performances.

Figure 7.6 presents the data on the Percentage Duration with No Alarms. The same 5-by-2 factorial ANOVA revealed marginal effects of display group ($F(4,45)=2.09$, $p=0.09$) as well as significant effects of predictor presence ($F(1,45)=15.2$, $p<0.01$). No significant interaction effect between the two terms were found ($F(4,45)=1.52$, $p=0.21$). Similar to the *Alarm Presence Measurement* findings, The participants generally performed the same when not provided with any predictive aids at all, and that the presence of some of the predictors elicited greater control performance benefits than others.

Fig. 7.6. Graph showing Percentage Duration of performed scenarios with No Alarms. Error bars indicate 1 Std. Error.

Again, as with the Alarm Presence Measure, although the interaction between group and condition was not significant, it is possible that this non significance was the result of exceptionally high variance in one of the display conditions. On this basis, it was again decided to perform a series of planned, one-tailed pair-wise T-tests. Results revealed significant performance benefits with Qualitative Arrows, $t(9)=5.107$, $p<0.01$, as well as Predictive Indicator, $t(9)=1.81$, $p=0.05$. Marginally statistical effects were found for Range Indicator, $t(9)=1.576$, $p<0.08$. Effects from Mini-Trends ($t(9)=0.501$, $p=0.314$) and Direction-of-Change ($t(9)=0.968$, $p=0.18$) were not statistically significant. These findings further validated that the presence of Qualitative Arrows, Range Indicator and Predictive Indicator improved the operators' control performances.

Operator prediction performance

For each of the two scenario conditions (baseline versus predictor), each participant was probed 6 times to predict what a specific parameter's value would be one minute in the future. To analyze the quality of participants' prediction accuracy within each of their display groups, the absolute deviation between the predicted values and the actual values (derived one minute after the prediction was made) was calculated and averaged across all 6 probes for each participant. The average absolute deviations were first transformed using natural logarithm and then compared between the Baseline versus Predictor scenarios. Figure 7.7 plots the transformed average absolute deviation for the entire subgroup for both Baseline and Predictor scenarios.

Fig. 7.7. Graph showing transformed average absolute deviation between participants' predictions and actual parameter values 1-minute later. Error bars indicate 1 Std. Error.

The same two-way with one repeated measure ANOVA was conducted between the 5 different ROC visualizations and the 2 scenario conditions. There was no main effects for display groups, $F(4,45)=0.701$, $p=0.595$, as well as scenario types, $F(1,45)=2.628$, $p=0.112$. No significant interaction effects were found, $F(4,45)=0.123$, $p=0.974$. Despite an average increase in raw prediction accuracy of 11.4% across all predictors, the results were not statistically conclusive.

Perhaps with the help of the rate-of-change visualizations the operators were making accurate predictions at *that point in time* which thus facilitated effective control maneuvers. The operator could have predicted a deviation, thus taking remedial action within the one-minute span and resulting in a different parameter reading one minute in the future as compared to the operator's initial prediction response. It might thus be more appropriate to compare the operator's prediction with a computer-calculated prediction one minute into the future. We could use the filtered rate-of-change to derive a calculated prediction of the parameter readings given the current operation state at that point in time. The absolute difference between the participants' predictions and the calculated predictions served as a measure of accuracy. Pre-analysis checks revealed no outliers. Figure 7.8 plots the transformed average absolute deviation comparing between participants' predictions and the calculated predictions.

Fig. 7.8. Graph showing transformed average absolute deviation between participants' predictions and FROC-derived parameter values. Error bars indicate 1 Std. Error.

The 5-by-2 factorial ANOVA found significant main effect of scenario conditions, $F(1,45)=5.216$, $p<0.03$. No significant main effect of displays ($F(4,45)=0.68$, $p=0.61$) nor an interaction effect ($F(4,45)=1.506$, $p=0.22$) were found. However, analyzing the predictor-only conditions, as was done above, certain predictive aids generated significant performance effects. The Qualitative Arrows generated on average a 28% improvement in operators' predictions, and this improvement was statistically significant, $t(9)=3.07$, $p<0.01$. The Predictive Indicator

generated on average a 17.5% improvement in operators' predictions, and this improvement was marginally significant, $t(9)=1.62$, $p=0.07$. The Mini-Trends ($t(9)=0.08$, $p=0.468$), Direction-of-Change($t(9)=0.813$, $p=0.22$) and Range Indicator ($t(9)=0.97$, $p=0.17$) did not produce any significant results. With Qualitative Arrows and Predictive Indicator, operators were making more accurate predictions at the time of the probes, and these improved predictions during that moment in time altered the process' course of the future appear to have been productive.

Given the presence of improved prediction when predictive aids were present, the Predictor scenario data was used to explore *whether the quality of prediction improved control performance*. A one-tailed Pearson's Correlation analysis was conducted between participants' absolute prediction deviation (between participants' prediction and FROC-calculated values) and Alarm Presence Measure as well as Percentage Duration with No Alarms. We would expect that the *greater* the prediction deviations (i.e.: poorer predictions), the poorer the control performance would be, the *more* alarms will be triggered. Results met this expectation, revealing a moderate positive correlation between prediction deviations and Alarm Presence Measure, $r=0.281$, $p<0.05$. Expectations were also met for Percentage Duration with No Alarms (*increasingly* poorer predictions should lead to *lesser* "No Alarm" states), $r=-0.213$, but this was statistically less significant, $p=0.069$. Figure 6.9 plots each participant's average absolute prediction deviation with the respective Alarm Presence Measure score. Results showed that effective prediction improved control performance.

Fig. 7.9. Graph plotting participants' prediction deviations with their Alarm Presence Measure scores. Solid line indicates best-fit linear trend.

Next we were interested to find out whether participants would take less time to predict in the presence (versus absence) of predictor aids. An outlier check for values exceeding 2 standard deviation eliminated two outliers from the Range Indicator group. A series of Shapiro-Wilk Test of Normality was performed revealing no significant off-normal distribution. Figure 7.10 plots the average prediction time participants took in responding to prediction probes.

Fig. 7.10. Graph showing participants' average prediction time. Error bars indicate 1 Std. Error.

The 5-by-2 factorial ANOVA revealed marginal significant main effect of display, $F(4,43)=1.406$, $p=0.08$, indicating that the predictor shortened prediction time. No significant main effect of scenario conditions ($F(1,43)=0.991$, $p=0.325$) or interaction effect ($F(4,43)=1.406$, $p=0.248$) were found. A series of planned, one-tailed pair-wise T-test was conducted within each of the five display groups, and significant prediction time reductions were revealed for Qualitative Arrows (8.3% time reduction), $t(9)=1.98$, $p<0.05$, and Predictive Indicators (10.5% time reduction), $t(9)=1.964$, $p<0.05$.

Operator control activity

The number of control movements was tallied for each participant in each group and compared for informative trends. Two outliers (one from Direction-of-Change group, the other from Qualitative Arrows group) were removed, and normality tests showed no significant off-normal signs. Figure 7.11 shows the mean data for each display group separated between the baseline and predictor scenarios.

Fig. 7.11. Graph showing the number of control movements.
Error bars indicate 1 Std. Error.

The 5-by-2 ANOVA revealed significant main effects for display types ($F(4,43)=4.81$, $p<0.01$) as well as for scenario conditions ($F(1,43)=10.773$, $p<0.01$). No interaction effect was found, $F(4,43)=1.34$, $p=0.27$. As above, two-tail t-tests were conducted within each display group which revealed significantly more control activity for Mini-Trend ($t(9)=2.49$, $p<0.05$) and Direction-of-Change ($t(8)=2.74$, $p<0.05$), as well as a marginal increase for Qualitative Arrows ($t(8)=2.07$, $p=0.07$). No effects of predictor presence was found for Range Indicator ($t(9)=0.019$, $p=0.99$) and Predictive Indicator ($t(9)=0.82$, $p=0.43$). The results seem to indicate that there was increase in control movement during the presence of low-precision predictors, and as predictor precision increased, the increase in control activity with the predictors began to diminish.

7.5 DISCUSSION

This study sought to validate the benefits of the filtered rate-of-change algorithm in supporting proactive monitoring and anticipatory process control performance, as well as to explore possible FROC visualizations. The visualizations were designed such that each design is a progression from the previous design in terms of rate-of-change precision. That is to say, the rate-of-change in Mini-Trends was implied (the operator had to infer the rate-of-change), Direction-of-Change only showed the direction of the parameter's behavior, the Qualitative Arrows provided categorical rate-of-change information within a given direction, the Range Indicator displayed the full range of rate-of-change values, and the Predictive Indicator extrapolated what the future reading would be given the current rate-of-change.

The findings revealed consistent pattern of effects across dependent variables with Qualitative Arrows showing greatest benefits, Predictive Indicator showing next, and Range Indicator showing third, while no benefits at all were offered by Mini-Trends and Direction-of-Change. While these benefits did not entirely order themselves in the array of greater benefits due to more precise predictive information, a specific trend consistency was observed across all the data analyses. The two displays with least precision offered no benefits on any of the measures. The three with the greatest precision offered some benefit. However within the three displays with greater precision, the ordering was less consistent. The Qualitative Arrows, being the third-most précised, was clearly the best. There was also consistency across the other different dependent variables. To the extent that prediction was better (Figure 7.9), and faster (Figure 7.8), Qualitative Arrows also led to more effective control (Figures 7.5 and 7.6).

We suspect the success of Qualitative Arrows was based on the “sweet spot” between low and high precision of graphical rate-of-change representation. Simple, low-precision representations do not offer as much information, and this lesser information could therefore explain why these predictive aids were not as effective. Findings further showed that presenting more précised graphical representation generally elicited improved control performances. The remaining three higher-

precision representations differentiated from one another by its ability to support change detection. Jessa & Burns (2007) noted that dynamic monitoring in process control involves keeping track of small and large changes, and hence it is important for operators to detect changes as and when they happen. This detection process is further supported when visual shapes show abrupt changes to indicate when variables have changed direction, creating an attention-grabbing effect (Yantis & Jonides, 1984). The continuous nature of Range Indicator and Predictive Indicator lacked this advantage, especially when the process is slow and sluggish. Although gradual change of graphical data representations is still effective at portraying a process shift into abnormal states (Tharanathan, Bullemer, Laberge, Reising, McLain, 2010), it is not as efficient in attracting attention and eliciting change detection as compared to abrupt changes. Of course the Direction-of-Change predictor also had an abrupt discrete aspect, but it did not offer sufficient precision of predictive information to be useful, a deficit that may offset any benefit that the attention capture might provide. Overall, it can be said that the Qualitative Arrows did not have higher information content, but rather it had a higher “information value” in that it presented more pertinent information to the user (i.e.: “Look at me! There was a change. Now the parameter’s behavior is going in this general direction.”).

This “abrupt” characteristic of Qualitative Arrows probably mitigated the effects of change blindness. In the context of process monitoring of complex control systems, change blindness can be seen as important changes in visually presented information that are missed due to visual transient or distraction (Durlach, 2004; see Rensick, 2002, as well as Simons & Rensink, 2005 for more generic references). Specifically, change blindness can occur when the visual objects change gradually (Simons, Franconeri, & Reimer, 2000). Hence the “gradual changing” behavior of Range Indicator and Predictive Indicator may have inhibited change detection, although ultimately participants would still recognize the eventual emergent features which would indicate an imbalanced process state (i.e.: diagonal gradient of arrows or linear indicators). These factors, coupled with the small size of these predictor shapes as well as the attention-sharing nature of the job, could explain why participants fared better when using the Qualitative Arrows versus the more précised Range Indicator and Predictive Indicator. Such continuous-moving visual objects can incorporate

“abrupt” features that increase saliency of deviations and attract attention during larger deviations (Reising & Bullemer, 2008; Jessa & Burns, 2006). For instance, when the rate-of-change exceeds a certain value, the Range Indicator’s border may change into a different color, or its arrow may immediately turn darker, thicker and more salient.

The amount of control activity was analyzed to explore its relationship with participants’ performance, and results hinted that performance guided by highly-precise rate-of-change visualization reduced possible effects of over-control. In general, the presence of low-precision predictors led to increased control activity, and as the precision of representation increased this increase in control activity was mitigated. We associate this increased in control activity with possible inappropriate and ineffective over-control of the process. Examining the profiles in Figure 7.11, in conjunction with the performance profiles in Figures 7.5 and 7.6, clearly the control activity supported by the two continuous predictor displays was effective, even as it was not greater than the baseline.

The **increase** in control activity for the three less precise displays is more complex. For Qualitative Arrows this increase was apparently productive (or the productivity of this information was not offset by any negativity associated with over control). For the other two less precise displays, it was not productive. A possible post hoc explanation, plausible in light of the previous analysis, may be due to two factors which drove control frequency (and in particular, what we have called “over control”): **1.** The *absence of precision*, which created a sluggishness of change as the underlying variable fluctuates between the discrete boundaries, a sluggishness not shown by the two continuous displays; **2.** The *abrupt, discrete changes*, uniquely characteristic of the Direction-of-Change and Qualitative Arrows predictors. While only the absence of precision was present for Mini-trends, thus showing just a small increase of control activity over baseline, both factors were present for Direction of Change and Qualitative Arrows, and as such both showed a much larger control activity increase.

Both the Mini-Trends and Direction-of-Change, which generated no positive control performance, showed statistically significant over-control effects (although arguably, Mini-Trends resulted in much less percentage of control activity increase

than Direction-of-Change). Over-control was marginally significant in the moderately-precise Qualitative Arrows, while Range Indicator and Predictive Indicator did not elicit any over-control behaviors. It is also possible that the over-control behavior had conversely offset any benefits that the predictors might have offered, such that we do not see any performance advantages in the low-precision predictors of Mini-Trends and Direction-of-Change, and that the benefits seen in Qualitative Arrows might have been partially off-set by some over-control. Additional research will need to be conducted for more concrete conclusions.

This study has shown that given the right data visualization, the proposed FROC algorithm can be used effectively at supporting proactive monitoring and anticipatory behavior. As reviewed in Chapter Four, current predictive algorithms employed in today's applications are not always appropriate for our purpose of developing a predictive display. The proposed FROC algorithm is a single-variate calculation based on a moving window of the parameter's historical data. It does not depend on data from other related process parameters unlike multi-variate calculations. However given that the FROC algorithm relies on dynamic historical data, some amount of lag is expected. This study showed that the algorithm was able to provide performance benefits despite this lag.

Although data pertaining to participants' accuracy in prediction yield some findings, the general trend appeared positive. A review of the data collected showed that for most of the visualization objects, at least 50% of the participants saw improvement in their prediction accuracy (e.g.: reduced absolute difference between their predictions and the actual parameter reading one minute later). Our findings reinforced the notion that accurate prediction is harder to achieve as the look-ahead-time becomes further. The future is dynamic and uncertain, so much so that even computer simulator models include confidence bands when deriving predictions.

Overall, three high-precision predictors showed potential in supporting proactive monitoring on a schematic process control display. The Qualitative Arrows, Predictive Indicator, and to a lesser extent the Range Indicator, displayed positive overall results, with the Qualitative Arrows eliciting the most performance benefits

despite not having the highest representation precision. We suspect that the Mini-Trends, with its small display size, might be difficult for operators to detect deviation until much later when the deviation is more visible. That is not to say that Mini-Trends are useless, for they may be beneficial at other operator tasks such as situation diagnosis. Intuitively, we anticipated the Predictive Indicator which provided the most explicit predictive information to be most beneficial. One of the Predictive Indicator's weaknesses may also be attributed to cluttering effects caused by the crowding of multiple information within a small display shape (Rosenholtz, Li, Mansfield, Jin, 2005). This simulator study revealed, particularly in the context of tasks requiring shared attention, that benefits of high-precision graphical information can be limited, if it is unable to timely attract the user's attention. Discrete alerts with some qualitative information evidently supported attention-sharing better than visualizations that only showed continuous representations. Continuous representation formats should therefore also incorporate some form of discrete visual indicators, such as a change in color, thickness, shape etc. The change in features, along with the increased salience, would certainly attract user's attention during process deviations (Nunes, Wickens, Yin, 2006).

While the findings from this study are conclusive, future studies should include real process control operators as subjects to see if these results hold. Real processes are even more complex, and many parameters are interlinked and managed using closed-loop controls. The implementation and strategies for these rate-of-change predictors may potentially differ from those seen here. Using this study's results, more complex visualizations suitable for larger operations could also be built for further evaluation. It would also be interesting to observe the effectiveness of the FROC algorithm on actual processes as well as the effort required in implementing and calibrating such applications. This study is thus a solid stepping stone for such future explorations.

7.6 REFERENCES

- Durlach, P. (2004). Change Blindness and Its Implications for Complex Monitoring and Control Systems Design and Operator Training. *Human-Computer Interaction*, 19, 423-451.
- Jessa, M., & Burns, C. M. (2007). Visual sensitivity of dynamic graphical objects. *International Journal of Human-Computer Studies*. 65, 206-222.
- Nunes, A., Wickens, C. D., Yin, S. (2006). Examining the viability of the Neisser search model in the flight domain and the benefits of highlighting in visual search. In *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.
- Reising, D. C., & Bullemer, P. T. (2008). A direct perception, span-of-control overview display to support a process control operator's situation awareness: A practice-oriented design process. In *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, Santa Monica, CA: Human Factors Society
- Rensick, R. A. (2002). Change Detection. *Annual Review of Psychology*, 53, 245-277.
- Rosenholtz, R., Li, Y., Mansfield, J., Jin, Z., (2005). Feature congestion: A measure of display clutter. *SIGCHI 2005*, 761-770.
- Simons, D. J., Franconeri, S. L., & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception*, 29, 1143-1154.
- Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, 9, 16-20.
- Tharanathan, A., Bullemer, P., Laberge, J., Reising, V., McLain, R. (2010). Functional versus schematic overview displays: Impact on operator situation awareness in process monitoring, *Proceedings of the 54th Human Factors and Ergonomics Society Annual Meeting*. Santa Monica, CA: Human Factors Society
- Wickens, C. D. (1986). The effects of control dynamics on performance. In K. R. Boff, L. Kaufman & J. P. Thomas (eds.), *Handbook of Perception and Human Performance: Volume 2*. New York: Wiley.
- Yantis, S. & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 601-621.

Chapter Eight

Concluding Remarks

8.1 RESEARCH ACCOMPLISHMENTS

Industry process control can be very complex and challenging. Console operators based within a control room have to monitor and control large production units comprising of hundreds to thousands of components and parameters, and these operators do so from behind the distributed control system (DCS). Various DCS displays provide different information to give the operators a sense of the unit's health, from schematic layouts which give an overview of the operations, to alarm summaries that document triggered alarms in chronological order, to trend displays that plot data of parameters over time. Aside from the trends display which plots data over time, none of the other DCS visuals provided explicit information about what may happen to the plant in the near future. Furthermore, interpreting trend patterns required knowledge and experience which are limited to experienced operators. This project aimed to explore a viable predictive visualization which can be incorporated into existing DCS displays.

In the process of achieving this aim, a series of literature review as well as four empirical studies were conducted. While the literature review provided theoretical foundations and research direction, it also brought out knowledge gaps not currently available. Two qualitative investigations provided insights on how console operators interact with their mental models and their use of trend displays on the DCS. These findings led towards the first experimental study, which featured incorporating predictive visual elements within a trends display, powered by a multi-variate rate-of-change algorithm. The algorithm was further redesigned into a single-variate formula and tested in the final experimental study along with a series of predictive shapes in a DCS schematics display. In the end a predictive shape with moderate "data precision" stood out from the rest, proving to be a viable predictive visualization for process control consoles. Additional discussions on the project's overall original findings are further discussed below.

Review Current Knowledge on Human Prediction

Literature pertaining to prediction is scattered across different topic domains, hence Chapter Two sought to cover as many bases as possible, and integrate these different knowledge into one coherent picture. Human prediction may either be top-down (expectancy-driven) or bottom-up (cue-perception), and this project chose to tackle bottom-

up prediction which was more relevant to the proactive monitoring tasks of process control operators. The literature review was able to piece together components of bottom-up prediction. Perceived cues are processed along with the individual's mental model, and through this mental simulation a situation diagnosis (e.g.: how it got to this state) can be achieved. Prediction is derived as the individual factors in the span of prediction, and the further into the future the more effort is required in coming up with an accurate prediction. Nonetheless, experts have shown, through recognition-primed decision-making, that diagnoses and predictions can be done almost automatically without relying on mental simulations.

This project contributed a theoretical synthesis of bottom-up prediction components. Many studies often pair relevant constructs together, such as mental models and situation awareness (Endsley, 2000; Hrebec & Stiber, 2001; Mogford, 1997) or mental models and expertise (Bellenkes, Wickens, Kramer, 1996; Serfaty, Macmillan, Entin, Entin, 1997) etc. This project reviewed many of these papers to find and associate these links to form a holistic process representation. Although putting the entire model under a controlled environment for empirical experimentation can be quite challenging, many of these constructs have already been individually researched fairly extensively and their roles as components to the overall concept can be seen clearly from these researches. This construction would be a good foundation for future study into the processes of human prediction.

Investigate Predictive Applications in Other Domains and in Process Control

Before developing a prototype predictive display for process control, a review was conducted on existing concepts from other domains to establish a predictive display categorization. Predictive tools of varying functions have been implemented across many different applications, from aviation to cars to weather forecasting. Information can either be discrete (one-off event) or continuous (non-stop process), and the predictive aid can be implicit (requires manual extrapolation) or explicit (prediction is derived automatically). Regardless, the review showed how predictive tools shared a common trait, which was that the further the span of prediction or “look-ahead time” was for the prediction to be made, the less accurate the prediction would be. The automated computation to derive these predictions may itself be inaccurate, further compounding the issues of imperfect predictive aids.

A separate review on some of the predictive computations currently available in process control further highlighted the problems of imperfect automation. Highly robust predictive algorithms and computations have been developed and used in today’s process control. However these systems were often designed for steady-state operations in which process control operators would activate so that the computer would take over controls allow for more sensitive production optimization. They were not specifically designed to handle abnormal situations, and in fact operators do switch them off when processes were off-normal and reverted back to manual control. Ironically it is also during abnormal situations which operators would benefit from predictive aids the most. These algorithms would not serve well for the project’s purpose in developing a predictive display, given its varying reliability and potential performance decrement during critical abnormal situations, thus explained the push towards deriving a single-variate predictive algorithm.

Both reviews created an awareness of the limitations faced in pursuing this project, as well as subsequent initiatives pertaining to predictive technologies in process control. These reviews reflected the difference in characteristics compared to other domains like hurricane forecasting or aviation, in which there are tens, if not

hundreds, of variables that can change the future of a parameter's trajectory, as well as the computational challenges.

Analyze Cognitive Contributors and Current Tools for Proactive Monitoring

Two qualitative investigations were conducted with process control console operators. The first study documented how ten console operators derive, update, and use their mental models throughout their shifts. Mental models are hard to quantify, and currently no effort has been made to understand process control operators' mental models, which we have identified as a key component of cue-based predictions and proactive monitoring. Site sources have indicated that expert console operators with at least ten years of working experience were known to behave proactively. Literature on expertise also reported on improved development and usage of mental models among domain experts (Klein & Crandall, 1995; Williams, Ward, Knowles, Smeeton, 2002; Bellenkes, Wickens, Kramer, 1996). These factors motivated this first study to be conducted, thus providing fundamental insights into the research project's target audience. Findings highlighted key activities like performing, early at the start of the shift, a mental simulation of a "virtual round" or walkabout by paging through the various information displays. Their mental models were often associated with "process knowledge", which they commented were derived through years of working in the production field outside the control room. Similar to other domain experts, the interviewees noted their more efficient scanning strategies as compared to their novice counterparts, and oftentimes it was this efficient scanning coupled with their strong process knowledge that enabled these experts to monitor and response proactively.

The second qualitative investigation focused on the trend display that expert console operators are known to rely on for proactive monitoring: Trends. Trend displays plot the data reading of parameters into graphs over time. Many reports have been made on the benefits of graph-based data displays (Meyer, Shinar & Leiser, 1997; Wickens & McCarley, 2008; Porat, Oron-Gilad, Meyer, 2009) but actual accounts on the use of trend displays in process control were less common (Hajdukiewicz & Wu, 2006). This study investigated how trend displays revealed emergent features that operators picked up early while monitoring, as well as provided

patterns which experts were able to use to diagnose the plant's current status or anticipate potential problems. Similar to the results from Hajdukiewicz & Wu (2006), the study found that trend displays was most beneficial for proactive monitoring for it supported the extrapolation of rate-of-change information.

Both studies contribute valuable information towards the development of a predictive display in process control which were not directly available in the literature. The value of trend displays particularly in process control was documented. While studies were done on operators in control rooms (Mumaw, Roth, Vincente, Burns, 2000; Sheridan, 2006) few targeted specifically at proactive monitoring behavior and anticipatory control. The choice of investigating mental models and data displays was driven by our understanding of cue-based predictions, given the two basic components being the mental model and cue perception.

Explore the Viability of a Predictive Display to Support Proactive Monitoring

Two experimental studies were conducted to explore two different prototype displays. The first study featured a trend display with either a linear or numerical rate-of-change indicator. The rate-of-change is calculated using a multiple-input-single-output algorithm that factors in all the variables related to the parameter. Performance benefits were found when participants were exposed to these rate-of-change cues, and survey revealed participants' subjective preference for linear display format. The benefits appeared evident even though trend displays were known to implicitly provide rate-of-change information. This study thus demonstrated the potentials of using rate-of-change based visualizations as the basis for a process control predictive display. The engineering challenge which was previously highlighted was that multi-variate algorithms, although simple to form in our Honey Mixer system, tend to be very complex and difficult to maintain in actual industrial units. Although this study established a proof-of-concept, more work was needed to validate the viability of a rate-of-change predictive display if we want to see it applied successfully in industries.

A second experimental study incorporated same Honey Mixer process into an actual Honeywell ExperionTM DCS schematic display. Five different rate-of-change

visualization objects were designed, each with different data resolution pertaining to how the rate-of-change is being presented. That is to say, on one end of the spectrum rate-of-change information has to be inferred, on the other end a 1-minute prediction of the future based on the rate-of-change was explicitly illustrated. Given the complexity in implementing multi-variate algorithms, a single-variate filtered rate-of-change algorithm was developed to drive the predictive visualizations.

Results indicated that although performance benefits generally existed with display objects that had higher data precision and abrupt feature changes garnered the most advantages. Qualitative Arrows' success could be attributed to higher informative value (not necessarily higher information content) and the relatively stronger attention-grabbing effect.

These prototype predictive displays demonstrated their effectiveness towards supporting proactive monitoring and anticipatory control for process control. In particular, the high-fidelity setup in the second experimental study further demonstrated the possibilities of rate-of-change objects as a actual predictive tool for the process control industry. Despite the inherent mild computational lag, the filtered rate-of-change algorithm was still sufficient at improving operator performance. Findings have already streamed into industry organizations, with the possibly of seeing similar predictive tools being implemented or incorporated into new research. This study marks the end of the project, given the accomplishments of these various objectives as well as the main goal of improving process control console operator's performance through means of facilitating proactive monitoring.

8.2 PROJECT LIMITATIONS

Although the overall results provided some evidence that the proposed predictive visualization tool may be useful, some limitations need to be acknowledged. For one, the experiments utilized university students instead of real operators as subjects. During the discussions with project sponsors (ASM Consortium) it was

concluded that using university students instead of actual operators would generate a relatively larger sample size, and reduce the resources required in setting up the experiment (simpler simulation, no need for complex interfacing with prototype displays and data-collection). While current findings hinted positive performance benefits, the limitation to conclude decisively that the proposed prototype would truly support actual operators in real-world process control is acknowledged. Comparing between expert and novice operators would also allow deeper understanding regarding the conceptual model of cue-based prediction, given that experts are known to practice recognition-primed decision-making.

In order to maintain a controlled testing environment, the “look-ahead time” had to be fixed. In the first experiment (trend-based display) a pre-determined linear extrapolation of up to 1.5min from the current moment was made so as to allow for ample visual change to be perceived. The span of prediction of 1.5-minute was arbitrary at that stage of the project given it was to initially establish benefits of predictive aids in process control (versus none at all). The subsequent experiment (schematic-based display with various predictive shapes) revolved around a 1-minute span of prediction because during simulation calibration, it was found that it took on average approximately 1 minute after the start of a process deviation to trigger the first alarm (i.e.: a salient visual alert). While different people may prefer and benefit from different prediction span, this variable was controlled in the experiments to allow for a cleaner evaluation of the predictive shapes, as well as to minimize cognitive confusion when attempting the experiment. Admittedly, a manipulation of this parameter would provide valuable insights on how people utilize predictive tools, since the further the “look-ahead time”, the more variable the outcome may be.

The experiments assumed 100% reliability on the part of the predictive algorithm and visualization. Unreliability can come in many forms: as discussed above, the further the “look-ahead time”, the lower the pinpoint accuracy of the prediction; the algorithm may not operate as expected during specific conditions (e.g.: steady-state algorithms during abnormal situations); rate of occurrence for discrete error (e.g.: how often the reality differed significantly from the predicted). Notably, Wickens & Dixon (2007) compiled and analyzed the many various studies of imperfect automated

predictors, and a project contributing towards understanding imperfect automated predictions for process control would be secondary as compared to first establishing performance benefits of predictor prototypes. Nonetheless, while presenting system confidence may introduce new failure modes, some existing research do suggest that presenting the computer's decision confidence do improve user trust of the system, and an improved user performance in general (McGuirl & Starter, 2006; Antifakos, Kern, Schiele, Schwaninger, 2005).

8.3 PREDICTING FUTURE WORKS

The discussion on some of the limitations of this project has opened up possible ideas for future work. Regardless of domains, everyone wants to be able to foresee the future so as to *react* preemptively (already a non-sequitor, “proact” perhaps?). Reliable predictive tools and displays will be of interest to many operators of complex systems. Building on current findings, using the rate-of-change of parameters as a predictive cue seems all too simple. Certainly in the future more intelligent predictive algorithms will be developed. Ideally these algorithms will be easy to implement and maintain, yet robust enough to make predictions even during abnormal situations. Future algorithms may even reliably estimate the future readings of parameters, or the time it would take for a parameter to reach certain limits.

Similar research into decoding cue-based prediction can also be performed in other environments, particularly in aviation. Aviation also features complex systems when controlling aircrafts, requiring pilots to predict and “stay ahead of the plane”. The cognitive constructs involved in piloting aircrafts are similar. Pilots need to maintain situation awareness by perceiving information inside and outside the cockpit in order to comprehend the current state and project where the plane will be in the near future. The disorientation experienced by novice pilots during instrument-flight conditions (IFR conditions, i.e.: relying solely on information displayed within the cockpit) reflects their limited mental models and their weak abilities in performing mental simulations. Research into aviation offers easy access to both low- and high-fidelity flight simulators. A viable research direction could focus on investigating the differences in simulated final approaches between novice and expert pilots, to see how

cue-based predictions and mental models differ between the two groups. Pilots are trained to look at specific cues to help them land the plane safely, and oftentimes novice pilots may often overlook these cues, or possess a weaker mental model (partly due to inexperience), which could lead to flawed predictions (or the inability to “stay ahead of the plane”) resulting in poor landing performances.

Future efforts may also look into higher-fidelity testing within process control, such as the setting up of predictor shapes in industry simulators and using actual process control operators. The HoneyMixer micro-system was designed so that it support easy learning by university undergraduates, but yet featured the non-linear process complexity found in actual process plants. This micro-system may be too simple for actual process operators to fully explore the benefits of predictive aids. Through the help of industry partners like Honeywell as well as various site sponsors, existing computer-based advanced training simulators found in actual process plants could be used to power DCS displays. Having actual process control operators should also allow for deeper exploration on how expertise and experience play a role in mental prediction, possibly validating or enhancing the current theoretical understanding of cue-based prediction. Should using such a method to derive quantitative results appears be resource intensive, other more qualitative approaches such as focus groups and interviews with actual operators should gather useful information on the prototype’s viability, as well as amendments before actual industry implementation.

Arguably, as processes become more and more complex, and controls become more and more automated, is there even a place for these rate-of-change predictive displays? Today’s process sites feature “Automatic Process Control” programs that serve as “auto-pilots” for the operators, taking over production control from operators during steady-state operations while optimizing the production for the overall good of the company. Yet there is a need for “resilience engineering” (Hollnagel, Woods, Leveson, 2006), where failures are due to the inability to adapt necessarily so as to cope with the complexity of the processes. Rigid automated processes and lean, “optimized” operations may not be designed to handle unexpected, dynamic situations. As such, during off-normal or sensitive operations like start-ups, manual control still

dominates as humans are more adaptable to changes, and it is also during such occasions that operators benefit from a reliable tool which predicts the behaviors of critical parameters. Until the days when production processes are fully automated without human-in-the-loop interventions, there is value in developing predictive displays.

Conceivably, more complex display objects may be developed through the incorporation of these rate-of-change cues along with other forms of information visualization. Referring back to the recommendations for design by Endsley, Bolte & Jones (2003), the futuristic display should allow easy retrieval of specific data, provide a holistic diagnosis of the current situation, and assist operators in making future projections. The second experimental study tested fundamental design concepts concerning the displayed rate-of-change's "data resolution". It would be interesting to see how new display objects can be derived through integrating these basic rate-of-change shapes to form more complex, more informative visualizations.

In the end, as physics Nobel laureate Niels Bohr would say, "It is difficult to make predictions, particularly about the future".

8.4 REFERENCES

- Antifakos, S., Kern, N., Schiele, B., and Schwaninger, A. (2005). Towards improving trust in context aware systems by displaying system confidence. *ACM International Conference Proceeding Series*, 111, 9-14.
- Bellenkes, A., Wickens, C.D., & Kramer, A. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space and Environmental Medicine*, 68, 569-579.
- Endsley, M. R.: Situation Models: An Avenue to the Modeling of Mental Models. *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA, 2000b.
- Endsley, M.R., Bolté, B. & Jones, D. G. (2003). *Designing for Situation Awareness: An Approach to User-Centered Design*. London, UK: Taylor & Francis.
- Hajdukiewicz, J. & Wu, P (2006). Beyond trends: A framework for mapping time-based requirements and display formats for process operations. *Human Factors and Ergonomics Society Annual Meeting Proceedings 2007*, 1885-1889.
- Hrebec, D. G. & Stiber, M. A. (2001). Survey of System Administrator Mental Models and Situation Awareness, In *Proceedings of the ACM Computer Personnel Research (SIGCPR'01)*, 166-172. New York, NY, 2001: ACM Press
- Klein, G. A. & Crandall, B. W. (1995). The role of mental simulation in naturalistic decision making. In P. Hancock, J. Flach, J. Caird, and K. Vicente (eds.), *Local Applications of the Ecological Approach to Human Machine Systems (vol. 2)*. Hillsdale, NJ: Erlbaum.
- McGuirl, J. M., & Sarter, N. B. (2006). Supporting trust calibration and the effective use of decision aids by presenting dynamic system confidence information. *Human Factors*, 48(4), 656–665.
- Meyer, J., Shinar, D., Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39, 268-286.
- Mogford, R.H. (1977). Mental models and situation awareness in air traffic control. *International Journal of Aviation Psychology*, 7, 331-342.
- Mumaw, R. J., Roth, E. M., Vicente, K. J., Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors*, 42, 36-55.
- Serfaty, D., Macmillan, J., Entin, E. E., & Entin, E. B. (1997). The decision making expertise of battle commanders. In C. E. Zsombok & G. A. Klein (Eds.), *Naturalistic Decision Making*. Mahwah, NJ: Lawrence Earlbaum.
- Sheridan, T. B. (2006) Supervisory Control, in G. Salvendy (ed.), *Handbook of Human Factors and Ergonomics*, Third Edition, John Wiley & Sons, Inc.. NJ: Hoboken.
- Talya Porat, T., Oron-Gilad, T., Meyer, J. (2009). Task-dependent processing of tables and graphs. *Behaviour & IT*, 28, 293-307

Wickens, C. D., & McCarley, J. (2008). *Applied attention theory*. Boca-Raton, FL: Taylor & Francis.

Williams, A. M., Ward, P., Knowles, J. M., Smeeton, N. J. (2002). Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology: Applied*, 8, 259-270.

APPENDIX A

Subject ID:

Date:

Subjective Feedback Form

After performing all three experimental scenarios, please rate the following statements.

		Strongly Disagree				Strongly Agree			
1	I found the production process easy to understand.	1	2	3	4	5	6	7	
2	I understood how the interface worked.	1	2	3	4	5	6	7	
3	The process scenarios were difficult to manage.	1	2	3	4	5	6	7	
4	There were too many things to look out for during the scenarios.	1	2	3	4	5	6	7	
5	I could have managed the process better.	1	2	3	4	5	6	7	
6	The recycled line and mixer output changed too frequently.	1	2	3	4	5	6	7	
7	It was difficult to spot process changes without any Predictors.	1	2	3	4	5	6	7	
8	The Line Predictor was easy to understand	1	2	3	4	5	6	7	
9	The Number Predictor was easy to understand	1	2	3	4	5	6	7	
10	I did not need the Predictors to perform well	1	2	3	4	5	6	7	
11	It was easy to anticipate future process values without Predictors.	1	2	3	4	5	6	7	
12	Right now I do not feel fatigued after completing all the scenarios.	1	2	3	4	5	6	7	
13	I enjoyed the entire experiment.	1	2	3	4	5	6	7	

Of the three display variants that you have encountered, please rank them according to your most favorite to your least favorite: No Predictor, Number Predictor, Line Predictor.

APPENDIX B

Univariate Analysis of Variance (Baseline Easy vs Baseline Hard)

Between-Subjects Factors

	N
condition 1.00	22
2.00	22

Tests of Between-Subjects Effects

Dependent Variable:baseline

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2.077E6	1	2077127.273	19.184	.000
Intercept	6241617.818	1	6241617.818	57.648	.000
condition	2077127.273	1	2077127.273	19.184	.000
Error	4547416.909	42	108271.831		
Total	1.287E7	44			
Corrected Total	6624544.182	43			

a. R Squared = .314 (Adjusted R Squared = .297)

Mixed-Plot ANOVA

Within-Subjects Factors

Measure:MEASURE_1

DisplayType	Dependent Variable
1	baseline
– 2	number
3	line

Between-Subjects Factors

	N
condition 1.00	22
2.00	22

Descriptive Statistics

condition	Mean	Std. Deviation	N
baseline 1.00	159.3636	143.55306	22
– 2.00	593.9091	442.64679	22
Total	376.6364	392.50372	44
number 1.00	119.7727	102.56284	22
– 2.00	470.9091	411.20275	22
Total	295.3409	345.33462	44
line 1.00	112.5455	120.93986	22
– 2.00	381.4545	341.47268	22
Total	247.0000	287.38047	44

Multivariate Tests^b

Effect		Value	F	Hypothesis df
DisplayType	Pillai's Trace	.217	5.680 ^a	2.000
	Wilks' Lambda	.783	5.680 ^a	2.000
	Hotelling's Trace	.277	5.680 ^a	2.000
	Roy's Largest Root	.277	5.680 ^a	2.000
DisplayType * condition	Pillai's Trace	.095	2.143 ^a	2.000
	Wilks' Lambda	.905	2.143 ^a	2.000
	Hotelling's Trace	.105	2.143 ^a	2.000
	Roy's Largest Root	.105	2.143 ^a	2.000

a. Exact statistic

b. Design: Intercept + condition

Within Subjects Design: DisplayType

Multivariate Tests^b

Effect		Error df	Sig.
DisplayType	Pillai's Trace	41.000	.007
	Wilks' Lambda	41.000	.007
	Hotelling's Trace	41.000	.007
	Roy's Largest Root	41.000	.007
DisplayType * condition	Pillai's Trace	41.000	.130
	Wilks' Lambda	41.000	.130
	Hotelling's Trace	41.000	.130
	Roy's Largest Root	41.000	.130

b. Design: Intercept + condition

Within Subjects Design: DisplayType

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
– DisplayType	.906	4.029	2	.133

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

b. Design: Intercept + condition

Within Subjects Design: DisplayType

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Epsilon ^a		
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
– DisplayType	.914	.977	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept + condition

Within Subjects Design: DisplayType

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square
DisplayType	Sphericity Assumed	377686.924	2	188843.462
	Greenhouse-Geisser	377686.924	1.829	206517.799
	Huynh-Feldt	377686.924	1.953	193352.488
	Lower-bound	377686.924	1.000	377686.924
DisplayType * condition	Sphericity Assumed	150897.288	2	75448.644
	Greenhouse-Geisser	150897.288	1.829	82510.073
	Huynh-Feldt	150897.288	1.953	77250.135
	Lower-bound	150897.288	1.000	150897.288
Error(DisplayType)	Sphericity Assumed	3672112.455	84	43715.624
	Greenhouse-Geisser	3672112.455	76.811	47807.080
	Huynh-Feldt	3672112.455	82.041	44759.425
	Lower-bound	3672112.455	42.000	87431.249

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		F	Sig.
DisplayType	Sphericity Assumed	4.320	.016
	Greenhouse-Geisser	4.320	.019
	Huynh-Feldt	4.320	.017
	Lower-bound	4.320	.044
DisplayType * condition	Sphericity Assumed	1.726	.184
	Greenhouse-Geisser	1.726	.187
	Huynh-Feldt	1.726	.185
	Lower-bound	1.726	.196

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	DisplayType	Type III Sum of Squares	df	Mean Square
DisplayType	Linear	369722.909	1	369722.909
	Quadratic	7964.015	1	7964.015
DisplayType * condition	Linear	150894.727	1	150894.727
	Quadratic	2.561	1	2.561
Error(DisplayType)	Linear	1559144.364	42	37122.485
	Quadratic	2112968.091	42	50308.764

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	DisplayType	F	Sig.
DisplayType	Linear	9.960	.003
	Quadratic	.158	.693
DisplayType * condition	Linear	4.065	.050
	Quadratic	.000	.994

Levene's Test of Equality of Error Variances^a

	F	df1	df2	Sig.
baseline	13.793	1	42	.001
number	10.042	1	42	.003
line	7.253	1	42	.010

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

Levene's Test of Equality of Error Variances^a

	F	df1	df2	Sig.
baseline	13.793	1	42	.001
number	10.042	1	42	.003
line	7.253	1	42	.010

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + condition

Within Subjects Design: DisplayType

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1.239E7	1	1.239E7	70.273	.000
condition	4077927.280	1	4077927.280	23.136	.000
Error	7402879.045	42	176259.025		

Estimated Marginal Means

1. Grand Mean

Measure: MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
306.326	36.542	232.582	380.070

2. condition

Measure: MEASURE_1

condition	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1.00	130.561	51.678	26.271	234.851
2.00	482.091	51.678	377.801	586.381

3. DisplayType

Measure: MEASURE_1

DisplayType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	376.636	49.606	276.528	476.745
2	295.341	45.177	204.169	386.512
3	247.000	38.617	169.068	324.932

4. condition * DisplayType

Measure: MEASURE_1

condition	DisplayType	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1.00	1	159.364	70.153	17.789	300.938
	2	119.773	63.890	-9.163	248.709

	3	112.545	54.612	2.333	222.757
2.00	1	593.909	70.153	452.335	735.484
	2	470.909	63.890	341.973	599.845
	3	381.455	54.612	271.243	491.667

Profile Plots

General Linear Model (Within Easy Scenarios)

Within-Subjects Factors	
Measure: MEASURE_1	
DisplayType	Dependent Variable
1	baseline
2	number
3	line

Descriptive Statistics

	Mean	Std. Deviation	N
baseline	159.3636	143.55306	22
number	119.7727	102.56284	22
line	112.5455	120.93986	22

Multivariate Tests^b

Effect		Value	F	Hypothesis df	Error df	Sig.
DisplayType	Pillai's Trace	.079	.857 ^a	2.000	20.000	.439
	Wilks' Lambda	.921	.857 ^a	2.000	20.000	.439
	Hotelling's Trace	.086	.857 ^a	2.000	20.000	.439
	Roy's Largest Root	.086	.857 ^a	2.000	20.000	.439

a. Exact statistic

b. Design: Intercept

Within Subjects Design: DisplayType

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
– DisplayType	.976	.478	2	.787

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

b. Design: Intercept

Within Subjects Design: DisplayType

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Epsilon ^a		
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
– DisplayType	.977	1.000	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept

Within Subjects Design: DisplayType

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square
DisplayType	Sphericity Assumed	27951.848	2	13975.924
	Greenhouse-Geisser	27951.848	1.954	14305.964
	Huynh-Feldt	27951.848	2.000	13975.924
	Lower-bound	27951.848	1.000	27951.848
Error(DisplayType)	Sphericity Assumed	568157.485	42	13527.559
	Greenhouse-Geisser	568157.485	41.031	13847.011
	Huynh-Feldt	568157.485	42.000	13527.559
	Lower-bound	568157.485	21.000	27055.118

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		F	Sig.
DisplayType	Sphericity Assumed	1.033	.365
	Greenhouse-Geisser	1.033	.363
	Huynh-Feldt	1.033	.365
	Lower-bound	1.033	.321

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	DisplayType	Type III Sum of Squares	df	Mean Square	F	Sig.
DisplayType	Linear	24111.364	1	24111.364	1.572	.224
	Quadratic	3840.485	1	3840.485	.328	.573
Error(DisplayType)	Linear	322022.636	21	15334.411		
	Quadratic	246134.848	21	11720.707		

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1125040.742	1	1125040.742	60.169	.000
Error	392656.924	21	18697.949		

Estimated Marginal Means

1. Grand Mean

Measure: MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
130.561	16.832	95.557	165.564

2. DisplayType

Measure: MEASURE_1

DisplayType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	159.364	30.606	95.716	223.011
– 2	119.773	21.866	74.299	165.247
3	112.545	25.784	58.924	166.167

General Linear Model (Within Hard Scenarios)

Within-Subjects Factors

Measure: MEASURE_1

DisplayType	Dependent Variable
1	baseline
– 2	number
3	line

Descriptive Statistics

	Mean	Std. Deviation	N
baseline	593.9091	442.64679	22
number	470.9091	411.20275	22
line	381.4545	341.47268	22

Multivariate Tests^b

Effect		Value	F	Hypothesis df	Error df	Sig.
DisplayType	Pillai's Trace	.323	4.774 ^a	2.000	20.000	.020
	Wilks' Lambda	.677	4.774 ^a	2.000	20.000	.020
	Hotelling's Trace	.477	4.774 ^a	2.000	20.000	.020
	Roy's Largest Root	.477	4.774 ^a	2.000	20.000	.020

a. Exact statistic

b. Design: Intercept

Within Subjects Design: DisplayType

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
– DisplayType	.851	3.233	2	.199

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

b. Design: Intercept

Within Subjects Design: DisplayType

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Epsilon ^a		
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
– DisplayType	.870	.942	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept

Within Subjects Design: DisplayType

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square
DisplayType	Sphericity Assumed	500632.364	2	250316.182

	Greenhouse-Geisser	500632.364	1.740	287680.083
	Huynh-Feldt	500632.364	1.884	265729.235
	Lower-bound	500632.364	1.000	500632.364
Error(DisplayType)	Sphericity Assumed	3103954.970	42	73903.690
	Greenhouse-Geisser	3103954.970	36.545	84935.059
	Huynh-Feldt	3103954.970	39.564	78454.261
	Lower-bound	3103954.970	21.000	147807.380

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		F	Sig.
DisplayType	Sphericity Assumed	3.387	.043
	Greenhouse-Geisser	3.387	.051
	Huynh-Feldt	3.387	.047
	Lower-bound	3.387	.080

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	DisplayType	Type III Sum of Squares	df	Mean Square	F	Sig.
DisplayType	Linear	496506.273	1	496506.273	8.428	.009
	Quadratic	4126.091	1	4126.091	.046	.832
Error(DisplayType)	Linear	1237121.727	21	58910.558		
	Quadratic	1866833.242	21	88896.821		

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1.534E7	1	1.534E7	45.950	.000
Error	7010222.121	21	333820.101		

Estimated Marginal Means

1. Grand Mean

Measure: MEASURE_1

Mean	Std. Error	95% Confidence Interval	
		Lower Bound	Upper Bound
482.091	71.119	334.191	629.991

2. DisplayType

Measure: MEASURE_1

DisplayType	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	593.909	94.373	397.650	790.168
– 2	470.909	87.669	288.592	653.226
3	381.455	72.802	230.054	532.855

Pairwise T-Test (Within Hard Scenario)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	baseline	593.9091	22	442.64679	94.37261
	line	381.4545	22	341.47268	72.80222
Pair 2	baseline	593.9091	22	442.64679	94.37261
	number	470.9091	22	411.20275	87.66872
Pair 3	line	381.4545	22	341.47268	72.80222
	number	470.9091	22	411.20275	87.66872

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	baseline & line	22	.644	.001
Pair 2	baseline & number	22	.671	.001
Pair 3	line & number	22	.288	.194

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	baseline - line	212.45455	343.25081	73.18132
Pair 2	baseline - number	123.00000	347.43715	74.07385
Pair 3	line - number	-89.45455	452.64605	96.50446

Paired Samples Test

		Paired Differences		t	df	Sig. (2-tailed)
		95% Confidence Interval of the Difference				
		Lower	Upper			
Pair 1	baseline - line	60.26566	364.64343	2.903	21	.009
Pair 2	baseline - number	-31.04500	277.04500	1.661	21	.112
Pair 3	line - number	-290.14656	111.23747	-.927	21	.364

APPENDIX C

General Linear Model (Alarm Presence Measure)

Within-Subjects Factors

Measure: MEASURE_1

PredictorPresence	Dependent Variable
1	Balarmpres
2	Palarmpres

Between-Subjects Factors

	N
Predictor 2.00	10
3.00	10
4.00	10
5.00	10
6.00	10

Multivariate Tests^b

Effect		Value	F	Hypothesis df
PredictorPresence	Pillai's Trace	.201	11.307 ^a	1.000
	Wilks' Lambda	.799	11.307 ^a	1.000
	Hotelling's Trace	.251	11.307 ^a	1.000
	Roy's Largest Root	.251	11.307 ^a	1.000
PredictorPresence * Predictor	Pillai's Trace	.089	1.092 ^a	4.000
	Wilks' Lambda	.911	1.092 ^a	4.000
	Hotelling's Trace	.097	1.092 ^a	4.000
	Roy's Largest Root	.097	1.092 ^a	4.000

a. Exact statistic

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Multivariate Tests^b

Effect		Error df	Sig.
PredictorPresence	Pillai's Trace	45.000	.002
	Wilks' Lambda	45.000	.002
	Hotelling's Trace	45.000	.002
	Roy's Largest Root	45.000	.002
PredictorPresence * Predictor	Pillai's Trace	45.000	.372
	Wilks' Lambda	45.000	.372
	Hotelling's Trace	45.000	.372
	Roy's Largest Root	45.000	.372

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
— PredictorPresence	1.000	.000	0	.

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Epsilon ^a
------------------------	----------------------

	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
– PredictorPresence	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square
PredictorPresence	Sphericity Assumed	29584.000	1	29584.000
	Greenhouse-Geisser	29584.000	1.000	29584.000
	Huynh-Feldt	29584.000	1.000	29584.000
	Lower-bound	29584.000	1.000	29584.000
PredictorPresence * Predictor	Sphericity Assumed	11431.600	4	2857.900
	Greenhouse-Geisser	11431.600	4.000	2857.900
	Huynh-Feldt	11431.600	4.000	2857.900
	Lower-bound	11431.600	4.000	2857.900
Error(PredictorPresence)	Sphericity Assumed	117736.400	45	2616.364
	Greenhouse-Geisser	117736.400	45.000	2616.364
	Huynh-Feldt	117736.400	45.000	2616.364
	Lower-bound	117736.400	45.000	2616.364

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		F	Sig.
PredictorPresence	Sphericity Assumed	11.307	.002

	Greenhouse-Geisser	11.307	.002
	Huynh-Feldt	11.307	.002
	Lower-bound	11.307	.002
PredictorPresence * Predictor	Sphericity Assumed	1.092	.372
	Greenhouse-Geisser	1.092	.372
	Huynh-Feldt	1.092	.372
	Lower-bound	1.092	.372

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	PredictorPresence	Type III Sum of Squares	df	Mean Square
PredictorPresence	– Linear	29584.000	1	29584.000
PredictorPresence * Predictor	– Linear	11431.600	4	2857.900
Error(PredictorPresence)	– Linear	117736.400	45	2616.364

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	PredictorPresence	F	Sig.
PredictorPresence	– Linear	11.307	.002
PredictorPresence * Predictor	– Linear	1.092	.372

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1807949.160	1	1807949.160	155.690	.000
Predictor	102615.440	4	25653.860	2.209	.083

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1807949.160	1	1807949.160	155.690	.000
Predictor	102615.440	4	25653.860	2.209	.083
Error	522563.400	45	11612.520		

Profile Plots

General Linear Model

(Percentage Duration with No Alarms)

Within-Subjects Factors

Measure: MEASURE_1

Predictor	Presence	Dependent Variable
1	—	Bpercent
2		Ppercent

Between-Subjects Factors

	N
Predictor 2.00	10
3.00	10
4.00	10
5.00	10
6.00	10

Multivariate Tests^b

Effect		Value	F	Hypothesis df
PredictorPresence	Pillai's Trace	.252	15.178 ^a	1.000
	Wilks' Lambda	.748	15.178 ^a	1.000
	Hotelling's Trace	.337	15.178 ^a	1.000
	Roy's Largest Root	.337	15.178 ^a	1.000

PredictorPresence * Predictor	Pillai's Trace	.119	1.516 ^a	4.000
	Wilks' Lambda	.881	1.516 ^a	4.000
	Hotelling's Trace	.135	1.516 ^a	4.000
	Roy's Largest Root	.135	1.516 ^a	4.000

a. Exact statistic

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Multivariate Tests^b

Effect		Error df	Sig.
PredictorPresence	Pillai's Trace	45.000	.000
	Wilks' Lambda	45.000	.000
	Hotelling's Trace	45.000	.000
	Roy's Largest Root	45.000	.000
PredictorPresence * Predictor	Pillai's Trace	45.000	.214
	Wilks' Lambda	45.000	.214
	Hotelling's Trace	45.000	.214
	Roy's Largest Root	45.000	.214

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
– PredictorPresence	1.000	.000	0	.

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
– PredictorPresence	1.000	.000	0	.

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Epsilon ^a		
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
– PredictorPresence	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square
PredictorPresence	Sphericity Assumed	1785.063	1	1785.063
	Greenhouse-Geisser	1785.063	1.000	1785.063
	Huynh-Feldt	1785.063	1.000	1785.063
	Lower-bound	1785.063	1.000	1785.063

PredictorPresence * Predictor	Sphericity Assumed	713.207	4	178.302
	Greenhouse-Geisser	713.207	4.000	178.302
	Huynh-Feldt	713.207	4.000	178.302
	Lower-bound	713.207	4.000	178.302
Error(PredictorPresence)	Sphericity Assumed	5292.396	45	117.609
	Greenhouse-Geisser	5292.396	45.000	117.609
	Huynh-Feldt	5292.396	45.000	117.609
	Lower-bound	5292.396	45.000	117.609

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		F	Sig.
PredictorPresence	Sphericity Assumed	15.178	.000
	Greenhouse-Geisser	15.178	.000
	Huynh-Feldt	15.178	.000
	Lower-bound	15.178	.000
PredictorPresence * Predictor	Sphericity Assumed	1.516	.214
	Greenhouse-Geisser	1.516	.214
	Huynh-Feldt	1.516	.214
	Lower-bound	1.516	.214

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	PredictorPresence	Type III Sum of Squares	df	Mean Square
PredictorPresence	— Linear	1785.063	1	1785.063
PredictorPresence * Predictor	— Linear	713.207	4	178.302
Error(PredictorPresence)	— Linear	5292.396	45	117.609

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	PredictorPresence	F	Sig.
PredictorPresence	— Linear	15.178	.000
PredictorPresence * Predictor	— Linear	1.516	.214

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	392865.704	1	392865.704	595.870	.000
Predictor	5513.285	4	1378.321	2.091	.098
Error	29669.136	45	659.314		

Profile Plots

Paired Sample T-Test (Mini-Trend)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres	198.6000	10	84.65380	26.76988
	Palarmpres	185.4000	10	80.60907	25.49083
Pair 2	Bpercent	48.3000	10	20.36664	6.44050
	Ppercent	50.7700	10	15.56970	4.92357

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Balarmpres & Palarmpres	10	.415	.233
Pair 2	Bpercent & Ppercent	10	.653	.041

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres - Palarmpres	13.20000	89.44992	28.28655
Pair 2	Bpercent - Ppercent	-2.47000	15.59324	4.93101

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	Balarmpres - Palarmpres	-50.78862	77.18862	.467	9
Pair 2	Bpercent - Ppercent	-13.62473	8.68473	-.501	9

Paired Samples Test

		Sig. (2-tailed)
Pair 1	Balarmpres - Palarmpres	.652
Pair 2	Bpercent - Ppercent	.628

Paired Sample T-Test

(Direction of Change)

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
--	------	---	----------------	-----------------

Pair 1	Balarmpres	135.2000	10	84.88528	26.84308
	Palarmpres	119.0000	10	71.20705	22.51765
Pair 2	Bpercent	61.8300	10	18.90421	5.97803
	Ppercent	66.1500	10	16.59372	5.24740

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Balarmpres & Palarmpres	10	.770	.009
Pair 2	Bpercent & Ppercent	10	.691	.027

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres - Palarmpres	16.20000	54.44018	17.21550
Pair 2	Bpercent - Ppercent	-4.32000	14.10617	4.46076

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	Balarmpres - Palarmpres	-22.74416	55.14416	.941	9
Pair 2	Bpercent - Ppercent	-14.41095	5.77095	-.968	9

Paired Samples Test

	Sig. (2-tailed)
Pair 1 Balarmpres - Palarmpres	.371
Pair 2 Bpercent - Ppercent	.358

Paired Sample T-Test (Qualitative Arrows)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres	145.3000	10	108.24463	34.22996
	Palarmpres	72.1000	10	58.52340	18.50673
Pair 2	Bpercent	60.5400	10	23.93840	7.56999
	Ppercent	78.1300	10	17.13029	5.41707

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Balarmpres & Palarmpres	10	.865	.001
Pair 2	Bpercent & Ppercent	10	.912	.000

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres - Palarmpres	73.20000	64.70755	20.46232
Pair 2	Bpercent - Ppercent	-17.59000	10.89255	3.44453

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	Balarmpres - Palarmpres	26.91101	119.48899	3.577	9

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	Balarmpres - Palarmpres	26.91101	119.48899	3.577	9
Pair 2	Bpercent - Ppercent	-25.38206	-9.79794	-5.107	9

Paired Samples Test

		Sig. (2-tailed)
Pair 1	Balarmpres - Palarmpres	.006
Pair 2	Bpercent - Ppercent	.001

Paired Sample T-Test (Range Indicator)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres	119.6000	10	83.60914	26.43953
	Palarmpres	84.8000	10	65.91038	20.84269
Pair 2	Bpercent	64.6100	10	21.47238	6.79016
	Ppercent	75.2500	10	17.94357	5.67425

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Balarmpres & Palarmpres	10	.317	.371
Pair 2	Bpercent & Ppercent	10	.425	.221

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres - Palarmpres	34.80000	88.51968	27.99238
Pair 2	Bpercent - Ppercent	-10.64000	21.34730	6.75061

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	Balarmpres - Palarmpres	-28.52316	98.12316	1.243	9
Pair 2	Bpercent - Ppercent	-25.91094	4.63094	-1.576	9

Paired Samples Test

		Sig. (2-tailed)
Pair 1	Balarmpres - Palarmpres	.245
Pair 2	Bpercent - Ppercent	.149

Paired Sample T-Test (Predictive Indicator)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres	159.6000	10	111.00470	35.10277
	Palarmpres	125.0000	10	80.03610	25.30964
Pair 2	Bpercent	56.9900	10	24.10839	7.62374
	Ppercent	64.2200	10	19.07708	6.03270

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Balarmpres & Palarmpres	10	.875	.001
Pair 2	Bpercent & Ppercent	10	.854	.002

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	Balarmpres - Palarmpres	34.60000	56.35443	17.82084
Pair 2	Bpercent - Ppercent	-7.23000	12.63329	3.99500

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	Balarmpres - Palarmpres	-5.71353	74.91353	1.942	9
Pair 2	Bpercent - Ppercent	-16.26731	1.80731	-1.810	9

Paired Samples Test

		Sig. (2-tailed)
Pair 1	Balarmpres - Palarmpres	.084
Pair 2	Bpercent - Ppercent	.104

General Linear Model

(Prediction Probe Response Time)

Within-Subjects Factors

Measure: MEASURE_1

PredictorPresence	Dependent Variable
1	BasePredictRT
2	PredictorPredict RT

Between-Subjects Factors

	N
Predictor 2.00	10
3.00	10
4.00	10
5.00	8
6.00	10

Multivariate Tests^b

Effect		Value	F	Hypothesis df
PredictorPresence	Pillai's Trace	.032	1.420 ^a	1.000
	Wilks' Lambda	.968	1.420 ^a	1.000
	Hotelling's Trace	.033	1.420 ^a	1.000
	Roy's Largest Root	.033	1.420 ^a	1.000
PredictorPresence * Predictor	Pillai's Trace	.083	.978 ^a	4.000
	Wilks' Lambda	.917	.978 ^a	4.000
	Hotelling's Trace	.091	.978 ^a	4.000
	Roy's Largest Root	.091	.978 ^a	4.000

a. Exact statistic

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Multivariate Tests^b

Effect		Error df	Sig.
PredictorPresence	Pillai's Trace	43.000	.240
	Wilks' Lambda	43.000	.240
	Hotelling's Trace	43.000	.240
	Roy's Largest Root	43.000	.240
PredictorPresence * Predictor	Pillai's Trace	43.000	.430
	Wilks' Lambda	43.000	.430
	Hotelling's Trace	43.000	.430
	Roy's Largest Root	43.000	.430

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
— PredictorPresence	1.000	.000	0	.

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Mauchly's Test of Sphericity^b

Measure: MEASURE_1

Within Subjects Effect	Epsilon ^a		
	Greenhouse-Geisser	Huynh-Feldt	Lower-bound
— PredictorPresence	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept + Predictor

Within Subjects Design: PredictorPresence

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square
PredictorPresence	Sphericity Assumed	.495	1	.495
	Greenhouse-Geisser	.495	1.000	.495
	Huynh-Feldt	.495	1.000	.495
	Lower-bound	.495	1.000	.495
PredictorPresence * Predictor	Sphericity Assumed	1.363	4	.341
	Greenhouse-Geisser	1.363	4.000	.341
	Huynh-Feldt	1.363	4.000	.341
	Lower-bound	1.363	4.000	.341
Error(PredictorPresence)	Sphericity Assumed	14.982	43	.348
	Greenhouse-Geisser	14.982	43.000	.348
	Huynh-Feldt	14.982	43.000	.348
	Lower-bound	14.982	43.000	.348

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		F	Sig.
PredictorPresence	Sphericity Assumed	1.420	.240
	Greenhouse-Geisser	1.420	.240
	Huynh-Feldt	1.420	.240
	Lower-bound	1.420	.240

PredictorPresence * Predictor	Sphericity Assumed	.978	.430
	Greenhouse-Geisser	.978	.430
	Huynh-Feldt	.978	.430
	Lower-bound	.978	.430

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	PredictorPresence	Type III Sum of Squares	df	Mean Square
PredictorPresence	— Linear	.495	1	.495
PredictorPresence * Predictor	— Linear	1.363	4	.341
Error(PredictorPresence)	— Linear	14.982	43	.348

Tests of Within-Subjects Contrasts

Measure:MEASURE_1

Source	PredictorPresence	F	Sig.
PredictorPresence	— Linear	1.420	.240
PredictorPresence * Predictor	— Linear	.978	.430

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	1150.887	1	1150.887	625.281	.000
Predictor	12.505	4	3.126	1.698	.168
Error	79.145	43	1.841		

Profile Plots

T-Test (Prediction Probe Response Time: Mini-trends)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT	3.6000	10	1.38154	.43688
	PredictorPredictRT	3.7500	10	.98836	.31255

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	BasePredictRT & PredictorPredictRT	10	.692	.027

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT - PredictorPredictRT	-.15000	.99830	.31569

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	BasePredictRT - PredictorPredictRT	-.86414	.56414	-.475	9

Paired Samples Test

		Sig. (2-tailed)
Pair 1	BasePredictRT - PredictorPredictRT	.646

T-Test (Prediction Probe Response Time: Direction of Change)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT	4.0667	10	1.10331	.34890
	PredictorPredictRT	3.7667	10	1.23278	.38984

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	BasePredictRT & PredictorPredictRT	10	.714	.020

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT - PredictorPredictRT	.30000	.89166	.28197

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	BasePredictRT - PredictorPredictRT	-.33786	.93786	1.064	9

Paired Samples Test

		Sig. (2-tailed)
Pair 1	BasePredictRT - PredictorPredictRT	.315

T-Test (Prediction Probe Response Time: Qualitative Arrows)

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
--	------	---	----------------	-----------------

Pair 1	BasePredictRT	3.6000	10	1.14180	.36107
	PredictorPredictRT	3.2833	10	1.00937	.31919

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	BasePredictRT & PredictorPredictRT	10	.897	.000

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT - PredictorPredictRT	.31667	.50583	.15996

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	BasePredictRT - PredictorPredictRT	-.04518	.67852	1.980	9

Paired Samples Test

		Sig. (2-tailed)
Pair 1	BasePredictRT - PredictorPredictRT	.079

T-Test (Prediction Probe Response Time: Range Indicator)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT	2.7083	8	.86717	.30659
	PredictorPredictRT	2.8542	8	.62002	.21921

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	BasePredictRT & PredictorPredictRT	8	.020	.962

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT - PredictorPredictRT	-.14583	1.05574	.37326

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	BasePredictRT - PredictorPredictRT	-1.02845	.73679	-.391	7

Paired Samples Test

		Sig. (2-tailed)
Pair 1	BasePredictRT - PredictorPredictRT	.708

T-Test (Prediction Probe Response Time: Predictive Indicator)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT	3.7667	10	.83222	.26317
	PredictorPredictRT	3.3667	10	.96801	.30611

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	BasePredictRT & PredictorPredictRT	10	.754	.012

Paired Samples Test

		Paired Differences		
		Mean	Std. Deviation	Std. Error Mean
Pair 1	BasePredictRT - PredictorPredictRT	.40000	.64406	.20367

Paired Samples Test

		Paired Differences		t	df
		95% Confidence Interval of the Difference			
		Lower	Upper		
Pair 1	BasePredictRT - PredictorPredictRT	-.06073	.86073	1.964	9

Paired Samples Test

		Sig. (2-tailed)
Pair 1	BasePredictRT - PredictorPredictRT	.081