Exploring Automated Trading: Modeling a Financial System with a Variable Degree of Automation, Display Design, and Evaluation

by

Yeti Li

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Systems Design Engineering

Waterloo, Ontario, Canada, 2017

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	DAVID B. KABER Distinguished Professor Edward P. Fitts Industrial and Systems Engineering North Carolina State University
Supervisor	CATHERINE M. BURNS Professor Systems Design Engineering University of Waterloo
Internal Members	SHI CAO Assistant Professor Systems Design Engineering University of Waterloo
	KUMARASWAMY PONNAMBALAM Professor Systems Design Engineering University of Waterloo
Internal-external Member	ALAN HUANG Associate Professor Accounting and Finance University of Waterloo

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The following paper is incorporated in chapter 5 this dissertation. Figure and table numbers and references were modified for consistency. All American spellings in the original paper were changed to Canadian spellings.

Li, Y., & Burns, C. M. (2017). Modeling Automation with Cognitive Work Analysis to Support Human-Automation Coordination. *Journal of Cognitive Engineering and Decision Making, 11*(4), , pp. 299 – 322.

Contributor	Statement of Contribution
Li, Y. (Candidate)	Model development (80%)
	Writing and editing (80%)
Burns, C. M. (Supervisor)	Model development (20%)
	Writing and editing (20%)

The two experiments introduced in chapter 3 and 4 were designed in collaboration with several others. The following manuscript includes the content of chapter 3 of this dissertation:

Li, Y., Wang, X. & Burns, C. M. (2017). Ecological Interface Design for Financial Trading: Trading Performance and Risk Preference Effects. Presented at the *2017 IEEE International Conference on Systems, Man, and Cybernetics*, Banff, Canada.

Contributor	Statement of Contribution
Li, Y. (Candidate)	Experiment design (80%)
	Data analysis (100%)
	Writing and editing (80%)
Wang, X.	Subject-domain knowledge (100%)
Burns, C. M. (Supervisor)	Experimental design (20%)
	Writing and editing (20%)

Abstract

Automation has become more pervasive and started to replace human operation in all information-processing stages (e.g., machine-learning based systems). Typically, automated systems have a variable degree of automation (DOA) that challenges modeling and design. Cognitive Work Analysis (CWA) is a modeling approach to support analysts in coping with the complexity of socio-technical domains and has shown success in providing implications for developing ecological displays that are effective in unanticipated situations. CWA allows space for modeling automated systems but has not been well developed to describe variable DOAs. This dissertation explores this problem by focusing on the case of automated trading which is underexplored in both human factors research and finance research. This dissertation starts with developing a modeling approach then focuses on design and evaluation.

Modeling. A DOA layering approach on the decision ladder was developed to serve as the theoretical foundation of this research. Two cases of automated trading – basket trading and trend following trading, each using a different DOA – were presented to facilitate the development of this approach. With this approach, the two most commonly used modeling tools in the CWA, the abstraction hierarchy, and the decision ladder each adopted an additional layer for representing human-administrated functions and automated functions. Also, the four information-processing stages in which the automation could take place were marked on the decision ladder to allow for a more detailed level mapping, which is unique in the CWA research. The DOA layering approach was demonstrated to extend the use case of CWA to include automated systems with a variable DOA and have important implications for ecological display design as well as automation design.

Design and evaluation. The experimental approach presented later in this dissertation further explored automation and display design implications of the DOA layering approach using AUTRASS (AUtomated TRAding System Simulation), a simulation developed as part of this research program. Two experimental studies on trend following trading are reported in which the design concepts were evaluated.

In Experiment 1, inspired by the CWA models, automation was designed as two configurations to represent distinct DOA situations. The moderate DOA configuration simulated a trading situation in which the participants performed a flexible trading task. The high DOA configuration represented a higher DOA situation where a trading algorithm that was unfamiliar to the participants traded in a similar market condition, and the participants monitored the automation and performed a fault detection task. Two types of displays were designed. Conventional displays were typical in information content and form to current trading displays and should support the basic use of the automation. Ecological displays were implemented from the CWA models to support monitoring for unanticipated situations for each DOA

condition. Four scenario types were developed by combing the two DOA configurations with the two display types. Experiment 1 involving 24 participants was conducted to thoroughly examine the effectiveness of ecological displays with different DOAs. Based on the literature, the ecological displays were hypothesized to improve task performance and situation awareness and to trigger riskier actions without imposing higher workload. Results of Experiment 1showed that the ecological displays did not provide better support on either trading performance (moderate DOA) or fault detection performance (high DOA). Empirically, a trade-off of situation awareness and workload between the two DOA configurations seemed to exist without the influence of the ecological displays. Interestingly, the results of this study suggested a different pattern of risk preference compared to that in the literature. Specifically, the ecological displays imposed riskier financial trading decisions. The results of Experiment 1 provided implications for identifying system and contextual factors that could influence risk preference and demonstrated sufficient space for improving automation design.

Experiment 2 is a follow-up to the first experimental study with separately recruited 24 participants. The high DOA configuration previously used in Experiment 1 was improved with its traits inherited and its flaws in automation design eliminated (i.e., the improved-high DOA configuration was expected to be better supported by the ecological displays). An adaptive configuration was introduced to simulate adaptive automation in the automated trading setting. The conventional displays and the ecological displays continued to be used without any modifications, and similar hypotheses were examined with the two new DOA configurations. Results of Experiment 2 showed that with the improved automation design in the high DOA configuration, the ecological display significantly improved fault detection performance. A consistent pattern of risk preference was found in this study as in Experiment 1.

A comparison of the two experimental studies showed new opportunities to derive automation and display design from the DOA layering approach that can support fault detection performance in automated trading, and future research is warranted to explore the influence of ecological displays on risk preference. The development of AUTRASS also makes a unique contribution. The design of the different DOA configurations demonstrated the applicability of the DOA layering approach to guide automation design.

Overall, the following conclusions were reached by conducting this dissertation research: CWA has been effective in characterizing the complexity in automated trading that is associated with the variable DOA, and it can further support the design of automation and the ecological displays. Ecological displays may foster risky operation with moderate DOA as well as performance improvement with high DOA. The applicability of the proposed approaches spanning modeling, design, and evaluation should go beyond the limit of automated trading to the brave new world of artificially intelligent automation.

Acknowledgements

"*An Unexpected Journey*"

- J. R. R. Tolkien

First of all, I wish to thank my supervisor Catherine M. Burns for bringing me into the field of cognitive human factors. Her assistance, patience, and advice kept my research on the correct track. I would also like to thank my committee members, David B. Kaber, Shi Cao, Kumaraswamy Ponnambalam, and Alan Huang, for their wonderful comments on this dissertation.

This dissertation is a bold move on an extremely underexplored topic. I would like to acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Quantica Trading Inc. Special thanks to Xian Wang of Shenzhen Platinum Venture Capital and Rui Hu of Microsoft Corporation. This dissertation cannot be made possible without their knowledge, encouragement and friendships during the hardest time. Thanks to Travis Felker, Roni Hoffman and Drew Coles for providing their expert knowledge, Moe Omer, Aaron Clasky, Aditya Dahiya and Geoff Corvera for providing technical support and Ji Min Kim, Rachel Cao and Adeline Tian for moderating the experiment sessions. Thanks to all the research participants at the University of Waterloo, who spent their valuable time participating in the experiments.

I would like to thank Robert Arrabito and Geoffrey Ho of Defense Research and Development Canada, Toronto, for being my pathfinders in my research career. Special thanks to Carla España Lynch at Facebook Inc. for being the best intern mentor, leading me to the gate of user experience research and shaping my view of humanity and life.

I want to express my gratefulness to current and former members of the Advanced Interface Design Lab: Plinio P. Morita for being the most humorous lab manager and the best non-Chinese hot pot chef, Wayne Giang for helping me survive my master's, Behzad Aghaei for impressing me with sushi-eating and violin-playing, Maryam Ashoori and Vivek Kant for shaping my view on theoretical research, Leila S. Rezai for discussions on work and life, Anson Ho for exploring California together, Jessie Chin for being the best role model for researchers, Dev Minotra for completing the most difficult project together, and Damla Kerestecioglu and Murat Dikmen for sharing the joy and sorrow. Thanks to my old friends of the University of Waterloo Human Factors Group: Jingru Yan, Samantha X. Yuan, Cleyton de Vargas and Y.-L. Betty Chang.

Last but not the least, I am thankful for beloved Sisi Chen and the best cat Niu-Niu who have made all my efforts meaningful.

Dedication

To the late Wei Zhou (1924 – 2012), a sinologist and grandfather.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Full Name
CWA	Cognitive Work Analysis
EID	Ecological Interface Design
NDM	Naturalistic Decision Making
DOA	Degree of Automation
SA	Situation Awareness
WDA	Work Domain Analysis
ConTA	Control Task Analysis
StrA	Strategy Analysis
SOCA	Social Organization and Cooperation Analysis
WCA	Worker Competency Analysis
AH	Abstraction Hierarchy
DL	Decision Ladder
SRK	Skill-, Rule, and Knowledge taxonomy
AUTRASS	AUtomated TRAding System Simulation
DURESS	DUal REservoir System Simulation
NASDAQ	National Association of Securities Dealers Automated Quotations
API	Application Program Interface
SAGAT	Situation Awareness Global Assessment Technique

Dissertation Outline

This dissertation presents a research program that consists of four parts and a series of appendices.

Part A lays the foundation for this research program. The author provides the motivations, develop the research questions, and justify the research methods.

Part B presents the modeling work of this research program.

Part C develops a design approach based on the modeling work and presents two experimental studies to evaluate the design approach.

Part D concludes this dissertation, summarizing key contributions and discussing future work.

The appendices of this dissertation include a literature review, materials for the experimental studies, notes for the data analysis, and the curriculum vitae of the author.

Part A

Background

The first part of the dissertation is the introduction to the research program being proposed. In chapter 1, the author introduces the motivations of this research. Inspired by these motivations, the author proposed three research questions that will be examined in the rest of this dissertation. the author then provides a summary of research methods and an overview of this dissertation.

Chapter 1

Introduction

1.1 Motivations

1.1.1 The Era of Intelligent Automation

The era of artificially *intelligent automation* is here (Sheridan, 2017). Autonomous cars transport passengers to destinations (e.g., Dikmen & Burns, 2016; Endsley, 2017; National Highway Traffic Safety Administration, 2013). Conversational agents, such as Siri, Cortana, and Alexa, facilitate conversations and perform real-world tasks (Luger & Sellen, 2016; Pogue, 2012; Solaimani, Keijzer-Broers, & Bouwman, 2015). IBM Watson computers mimic various aspects of human brains to handle complex cognitive computing works (Modha et al., 2011), reshaping the worlds of scientific research (Chen, Elenee Argentinis, & Weber, 2016), healthcare (Fortune, 2016), and most recently manufacturing (Reuters, 2017). From steam turbines to intelligent automation, automation has become more pervasive, and the degree of automation (DOA) is increasing.

Human factors research proposed that automation could use four distinct stages of information-processing (i.e., information acquisition, information analysis, decision and action selection, and action implementation, Parasuraman, Sheridan, & Wickens, 2000). A higher DOA can be achieved by using both “later stages and higher levels within stages” (Onnasch, Wickens, Li, & Manzey, 2014), suggesting that intelligent automation may augment human decision-making and independently execute action choices. With increasing DOA, automation has benefits for reducing operator workload and improving system performance when the automation is reliable. However, increasing the DOA may degrade operator awareness and hurt system performance when the automation fails. There is the *automation trade-off*, statistically describing the relationship between automation and its human operator (Onnasch et al., 2014).

Based on machine learning algorithms, intelligent automation spans these last two stages to make sense of a large amount of information collected through sensors and to make sophisticated decisions. Although automation is more capable of taking over sophisticated tasks from human operators as it gains more independence, there may still be occasional but crucial human intervention which is "feeding in goals, criteria and other value information" (Sheridan, 2017). If, ultimately, the goal of human factors engineering research in the intelligent automation era is to facilitate "the design of interactive systems of people, machines, and environments to ensure their effectiveness, safety, and ease of performance" (Human Factors and Ergonomics Society, n.d.), efforts must be made to identify intervention opportunities for human operators and to develop support for human operators working with automation in a coordinated way. Arguably, it is logically necessary to start this endeavor by understanding a real-world system.

1.1.2 Automated Trading: A Complex Socio-Technical System

The legacy of the human factors research lies in safety- and life-critical domains (e.g., aviation, nuclear, and medical, for a review, see Parasuraman & Wickens, 2008). Arguably, there should be more attention in other domains which might not be life-critical, but have a significant social impact that has not been well explored in the literature.

The theoretical foundation of this dissertation (Part B) is built to understand a largely unexplored domain: automated trading. The trend of intelligent automation is emerging in the world of finance, where automated trading has started to take over thousands of Wall Street jobs. Automation in financial trading is not life-critical, but has a substantial social impact due to its vulnerability to technical, economic and political disturbances. Automation uses sophisticated computerized algorithms, powerful computers, and rapid telecommunication technologies, and has significantly changed the ways of financial trading. Automated trading is responsible for more than 50% of the trading volume of the

United States stock markets (Iati, 2009). It can capture extremely short windows of trading opportunities (seconds or milliseconds in the case of high-frequency trading, HFT) and complete many transactions with minimal human intervention. Today, artificial intelligence is being adopted by highly automated trading systems, with machine-learning based algorithms, making multiple market predictions and action choices based on different sources of information, and voting on the best course of action choice (Metz, n.d.). While Tesla cars, Siri and IBM Watson computers are under the spotlight and have prompted many debates about the rewards and risks of automation (e.g., "ban on human drivers", Dredge, 2015), on the other hand, relatively less attention has been paid to similar issues with automated trading (for a review, see Treleaven, Galas, & Lalchand, 2013).

From a human factors research point of view, studying automated trading would make unique contributions to understanding the social impact of automation. The social impact of automation has become more pervasive in automated trading and intelligent automation in general but has not been well addressed with legacy, life-critical domains in the literature. The DOA in automated trading is not only increasing as the technology advances, but can vary through the regulation, the knowledge and intention of the algorithm designer, and the technological capability. The flexibility in trading algorithm design brings in increasing regulatory pressure, as the abuse of automation is a serious disturbance that may lead to significant market crashes (e.g., the 2010 Flash Crash: Minotra & Burns, 2016; *N.D. Ill. v. Sarao, United States District Court Northern District of Illinois Eastern Division*, 2015). On the other hand, the flexibility in algorithm design creates a "*quality arbitrage*" phenomenon where automation profits from competitors who use less powerful technologies (Kumiega & Van Vliet, 2012). Trading algorithm designers always have different resources for developing trading algorithms, and the quality difference between trading algorithms will only be amplified when more artificially intelligent automation is onboard.

Automated trading, a domain with varying DOAs, has high complexity. It is apparent that the finance domain, in general, has become more technologically complex as the automation advances through time. Since also the DOA can vary through trader's skills and intentions, the coupling between trading algorithms and traders increases, suggesting that automated trading is not a pure technological domain. In the prior work, the author identified automated trading as a complex socio-technical system (Li, Burns, & Hu, 2015), similar to domains that have been well studied in the human factors literature (e.g., aviation and process control). The complexity of automated trading can be described using all eleven characteristics of complex socio-technical systems suggested by Vicente (1999, p. 14). It should be particularly noted that the complexity of each characteristic is amplified as the DOA varies. For example, traders have different knowledge backgrounds so that financial trading, in general, involves heterogeneous perspectives, which is one of the eleven characteristics. As part of the trading process, the identification of chart patterns (e.g., price chart) to predict market movements is a subjective process and should be related to the trader's skill and knowledge, as Murphy commented (1999): "The truth of the matter is that charting is very subjective, chart reading is an art (possibly the word 'skill' would be more to the point)". For automated trading, trading algorithms typically identify chart patterns using quantitative measures, which are supposed to add less subjective perspective to that charting process, as stated by Kumiega and Van Vliet (2012): "Computers, on the other hand, face no such subjectivity. They can be expected to follow the rules, and they can form objective, unbiased estimates of risk". Trading algorithms may not be as subjective to decision bias as traders. However, the fact is automation and traders must work coordinately in a tightly coupled manner. At a low DOA, automated trading uses ad-hoc trading algorithms which strictly follow the rules set up by the trader. Therefore, these trading algorithms are merely representations of the trader's perspective, and carry over the trader's market estimates with bias. The rules that the ad-hoc trading algorithms follow may prove excessively cumbersome to adapt to unanticipated market situations, which obviously requires human intervention.

Highly automated trading algorithms, typically machine-learning based, are potentially more adaptive to unanticipated market situations. However, like in other machine-learning based systems (Alaieri & Vellino, 2016), these trading algorithms may produce unpredictable consequences that may not be easily explained by traders, and at the end of the day, any monetary losses will be in the trader's wallet.

In Part B of this dissertation, the modeling of automated trading is demonstrated using cognitive work analysis (CWA). CWA is a theoretical framework for analyzing functions and constraints for complex socio-technical systems. Despite its origin in process control industries (e.g., nuclear, Rasmussen, 1986; Vicente, 1999), CWA has received increasing attention in social domains (e.g., healthcare, Bisantz & Burns, 2009; Jiancaro, Jamieson, & Mihailidis, 2014). So far, the author's prior work has been the first to apply CWA to modeling automated trading (Li, Burns, & Hu, 2016; Li & Burns, 2017; Li et al., 2015). In particular, Li and Burns (2017), thoroughly described how to model automated trading systems and automated systems with varying DOAs in general. This paper is incorporated in chapter 5 of this dissertation.

1.1.3 Ecological Displays to Support Monitoring Performance for Automated Trading

To study monitoring performance in trading, an important facet of automated trading, the author explores the ecological interface design (EID) approach to develop graphical displays. Based on CWA models, EID is an approach that copes with the complexity of socio-technical systems and helps to design graphical displays to improve monitoring performance. In this regard, EID supplements CWA and would extend the scope of this dissertation to improving automated trading software with better display design. As reported in the literature, EID graphically represented constraints that can be extracted from CWA models, and has improved operator monitoring performance in certain cases in comparison to conventional displays (e.g., nuclear power station: Lau, Jamieson, Skraaning Jr., & Burns, 2008; petrochemical plant: Reising & Sanderson, 2000a, 2000b). In these examples, the conventional displays were industry state of art (Vicente, 2002) that represent the physical structure of the domain and typically mimicked plant diagrams (e.g., pipeline and reservoir). The ecological displays were developed to add functional information to the representation of the physical structure to support knowledge-based problem-solving. Although both the conventional and the ecological displays could effectively support monitoring for anticipated situations (Vicente & Rasmussen, 1992) where procedures to cope with these situations are documented by the system designer, the ecological displays were particularly useful in unanticipated situations where procedural support is not always available (Lau, Jamieson, et al., 2008). The unanticipated situations are likely where automated trading lies, as financial markets are so dynamic and trading algorithms are extremely flexible. On the other hand, human factors research into the finance domain has been rare and so far at the analysis/modeling phase (e.g., Achonu & Jamieson, 2003; Leaver & Reader, 2015; McAndrew & Gore, 2013; Minotra & Burns, 2016; Sundström & Hollnagel, 2011); therefore, a gap exists in supporting monitoring performance for automated trading.

To fill this gap, built on the CWA models the author has developed in Part B, Part C of this dissertation further explores design concepts following the EID principles in a computer simulation of automated trading with human participants. The author evaluated ecological displays that might support the improved monitoring performance in unanticipated situations. In the literature, the primary measures for evaluating the efficacy of ecological displays are task performance, awareness, and workload (Burns et al., 2008; Lau, Skraaning Jr, Jamieson, & Burns, 2008). These measures can be used to evaluate the effects of ecological displays in an automated trading setting, and if there is an effect, in which DOA this effect exists.

1.1.4 Ecological Displays to Influence Traders' Risk Preference

For evaluating ecological displays, a unique opportunity lies in adding new measures pertaining to trader's risk preference from the behavioural finance research to the evaluation of ecological displays.

For example, *prospect theory*, the foundation of behavioural finance research, described risk preference as people might perceive more pains at a prospect loss than an equal amount of prospect gain (Kahneman & Tversky, 1979). Risk preferences vary and, therefore, result in different description-based choices, whereby people were presented with different outcomes (i.e., prospect gains and losses) and their probabilities. Each probability distribution was associated with a certain risk level. In the real world, however, people make choices based on their experiences, and no outcomes and probabilities are explicitly stated (Hertwig & Erev, 2009). A description-experience gap (Hertwig & Erev, 2009) was observed whereby people make different risk preferences in a description-based choice or when "personal observation and feedback from the environment guided the outcomes generated and assessments of their relative probabilities", in other words, an experience-based choice (McAndrew & Gore, 2013). To study risk preference in experience-based choice, Hertwig and Erev (2009) designed a sampling paradigm that allowed people to explore two outcomes and their probability distributions before the risk preferences were examined. People gained experience through learning in Hertwig and Erev's sampling paradigm (McAndrew & Gore, 2013). McAndrew and Gore drew a distinction between the "experience through learning", which could be achieved in Hertwig and Erev's sampling paradigm, and the "experience through professional training", achieved by conducting a series of structured interviews with professionally-trained traders (McAndrew & Gore, 2013).

McAndrew and Gore's research made unique contributions by extending the behavioural finance research to the field of naturalistic decision making (NDM), showing environmental factors in a real-world setting can influence traders' risk preference. Both NDM and cognitive system engineering (Rasmussen, Pejtersen, & Goodstein, 1994), the branch of human factors to which CWA and EID belong, "stress the importance of real-world task settings for capturing and understanding the true nature of human cognition" (Endsley, Hoffman, Kaber, & Roth, 2007). McAndrew and Gore's success in understanding financial trading from the NDM perspective has inspired this dissertation research to explore further whether traders' risk preference can be influenced by other socio-technical factors (e.g., display, automation, and probably the financial trading software in general). The author argues that this further exploration would make a direct impact on the improvement of automated trading software. Since EID should improve operator performance for unanticipated events that are associated with risks, evaluating ecological displays in a financial trading setting can add a new perspective to the behavioural finance research on risk preference in experience-based choice.

In this dissertation research, traders' risk preference as previously observed by McAndrew and Gore was evaluated qualitatively with participants in the simulation study. In addition, traders' risk preference was analyzed quantitatively through understanding what strategies the participants adopted in a simulated financial trading task. The design of the quantitative measures was inspired by Borst, Flach and Ellerbroek's recent observations (2015) in the aviation domain. They found that pilots occasionally made risky decisions with ecological displays, and provided an explanation for this finding: the representation of the deep physical structure on ecological displays made the limits of system performance clear. With these limits directly perceived, pilots might be more likely to go beyond the limits than pilots who were uncertain about the existence of the boundary. On the other hand, as predicted by Borst et al., if the ecological displays were designed to make more aspects of the intentional structure (e.g., safety culture and regulation) visible, pilots may be able to balance the trade-off between efficiency and safety. As a continuing attempt of studying the automated trading work domain, care must be taken in this dissertation to determine whether the intentional structure of the work domain could be considered in the design of the ecological displays, and if not, whether the influence of the ecological displays follows the same pattern of risk-seeking as observed in the aviation domain.

These qualitative and quantitative measures are being explored in Part C of this dissertation.

1.2 Research Questions

Inspired by the motivations presented above, the author developed three research questions for this dissertation:

Research question 1: How can we model automated trading systems with a variable DOA using CWA?

Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

Research question 1 is being examined in Part B of this dissertation. Research questions 2 and 3 will be answered in Part C, with several hypotheses relevant to the experimental design being examined.

1.3 Research Methods

The author used five research methods to examine the research questions: literature review, using expert opinions, modeling, design, and simulation.

1.3.1 Literature Review

Although the author has very specific research questions, there is no doubt that the application domain is broad and unfamiliar to me and the human factors community in general; therefore, studying these research questions required gaining a large amount of new knowledge on automated trading. These knowledge bases are difficult to directly obtain from traders, because financial practitioners are legally required to maintain strict confidentiality of all client information provided to them. There are also practical concerns about ethically accessing trading algorithms which are expensive intellectual products (e.g., a former Goldman Sachs developer is allegedly guilty for stealing trading algorithms, Bloomberg, 2017). These difficulties were partially compensated for by reviewing publicly accessible textual sources (e.g., research papers, books, reports, and news). Further, since the research questions involve many existing research topics with each one alone being sophisticated research, the author reviewed a large amount of literature on human-automation interaction, CWA, and EID. Results of the literature review are presented in Appendix A.

1.3.2 Using Expert Opinions

The author gained a basic knowledge of the automated trading domain, identified theoretical gaps and generated initial research questions through the literature review. To ensure the validity of these research questions, the author was fortunate to receive help from several subject-matter experts whom he contacted from professional and personal sources. During the candidacy, the author participated in a Natural Sciences and Engineering Research Council of Canada (NSERC) collaborative research project with Quantica Trading to train myself as a user experience designer in designing automated trading software. With usability principles in mind, the author was part of a team to design an interactive interface for trading algorithm design and displays for traders who monitor back-testing and live-trading. These design deliverables later became cornerstones of a commercial product. As part of the collaboration, the author worked closely with the company's user experience lead and director of sales, the algorithm manager, and the project manager who previously held job positions in institutional brokerage firms. The modeling part of this dissertation research was completed at the time of this research collaboration, resulting in two co-authored publications with one of the subject-matter experts. The display design and evaluation parts of this dissertation research involve developing trading algorithm prototypes, a computer simulation of automated trading, and quantitative and qualitative measures for evaluation of performance. As the author worked on this part of the dissertation, he was closely guided by a trading algorithm designer working in the futures market.

The influence of expert opinions spanned the entire dissertation, including the modeling, the design, and the evaluation stages; therefore, expert opinions are not reported in a separate chapter in this dissertation. The design work the author conducted at Quantica Trading was commercialized so that its detail has not been revealed in this dissertation. Readers may refer to a media report for public information regarding this commercial product (Leaprate, 2014). The author declare that all development work was based on publicly accessible information and the author report no conflicts of interest with the institutes to which the subject-matter experts were affiliated.

1.3.3 Modeling

The author laid the theoretical foundation of this dissertation by modeling the automated trading domain with an approach that can handle its increasing social and technological complexity. CWA, a formative modeling approach that has been successfully used to model socio-technical domains (Vicente, 1999), seemed to be a reasonable starting point for understanding automated trading. CWA is a work-centred approach, focusing on modeling how operator behaviour is influenced by constraints in a work environment. Being able to model a variable DOA is an important feature for addressing the coupled relationship between traders and automation which is typical in automated trading; however, this feature has not been explored in the CWA literature. To add this feature to CWA, as part of this modeling exercise, the author proposed a *DOA layering* approach that extends CWA to handle the complexity of the variable DOA. Further, the proposed DOA layering approach, as the rest of this dissertation implies, is not limited to modeling automated trading and should be applied to other automated systems.

1.3.4 Design

Choosing EID as the display design approach is a logical decision because the core value of EID is to use formative CWA models to guide display design. Further, display design should “reveal the deeper structure of the work domain when automation is taking over more and more of the control activities" (Borst et al., 2015), suggesting that EID is potentially suitable in support of display design for automated trading. My DOA layering approach extends CWA to model a variable DOA, a situation that has not been explored in the CWA literature; therefore, the natural next step is to explore new ways of using EID to support improved design for this situation, which could make unique contributions. To design displays following EID principles, the author linked different aspects of the CWA models developed with the DOA layering approach to different types of ecological displays that are designed to support one or more DOAs; further, the author developed the design concepts.

1.3.5 Simulation

The design concepts were developed into displays and evaluated in a lab-control simulation. The author led a team of student software developers at the University of Waterloo and received help from subject-domain experts to develop the AUTRASS (AUtomated TRAding System Simulation) microworld to represent automated trading software used by institutional traders. AUTRASS was designed to be representative of many aspects of the complexity of automated trading, as identified in the author's prior work (Li et al., 2015). AUTRASS consists of a modular frontend that allows for a variety of displays being tested, an order processing back-end supporting both traders and automation, and a data feed using historical data. The development of AUTRASS was inspired by DURESS (DUal REservoir System Simulation), the first microworld designed for evaluating ecological displays (Vicente, 1991).

1.4 Dissertation Overview

This dissertation is an exploration of automated trading from a human factors perspective, with the chapters logically structured into four subjects concerning introduction, modeling, design and evaluation, and a conclusion.

This chapter is the only chapter in Part A. While this chapter briefly introduced the automated trading domain and presented the research questions, this dissertation assumes the readers have no prior knowledge of human factors or finance. Appendix A presents an in-depth literature review of automated trading, human-automation interaction and CWA.

Part B of this dissertation includes chapter 2, and it builds the theoretical foundation for this dissertation. The major part of this chapter is adopted from a paper published in the *Journal of Cognitive Engineering and Decision Making* with minor formatting changes in figure and table numbers and references. In chapter 2, CWA has been used to model automated trading, and during this modelling exercise, a DOA layering approach has been proposed to handle varying DOA situations. A summary of key findings is added to the end of this chapter to connect the development of the DOA layering approach to the other topics of this research program.

Part C presents the design and evaluation of ecological displays in automated trading inspired by the DOA layering approach. The author presents two experimental studies in which the design concepts were evaluated in AUTRASS. Chapter 3 starts with an introduction to the design approach based on the CWA models developed in chapter 2. Automation was designed with two configurations to demonstrate a variable DOA. Conventional displays were industry state of art and should support the basic use of the automation. Ecological displays were implemented based on the CWA models and should support the variable DOA. Experiment 1 evaluated a series of hypotheses derived from the research questions of this dissertation and is reported later in this chapter. An earlier version of this chapter has been accepted by the *2017 IEEE International Conference on Systems, Man, and Cybernetics* in Banff, Canada. Chapter 4 presents Experiment 2 which was a follow-up to the first experiment. Automation design was partially modified to eliminate the flaws identified in Experiment 1 results and to include an adaptive automation condition. A discussion on the results of Experiment 2 is presented with comparisons to the Experiment 1 results.

Part D contains chapter 5 and concludes this dissertation. In this chapter, the author summarizes key findings, highlight contributions of this research program and suggests directions for future research.

Part B

Modeling

Chapter 2 is the only chapter in Part B. The author proposes a theoretical approach that extends several analysis phases of the CWA. Section 2.1 to 2.7 were adopted from a manuscript entitled "Modeling Automation with Cognitive Work Analysis to Support Human-Automation Coordination". This manuscript has been published in the *Journal of Cognitive Engineering and Decision Making*. The theoretical approach proposed in this manuscript is called the DOA layering approach. As its name implies, the DOA layering approach aims at mapping functions allocated to the trader and functions allocated to the automation onto an AH and a DL. The author discusses what additional information and design implications can be captured with this approach in comparison to other ways of representing automation in the CWA literature.

Chapter 2

Modeling Automation with Cognitive Work Analysis to Support Human-Automation Coordination

Overview: Cognitive Work Analysis (CWA) is useful to develop displays for complex situations but has not been well explored in providing support for human-automation coordination. To fill this gap, in this paper, we proposed a degree of automation (DOA) layering approach, demonstrated by modeling an automated trading domain with a future goal of supporting interface design in this domain. The abstraction hierarchy and the decision ladder each adopted an additional layer mapping functions allocated to the trader and functions allocated to the automation. In addition to the mapping, we marked the four stages of automation on the decision ladder to provide guidance on representing the function allocation at the task level. Next, we compared the DOA layering approach to how automation was previously represented in the CWA literature. We found that a DOA-layered decision ladder, which included well-developed knowledge of the stages and levels of automation, can be suited to modern automated systems with different DOAs. This study suggested that the DOA layering approach has important implications for designing automation displays and deciding stages and levels of automation, and may be a useful approach for modeling adaptive automation.

Keywords: Cognitive work analysis; human-automation interaction; degree of automation; stages and levels of automation; automated trading; abstraction hierarchy; decision ladder.

2.1 Introduction

Automated systems are becoming more pervasive and the degree of automation (DOA) that is possible has been increasing. Recently, there has been a growing interest in artificially *intelligent automation* (Sheridan, 2017). The IBM Watson computer that defeated a human chess champion and self-driving cars are two examples suggested by Sheridan (2017). These two examples are highly automated systems powered by sophisticated machine learning algorithms (Sheridan, 2017). At a high DOA (i.e., later stages and higher levels within stages, Wickens, Li, Santamaria, Sebok, & Sarter, 2010), task dynamics are represented by automated features, but humans still take a supervisory role by initiating parameter changes that drive the control of the system. For example, financial trading algorithm designers specify the goals, knowledge and criteria to the development process. After the automated systems are implemented, users must sometimes regain full control when unexpected automation failures or environmental disturbances occur (e.g., the driver must manually reduce speed when a self-driving car enters a road construction zone that is undocumented in the navigation system). Therefore, for highly automated systems using all stages and levels of automation (Parasuraman et al., 2000), there is clearly a need for occasional human intervention (Sheridan, 2017). Understanding where, and how humans will interact with automation remains a challenge.

Two approaches for keeping humans in the loop have been to manipulate the DOA to either avoid high DOA situations, or to provide adaptive automation when users are in varying DOA contexts. These approaches are derived from the *automation trade-off* (e.g., Bainbridge, 1983; Sarter & Woods, 1995). There might be a third approach. A recent meta-analysis suggested that effective, or “ecological” displays may modify or even reverse the automation trade-off, which means a higher DOA could improve both routine and failure performance (Onnasch et al., 2014). Though preliminary, this suggestion has shown that designing better displays for automated systems can be a potentially useful approach. We argue though, that to design effective displays, or to choose an appropriate DOA, one should first develop models of the cognitive work that the user will experience in different DOA contexts. These models can begin to show the functions that the user must take over in cases where the automation must be ended. Further, these models could be used to help to derive the design requirements for displays that can help users work with higher levels of automation, without losing situation awareness.

Consistent with this idea, Kaber, Riley, Tan, and Endsley (2001) suggested that automated system displays must highlight the transition between system states and inform operators of the allocation of control responsibilities. This suggestion aligns with the goal of the Ecological Interface Design (EID) approach, of making control opportunities visible in order to retain skill and awareness (e.g., Borst et al., 2015; Furukawa & Parasuraman, 2003; Kaber et al., 2001). Borst et al. (2015) recently advanced the understanding of EID applied to automated systems, suggesting that ecological displays should coordinate with the increasing DOA by providing more information to support human-automation coordination. However, stronger approaches are needed to help determine what that information should be.

In this paper, we propose an approach to transform knowledge from the stages and levels of automation model to design requirements that could promote human-automation coordination. By integrating the stages and levels of automation model into an analysis, we can discover important properties of human-automation interaction that could be represented in better designs. Since Cognitive Work Analysis (CWA) has shown success in determining requirements for complex systems, it makes sense to explore how CWA could be used more effectively to generate design requirements for automated systems. In this paper, we demonstrate our approach in an automated trading domain. Financial systems present a fertile domain to explore human decision making, with complex dynamics and increasingly pervasive automation. There have been some, but not many human factors studies on financial systems in general (e.g., behaviour and performance modeling: Achonu & Jamieson, 2003; McAndrew & Gore, 2013; systematic safety: Sundström & Hollnagel, 2011; incident analysis: Leaver & Reader, 2016), but none specifically analyzed automated trading. Studying automated trading presents many potential research opportunities. First, applying CWA to automated trading expands the application of CWA to a complex market-based domain that operates on different principles from physical systems (e.g., process control). Second, studying automated trading presents an opportunity to address the automation trade-off in a financial domain. In financial markets, the majority number of transactions are now completed with automation technologies, mostly by using sophisticated trading algorithms (Iati, 2009). While trading algorithms improve the human ability to utilize small profitable opportunities (e.g., a small trading window may only last seconds or milliseconds), traders may encounter attentional failures while interacting with the trading algorithms or intentionally abuse the trading algorithms. An example of attentional failures in financial trading is a slip or lapse (e.g., Leaver & Reader, 2016), and an example of abuse of automation is "spoofing" - illegally profiting from market manipulation by generating fake supply or demand (e.g., *N.D. Ill. v. Sarao, United States District Court Northern District of Illinois Eastern Division*, 2015). The last research opportunity lies in the great flexibility in developing trading algorithms. Traders may develop trading algorithms using all stages and levels of automation. Here we give two examples of trading algorithms with different DOAs. The first example is high-frequency trading, using a rigid execution algorithm to trade in milliseconds. This algorithm typically has a high DOA that requires minimal human intervention, and therefore, it introduces a new risk of magnifying market value losses. For these reasons, high-frequency trading systems have received increased regulatory pressure, such as traders who utilize high-frequency technologies are being closely monitored by the regulators (Fabozzi, Focardi, & Jonas, 2011). In certain cases, traders may be more inclined to move towards developing intermediate DOA algorithms or manual trading (Li et al., 2015). As another example, it has been reported in the literature that traders using more advanced algorithms may completely outperform and profit from their competitors, who are equipped with less advanced technologies. This is a phenomenon known as "quality arbitrage" (Davis, Kumiega, & Van Vliet, 2013). We summarize the research opportunities discussed above: the complexities of automated trading suggest that human factors research in this area could contribute to both new understandings of human decision-making, and improvements to financial trading software.

An investigation on how traders interact with different DOA algorithms could improve the understanding of automated trading, and automation in general.

The remaining part of this paper is organized as follows. We first introduce two automated trading scenarios, one for low DOA the other for high DOA. After that, we propose a *DOA layering approach*, showing how CWA can be used to model these scenarios by extending the Work Domain Analysis (WDA) and Control Task Analysis (ConTA) to account for DOA. As part of this work, we discuss the difference in the proposed DOA layering approach and how automation was represented in the CWA literature. We then discuss the implications of using the DOA layering approach - implications for display design and implications for automation design. Finally, we discuss a future application for the DOA layering approach would be to represent function allocation that changes during DOA shifts, for modeling adaptive automation.

2.2 Automated Trading Scenarios

Two financial trading scenarios, basket trading and trend following trading, are used for this analysis. The two trading systems differ in their DOAs and were mainly inspired by the knowledge obtained from a literature review (e.g., Chan, 2009) and a discussion with subject-domain experts.

1. *Low DOA scenario: basket trading*. Basket trading systems are popular in the institutional trader community. To use a basket trading system, the trader first configures a data analysis and order generation algorithm to create a shortlist of financial products for trade. The trader then executes the algorithm to generate a basket of orders. On the completion of all orders in the basket, the trader may adjust their portfolio holdings without altering the portfolio allocation. As part of the purpose of basket trading, the basket of orders should be executed simultaneously, though price movements of the financial products are quick. The basket of orders must go through a trading platform in order to reach the market exchange. The trading platform is either provided by the

trader's brokerage firms (e.g., Interactive Brokers), or it is broker-neutral software (e.g., Bloomberg Terminal). Chan (2009) described the basket trading system as typically running "only a few times a day in order to generate one or a few waves of orders". This description showed that the basket trading system is a low DOA, semi-automated system. The asynchronous nature of basket trading (e.g., collecting data and generating orders) is related to information analysis and decision-making. In normal conditions, the order execution is synchronous with the financial market. In other cases, if it is not possible to execute all orders synchronously (e.g., in a volatile market), the basket trading system could fail. The trader may also make a wrong decision on the proportions of the financial product in the basket.

2. *High DOA scenario: trend following trading*. Trend following trading systems are a real-time trading system, typically based on a sophisticated technical analysis (e.g., Moving Averages: Ellis & Parbery, 2005; Bollinger Bands: Bollinger, 2001). Our automated trading experts described a hypothetical trend following system: a trading system uses a "scalping" algorithm based on a Moving Average technical analysis, seeking to make profitable trades based on arbitrage of small price gaps. The algorithm typically goes through a number of trade iterations. Once a trade iteration is completed, another iteration will begin automatically, limited only by a total number of iterations defined by the trader. The algorithm has distinct buy and sell logic. For example, the algorithm would wait to confirm a buying signal that the 50-day Simple Moving Average (SMA) crosses above the 200-day SMA on the day candles and the Relative Strength Index (RSI) in an oversold territory is below 30. Once a buying signal is identified, the algorithm would place multiple buying orders in 10 iterations to the market, buying a random quantity between 400 to 800 shares in each iteration. To use a scalping algorithm, the trading platform must perform real-time data collection, automatic decision-making, and rapid order placing. The scalping algorithm is perfect for exploiting a small market opportunity repeatedly

without manually re-launching the trading system. The trader typically evaluates the performance of the trend following system using a set of measures, such as Sharpe ratio, total profit and loss, and commissions. The trader has authority over monitoring every trade made by the scalping algorithm, but the monitoring is not required. The trader would typically intervene when the trading system achieves expected revenue, or when the scalping algorithm needs a performance upgrade. However, the trader may override the autonomous operation, if an algorithm bug or market disturbance occurs, by canceling or modifying an order, closing a position (e.g., selling off), or stopping the entire trading system. According to Chan's (2009) description, the trend following system has a high DOA.

2.3 Using the Work Domain Analysis to Model Automation

We have modeled these two automated trading scenarios using WDA. We first build a base AH from the domain, and then propose a DOA layering approach for representing the DOA.

2.3.1 Base AH

We propose a base AH should be developed as is typically done in CWA (Rasmussen, 1986; Vicente, 1999). The base AH should include the usual five levels of abstraction as in Rasmussen and Vicente's original AH approach. The scope of the base AH is limited to the system under control by the user or the automation and does not include the automation. Once developed, the base AH can serve as a template for mapping the influence of automation on the domain.

We developed a base AH to represent the financial trading domain, using the two automated trading scenarios. Since the descriptions of the scenarios are generally task-specific, we reviewed the scenarios with our subject-domain experts and distilled the scenarios into domain functions (e.g., the functions of buying and selling in both scenarios). Later, the domain functions were organized to fit the five levels of abstraction, excluding DOA-specific functions (e.g., the basket of orders in the low DOA scenario and the Moving Average technical analysis in the high DOA scenario).

As a result, the base AH shows the flow of securities is largely about buying and selling and is governed by principles such as the law of supply and demand and the flow of capital. In the next paragraphs, the base AH is described in more detail, with list numbers correspond to labels in Figure 1:

1. Functional Purpose shows the purposes of trading. Financial activities have a commonly accepted goal that is to make a profit. At the same time, financial activities receive regulatory constraints such as market principles and laws. The regulatory constraints shall ensure traders and automation are seeking to profit in legal ways;
2. Abstract Function defines principles, priorities, and values to follow in achieving the Functional Purpose. We identified two groups of Abstract Functions: financial decision-making principles, and market constraints. Financial decision-making principles include the law of supply and demand, a law governs that financial activities at the most fundamental level phrase in Adam Smith ‘s 1776 book *The Wealth of Nations*. We identified a priority that to balance gains and losses, acknowledging that the ideal balance point of gains and losses would interact with the profit goal as well as the acceptable risk level. For example, a trader may aim to maintain a diversified portfolio to protect the trader against the risk of volatility, while other traders may seek higher profits at greater risk. Financial products must be traded ethically according to the values of the trading system. Otherwise, there could be ethical problems, such as the market crash (Davis et al., 2013); In the second set of abstract functions, market constraints, we have represented the flow of capital, market information, and laws and regulation. The flow of capital influences trading in that no trader can trade beyond their authorized capital limit, and capital must flow between market participants to keep the market liquid. Market information must also

flow to enable decisions, following a certain protocol. For example, the Financial Information eXchange (FIX) protocol is the *de facto* electronic communication standard protocol introduced in 1992. The FIX protocol regulates the information flow in a financial market, exchanging real-time trading data related to securities, orders, and trades between traders and brokers (Hu & Watt, 2014). Further, the markets are subject to regulations and policies that may influence individual trades, securities, and market behavior as a whole;

3. At the Generalized Function level, we identified four main processes: 1) to buy, resulting in position gains of a portfolio; 2) to sell, resulting in position losses of a portfolio; 3) to obtain market information such as quotes and order books and 4) to develop successful trading strategies;
4. The Physical Function level shows physical components, including 1) exchange, a computerized auction market (e.g., New York Stock Exchange). Traders and automation may have access to multiple exchanges, allowing them to execute arbitrage strategies across exchanges; 2) buyer and 3) seller. They can be traders or automation representing a trade client; 4) securities, identifying which financial products are being traded. Multi-asset trading platforms use multiple securities at the same time; 5) order, showing instructions of a trading action. A bid order represents increasing a position. An ask order is used to decrease a position; 6) account and 7) position, showing trader's capacity in the form of cash and assets; 8) intermediaries, which are normally brokers offering services to a number of trade clients and market exchanges;
5. The bottom level, Physical Form, shows the operational conditions, or attributes. There are five categories: 1) cost, including variable and fixed costs to trade; 2) time, showing the life cycle of a trading strategy, market and order time, and latency; 3) state of the market and the position; 4) price, including market price, order price and price of portfolio in a certain currency and 5)

volume, including market, order and position volume in shares. Many of these attributes can be seen directly through the trading platform visualizations.

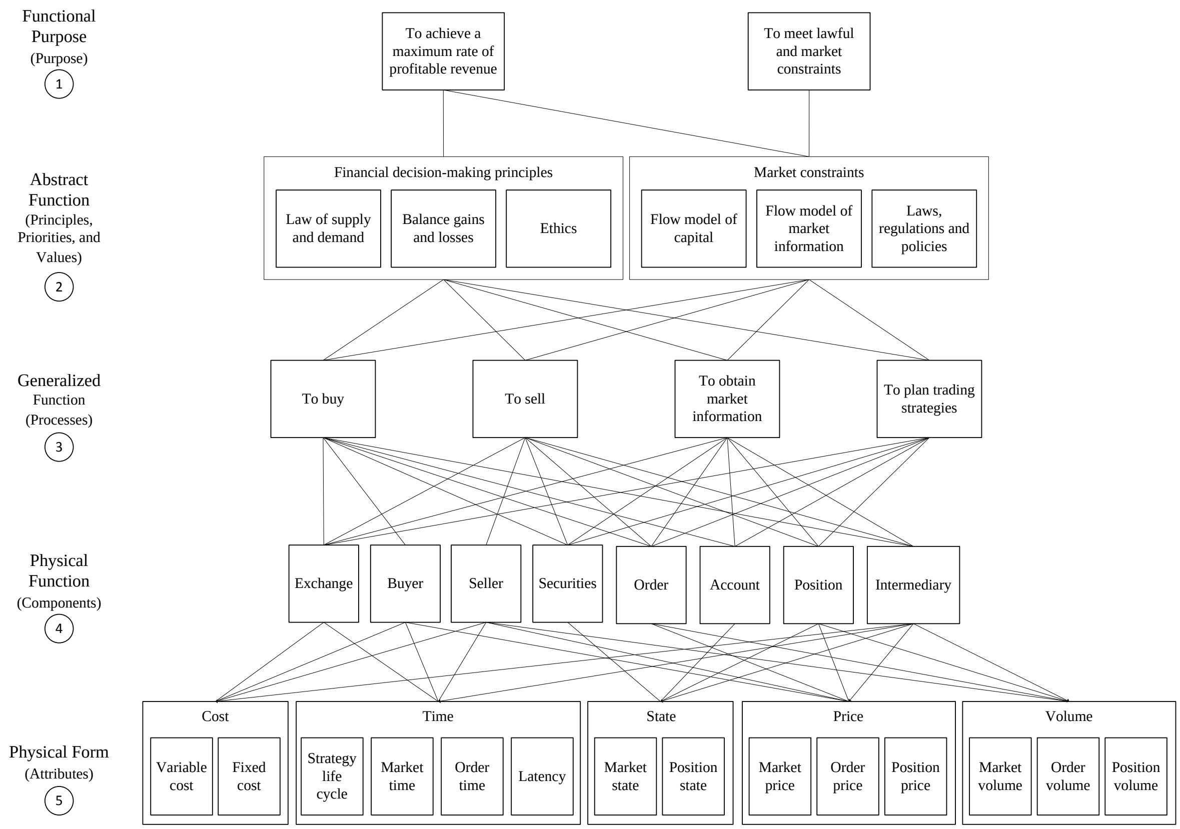


Figure 1. Base AH of financial trading.

2.3.2 DOA Layering on the Base AH

We have modeled the financial trading domain as broad as possible, thereby representing both automated financial scenarios with the same base AH. A consistent base AH can be used as the common ground for portraying DOA-specific information that was excluded from the base AH.

Having completed this model, we propose the DOA layering approach, layering automation on the base AH. The key to this approach is to identify the responsibility of each function in the base AH. A function in the base AH can be represented as either a sole responsibility of a trader or automation, or a shared responsibility. For simplicity, in the following examples of DOA layering, we represent functions that are solely allocated to the trader or the automation, excluding shared function allocations.

The function allocation was based on domain knowledge, with much of the knowledge coming from the literature review (e.g., Chan, 2009), ethnographic experience at a trading software company, and discussions with traders on staff at the company. The first author had been involved in an observational study at Quantica Trading Inc., an automated trading software company based in Kitchener, Canada. He was part of a multidisciplinary team, including staff traders, to redesign an automated trading platform. Details of this observational study were reported in a previous paper (Li et al., 2015). In certain cases, details that would be instrumental in determining the function allocation were not available in the literature. Particularly in this domain, details about a trading system are rarely publicized, as the finance industry is unique for its strict confidentiality and protection of institutional clients. This unique characteristic of the finance industry also led to a significant limitation in being able to directly observe professional traders. To mitigate these concerns, we discussed with subject-matter experts, staff traders, available at the company about function allocations that were missing in the two scenarios. For example, the high DOA scenario suggested that the “to achieve a maximum rate of profitable revenue” Functional Purpose would be allocated to the automation. The “to meet lawful and market constraints” Functional Purpose was not described in the scenario literature but, discussions with the subject matter experts suggested this function was best allocated to the trader.

As shown in Figure 2 and 3, functions of the base AH were assigned shades. Functions allocated to the automation were shaded and functions allocated to the trader were not shaded.

Figure 2 shows the low DOA function allocation. We can see the higher levels, namely, the Functional Purpose and the Abstract Function, are solely allocated to the trader. The trader is responsible for deciding the proportion of each financial product in the portfolio allocation. The lower levels, Generalized Function, Physical Function and Physical Form, are allocated to both the trader and the automation. The automation is not capable of controlling all aspects of trading, thereby requiring trader involvement.

In Figure 3 we present the high DOA function allocation. While the automation continues to share functions at the lower levels with the trader, it also plays a role in controlling functions at the higher levels. For example, the scalping algorithm is responsible for ensuring the profitability of the trading system (Functional Purpose). The algorithm may realistically achieve this purpose by balancing gains and losses (Abstract Function), even as the trader exercises authority over other Abstract Functions (e.g., “ethics” and “laws, regulations, and policies”). To manage both Generalized Functions of buying and selling, the algorithm must accurately choose entry and exit points into the market. A broader base of information is being considered by the algorithm, such as the market price, the order price and the latency of the order (Physical Functions and Physical Forms).

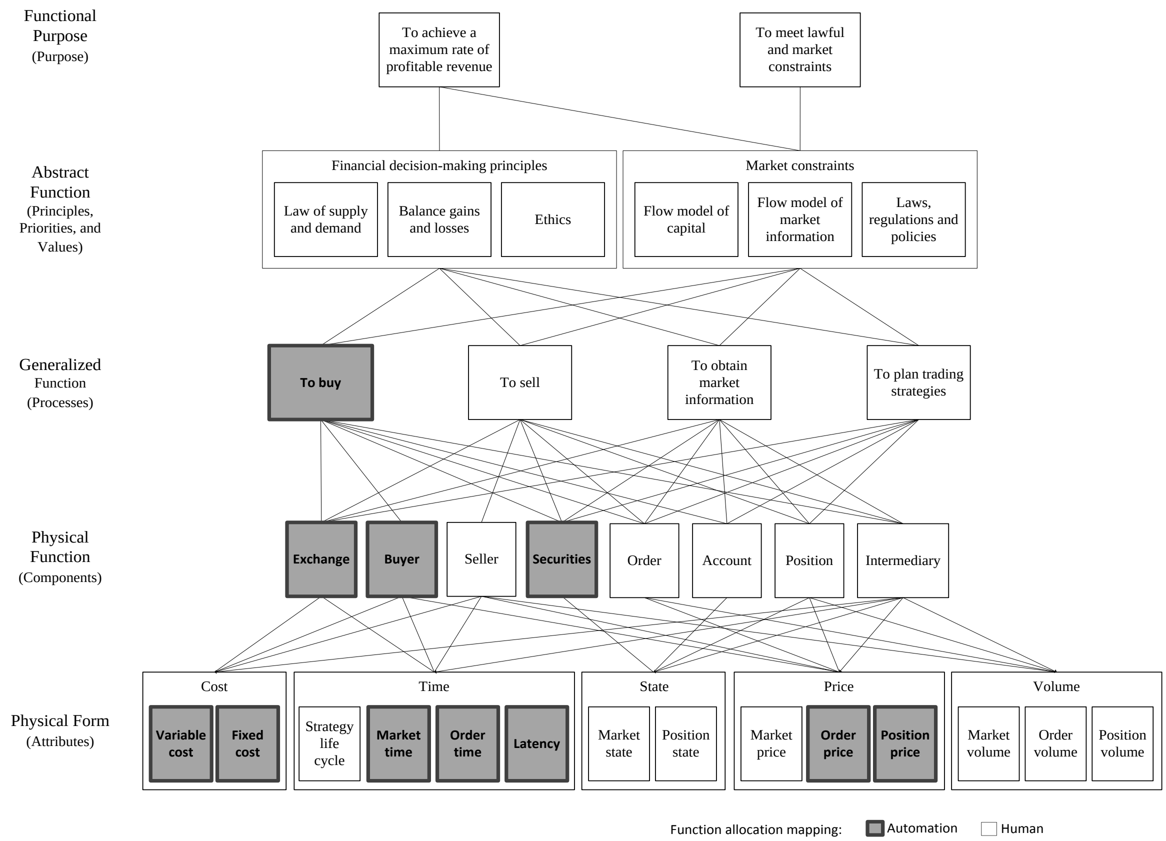


Figure 2. AH of basket trading (low DOA).

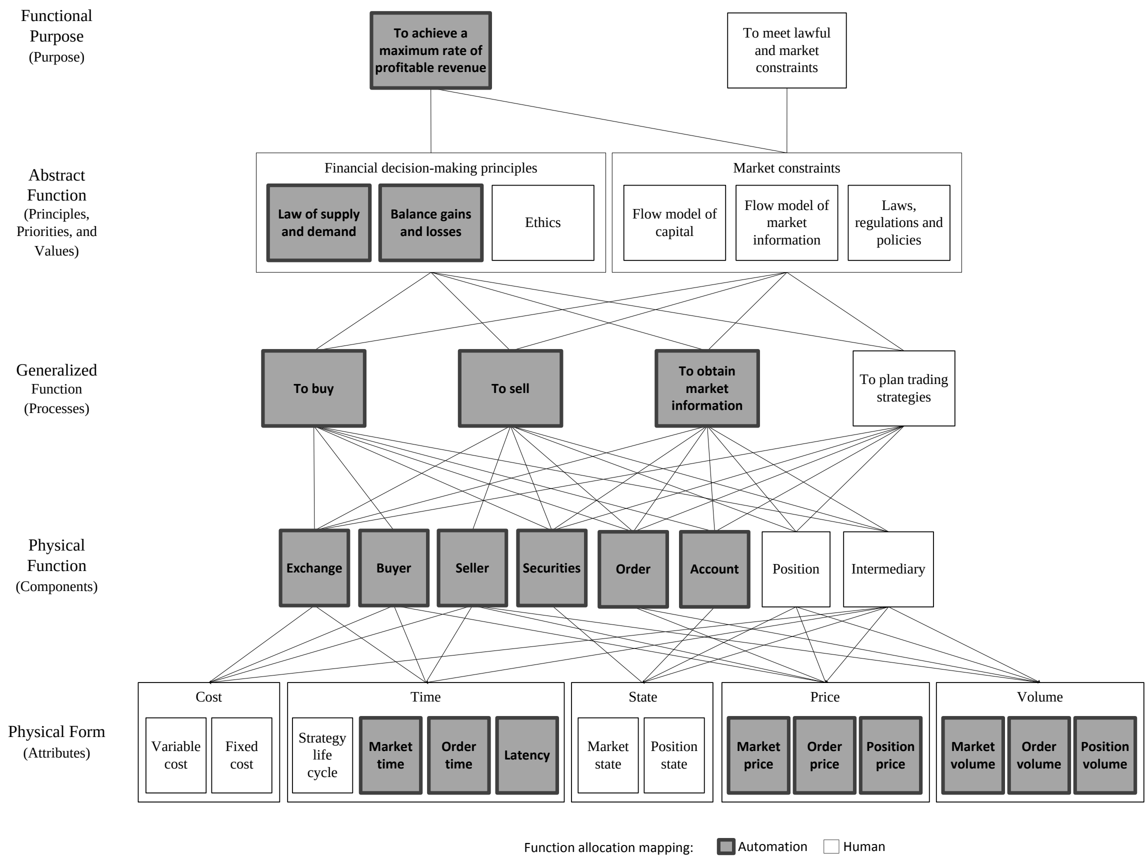


Figure 3. AH of trend following trading (high DOA).

The DOA layer adds a new dimension to the base AH, showing how human and automation work collaboratively at a certain DOA. Functions can be allocated to any actor of the work domain (human or automation). Shared allocations could also be included in the DOA layer, though the analyst may want to differentiate between shared allocation approaches. The greater breadth of physical functions and associated attributes of the DOA layer can show where situation awareness losses might occur. Effective salient display of this information, and the operation of the functions being performed by the automation may inform more effective displays in this situation.

2.4 Using the Control Task Analysis to Model Automation

The WDA, as was described in the previous section, can be used to map the function allocation of automation on domain structure, while the other phases of CWA can be used to illustrate the behavior of the automation. In particular, the ConTA looks at how information is processed, mapping those task stages on to the decision ladder (DL) and exploring various shortcuts that are possible in processing (McIlroy & Stanton, 2015; Vicente, 1999). In this section, we propose to utilize the stages and levels of automation information while conducting a ConTA. We first discuss how to represent the four stages of automation on a base DL. This base DL is a template having the usual ladder structure and shortcuts as in an original DL. We then model four cases using a layering approach on the base DL. The four cases include two automated trading scenarios (the low DOA and high DOA), each in two situations (routine operations and unanticipated situations).

2.4.1 Representing Four Stages of Automation on the Base DL

We divided the DL over four regions and mapped the four stages of automation on to the ladder. Automation of four stages includes acquisition automation, analysis automation, decision automation and action automation (Parasuraman et al., 2000).

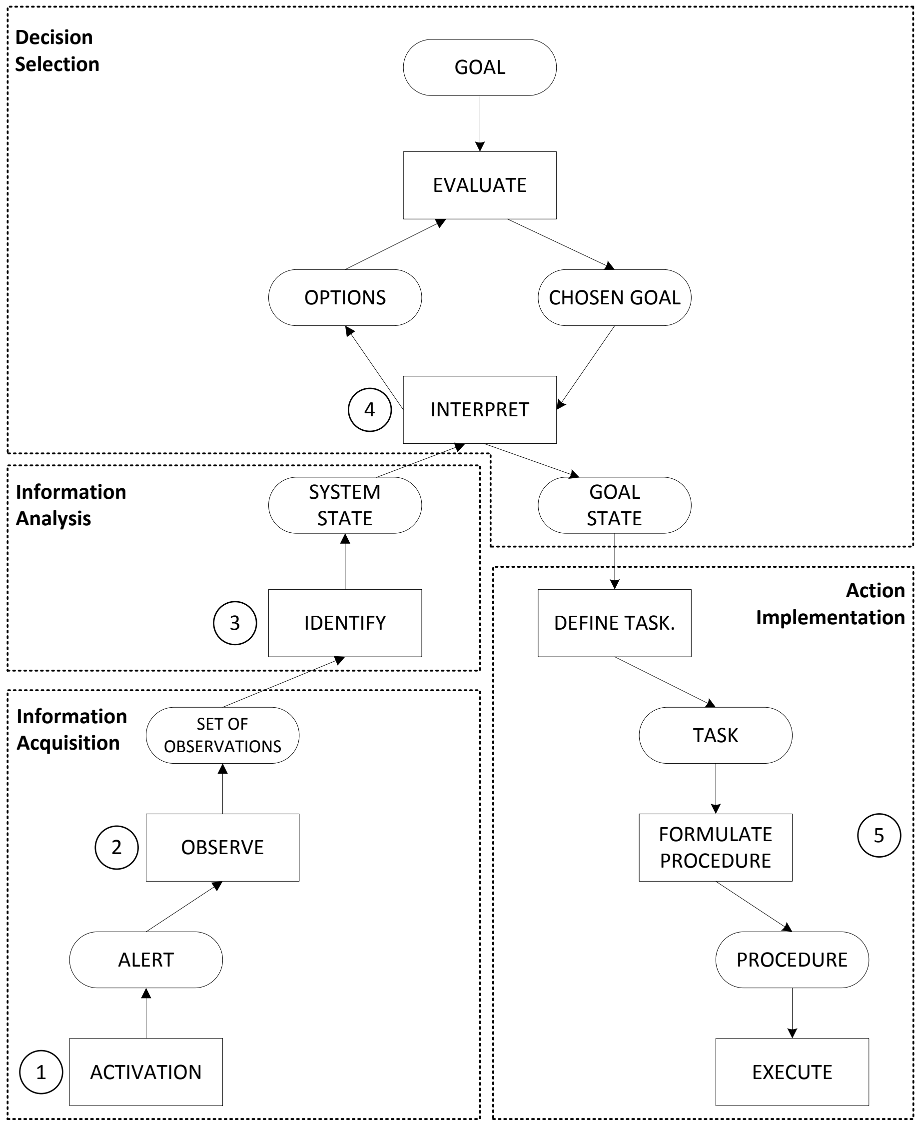


Figure 4. Representing stages of automation on a DL.

We describe DL steps (Rasmussen, 1974) and their affiliation with the four stages of automation, with list numbers correspond to labels in Figure 4. The following points explain the justification for this mapping and connect DL and stages of automation contents to financial trading examples. We correlate the DL steps to the five levels of abstraction of the base AH we previously presented.

1. *Activation*. The DL may start when traders are notified by environmental signals in the market (Physical Form: “market price”). If this DL step is automated, acquisition automation receives real-time quotes from the market (Physical Form: “market price”) when the market is open and a reliable data connection is established;
2. *Observe*. Traders observe alerts from the previous step (Generalized Function: “to plan trading strategies”) and reduce noise to form a set of observations (Physical Functions: “exchange”, “securities”, “account”, “position” and “intermediary”), based on a subconscious mental model. If this step involves acquisition automation, it will become an automated data processing step based on a pre-defined rule. For example, an algorithm prioritizes stocks depending on their volatility (Generalized Function: “to plan trading strategies”), then presents the priority list to the traders for further research;
3. *Identify*. At this step, traders identify the underlying state of the trading system (Abstract Function: “flow model of capital”). For example, traders may correlate the current state to a previously experienced state. In aviation and process control domains, trend displays are provided in analysis automation to help the operators make sense of the available information (Parasuraman et al., 2000). In financial trading, similar tools such as trend line and Moving Average are used to help traders identify market movements (Abstract Function: “flow model of market information”);
4. *Interpret and Evaluate.* Rasmussen (1974) pointed out that human decision-making is a “very complex mental process that requires a high level of abstraction of the domain knowledge” and expert operators may bypass this process if the system state is known. For example, a professional trader may find an association from the current state to a certain chart pattern that leads to a trading opportunity (Abstract Function: “flow model of capital”). A novice or non-trader does not have an ability to bypass the interpretation, and must actively look for possible

options (two Functional Purposes: “to achieve a maximum rate of profitable revenue” and “to meet lawful and market constraints”). Similarly, decision automation is related to varying numbers of options to choose from, depending on the level of automation (Parasuraman et al., 2000). For example, an ad hoc algorithm trades when a market indicator (e.g., price) meets certain criteria (Functional Purpose: “to achieve a maximum rate of profitable revenue”). The algorithm uses “if *x*, then *y*, else *z*” conditional logic. *y* and *z* are known states that can be mapped to decision *a* and *b* separately. If the state is unknown (Functional Purpose: “to meet lawful and market constraints”), there will be no decision. In another example, a machine learning algorithm uses a higher level of decision automation and could be more artificially intelligent than the ad hoc algorithm. The machine learning algorithm can learn without being explicitly programmed with a conditional logic. It has more options to choose from than the ad hoc algorithm does (Functional Purposes). This intelligent algorithm may even create new options by self-learning unidentified system states;

5. *Define Task, Formulate Procedures and Execute*. The right-hand side of the DL describes the execution process, and action automation describes the same. Both manual and automated trading require specific technological details of the trading system to complete the execution process. The process typically involves multiple steps, for example, to define the direction (Generalized Functions: “to buy”, “to sell”), to formulate the parameters (Physical Functions and Physical Forms) and to decide the destination (Physical Function: “exchange”).

Mapping the four stages of automation on a base DL provides guidance on representing the function allocation at the task level. “What is more automation” in each stage (Onnasch et al., 2014) can now be represented by annotating the boxes (steps) in the corresponding DL region. Table 1 is a summary of function allocation as seen in the two automated trading scenarios we have been considering. Functions in each stage are annotated with the names of DL steps.

Table 1. Function Allocation Mapped on the Four Stages of Automation (With DL Annotations in Bold).

Scenario	Stage 1. Information Acquisition	Stage 2. Information Analysis	Stage 3. Decision Selection	Stage 4. Action
Basket trading (low DOA)	Partially automated. The trader manually downloads historical market data (**activation**). A MATLAB algorithm organized the data (**observe**).	Automated. The trading software retrieves fundamental information of the shortlisted stocks (to buy or sell) from a database (**identify**).	Not automated. The trader decides all trades (**Interpret and Evaluate**).	Partially automated. The trader determines the parameters of the orders. Once submitted to the market exchange, the orders are proceeded automatically (**Define Task, Formulate Procedures, and Execute**).
Trend following trading (high DOA)	Mostly automated. A real-time data feed streamlines data collection (**activation**). The trader typically monitors the market data, but is not dependable in the data collection process (**observe**).	Automated. Sophisticated metrics are calculated in real-time (**identify)**.	Mostly automated. The trading algorithm interprets the situation by examine the metrics with a pre-determined criterion. The trader may stop trading (e.g., “panic button”) but is unable to modify the criterion in real-time (**Interpret and Evaluate**).	Automated. Orders are generated in milliseconds and are executed by the market exchange (**Define Task, Formulate Procedures, and Execute**).

2.4.2 DOA Layering on the Base DL

Similar to our way of representing DOA on the base AH, we used a DOA layering approach for the base DL. Likewise, functions allocated to the automation are shaded and functions allocated to the user are not shaded. In Figure 5 and 6, we use shaded boxes to represent information processes that are responsibilities of automation (e.g., trading algorithms). Boxes that are not shaded are human information-processing steps, assuming for simplicity that the operator is moving through all the steps of the DL. In this section, we present four cases - two scenarios (low DOA and high DOA) and two situations (routine operation and unanticipated situation).

2.4.2.1 Low DOA scenario: The routine operation situation (Case 1)

We represent two cases of the low DOA scenario (basket trading), a routine operation DL in Figure 5 and a DL showing unanticipated situations in Figure 6. We have looked at routine and unanticipated situations in order to show the challenges faced by the trader in intervening in the different automated trading scenarios. In each case, a data analysis and order generation algorithm is involved in the information acquisition and information analysis stages.

1. *Goal State*. The goal in basket trading is to hold many financial products in certain proportions. The basket of products must be bought or sold simultaneously, so that price movements for each product do not alter the portfolio allocation. The basket of products can be stated as follows:

$$\text{basket of products} = \sum_i asset\ allocation_{\ i}$$

Since the proportion of each product is normally preset and required by the trader or the trading institution, the trader begins this routine operation DL with knowledge of a desired goal state;

2. *Define Task*. Knowing the goal state, the trader then defines the task needs to be accomplished in basket trading. For basket trading, this step involves the trader setting up the criteria for short-listing products that will be traded, and deciding what trading action will be conducted;
3. *Task to Procedure Shortcut*. An expert trader may take this shortcut step to transfer knowledge of the task to certain procedure activities, without considering all details of the basket purchase allocation every time;
4. *Formulate Procedure*. To generate a complete basket, the trader must consider position limit and other administrative or trading restrictions on each product of the basket. In the futures market, for example, position limit is the highest number of futures contracts a trader may hold on the premise of deposit. In this case, the trader may fine tune the basket purchase allocation that does not violate the regulations;
5. *Procedure to Alert Shortcut*. Instead of manually carrying out the task, the trader can take this shortcut step to transfer knowledge of the task to a certain data form that will be later used by the algorithm;
6. *Activation*. The trader must download and submit historical data from the market into the automation at the beginning of each trading day;
7. *Alert*. The resulting alert will indicate to the automation that data are ready for analysis. The data are combined with the desired proportion of each financial product, and will be provided to the analyzing program;
8. *Observe*. The algorithm contains a MATLAB script (Chan, 2009) to organize (e.g., sort, rank, index and select) data according to pre-defined criteria into suitable formats. The resulting observations contain a shortlist of financial products. The resulting observations also involve orders will indicate to the automation that data are ready for analysis;

9. *Identify*. The algorithm generates a basket of orders that should lead toward a desired product allocation portfolio;
10. *Interpret*. The basket of orders must be interpreted by the trader before being submitted to the market. In financial trading, many system states are unique and could change in a short time period. At this step, the trader must undergo an interpretation of the consequences (Rasmussen, 1974);
11. *Goal State to Procedure Shortcut*. The trader decides to submit all orders. The task and all the procedures have been determined by the trader and the algorithm;
12. *Execute*. The brokerage's trading application (action implementation) submits the basket of orders to the market. The DL ends at this step.

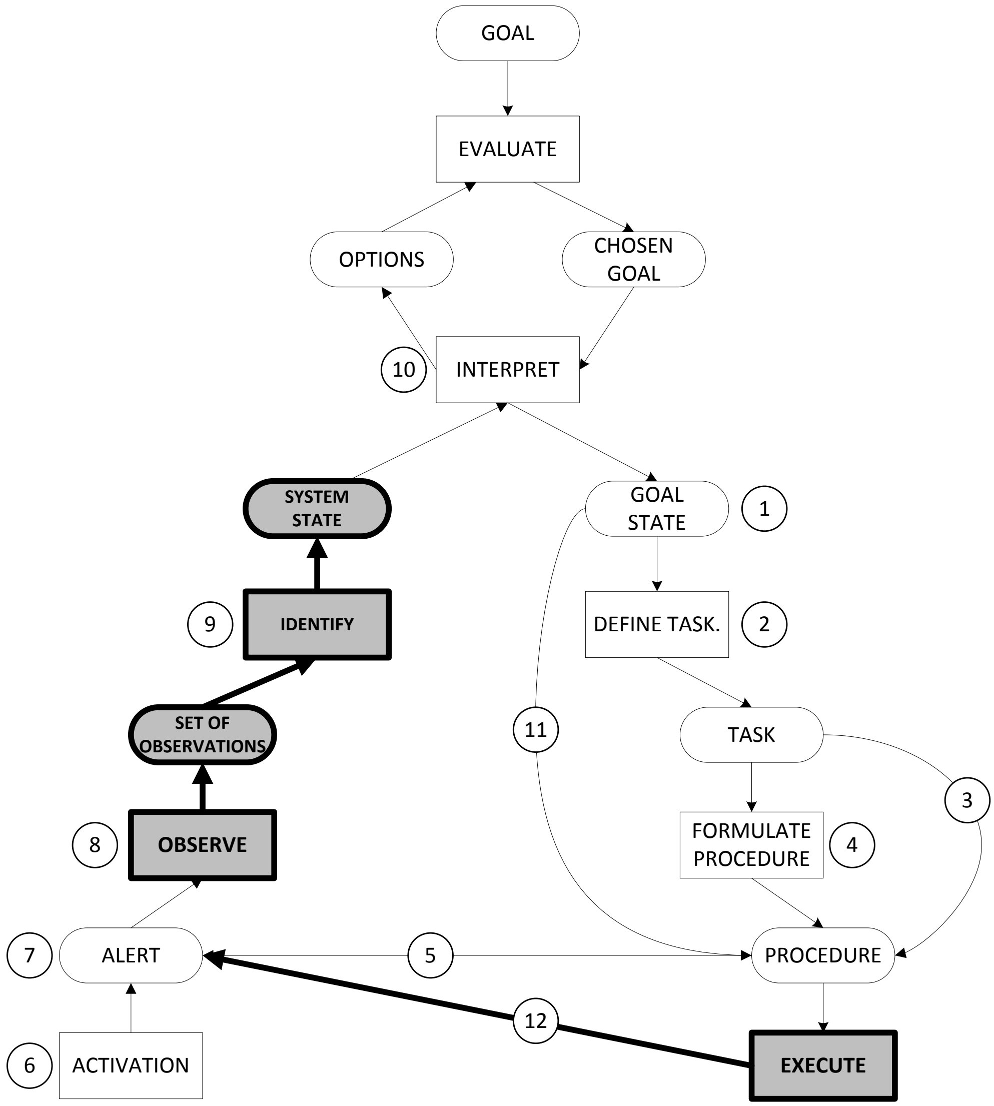


Figure 5. DL of basket trading (low DOA, routine operation).

2.4.2.2 Low DOA scenario: The unanticipated situation (Case 2)

The basket trading system can also be operated in an unanticipated mode if it is not possible to execute the basket trade on all products. As a consequence, it will be difficult to hold products in their correct proportions. Alternatively, the trader could make a wrong choice on the financial products or their proportions in the basket. A violent price fluctuation of a single product can nullify all the gains or expose the trader to losses. In this case, the basket trade cannot provide the trader protection against volatility. The unanticipated mode is represented in Figure 6. At a low DOA, the algorithm does not contain a diagnosis feature, therefore most of the decision-making is completed by the trader.

1. *Activation*. The DL starts in automated information acquisition, that is the brokerage trading software receives quotes from the market exchange via an electronic communication protocol (e.g., Financial Information eXchange – FIX). In an unanticipated situation, the resulting alerts contains quantitative data (e.g., unfilled order quantities) and the reason of order rejection (e.g., no financial product definition has been found for the purchase request);
2. *Observe*. The trader must observe the variables and compare the variables to their respective desired values to assess the fault. For example, the trader must try to probe the reason of order rejection;
3. *Identify*. The trader may identify the root cause of the fault and whether the fault is fixable. For example, if illiquidity is the principle problem and can not be mitigated in a short term, the trader may stop the attempt to rebalance the proportions of the products;
4. *Interpret, Evaluate and Re-evaluate*. The trader decides what action to take to manage the fault;
5. *System State to Goal State Shortcut*. A reoccurred situation provides knowledge that can accelerate the decision-making process;

6. *Define Task*. The trader now defines the necessary task - depending on the type of the unanticipated situation – that can move the system to the correct state;
7. *Formulate Procedure*. The trading platform takes over control from the trader. Procedures are formulated based on the decision made by the trader. The trading platform may either continue to purchase financial products in the basket, or hold on to the current portfolio;
8. *Execute*. The trading platform submits orders (e.g., buying orders or stop orders). The trader receives a confirmation message from the market exchange.

Essentially, in the unanticipated situation, with the low DOA, the trader must take over the observations and determination of system state, continuing to execute or not as required.

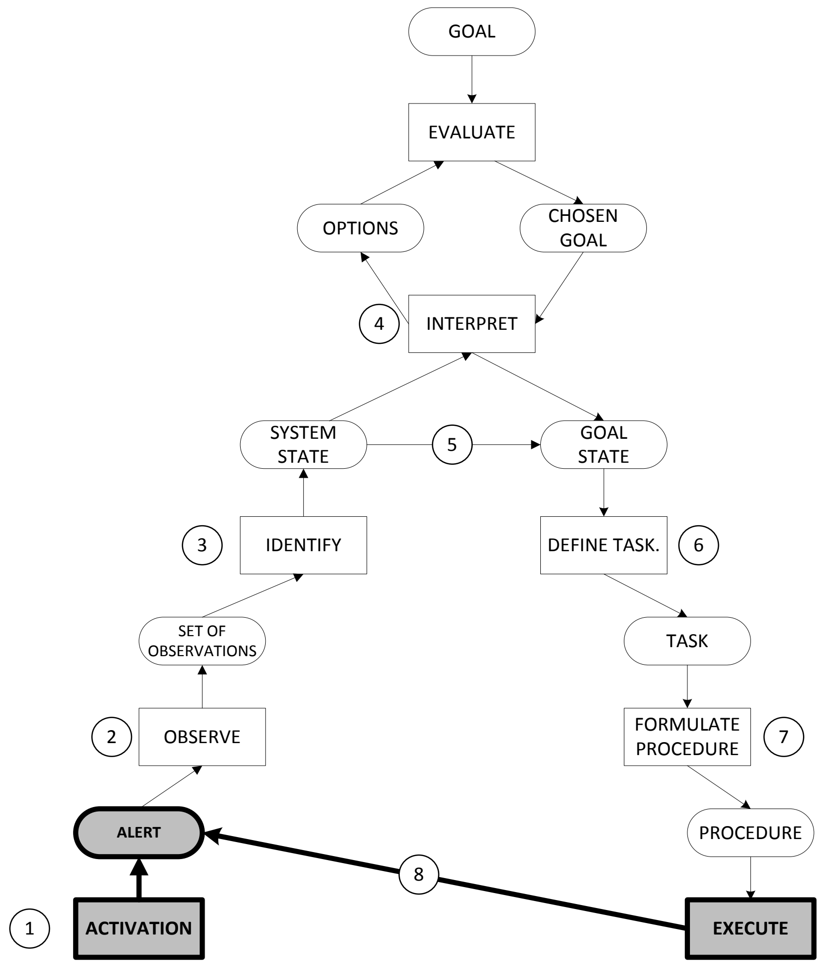


Figure 6. DL of basket trading (low DOA, unanticipated situation).

2.4.2.3 High DOA scenario: The routine operation situation (Case 3)

The high DOA scenario (trend following trading) is represented in Figure 7 (routine operation) and Figure 8 (unanticipated situation). In a routine operation mode, a scalping algorithm is first developed by an algorithm developer. At the beginning of each trading day, the trader first downloads data, strategies with other traders and clients and then starts up various applications including the scalping algorithm and the trading platform.

1. *Activation to Set of Observations shortcut*. A real-time data feed (information acquisition) receives quotes from the market;
2. *Observe*. Automation at this degree uses a pre-determined trading shortlist, therefore it does not filter the data. The trader may observe the real-time data, but this step is optional as it does not provide inputs for observations;
3. *Identify*. The quantitative trading algorithm (information analysis) calculates 50-day SMA and 25-day SMA and the RSI for preselected stocks. Note that this is an automated process, as the variables being calculated, the calculating methods and the stocks were determined before this trading task started;
4. *Interpret*. For each listed stock, when the 50-day SMA crosses above the 200-day SMA and RSI in an oversold territory is below 30, the quantitative trading algorithm (decision selection) interprets the situation as a buying signal. Automation at this DOA does not provide alternatives. It will not trade if there is no designated trading signal;
5. *Define Task*. The scalping algorithm (action automation) determines a buying task when the buying signal occurs;
6. *Formulate Procedure*. The scalping algorithm (action automation) randomizes the size of each order (400 to 800 shares), and determines other order parameters;

7. *Execute*. The scalping algorithm (action automation) places 10 iterations of orders to market then waits for a confirmation message from the market exchange.

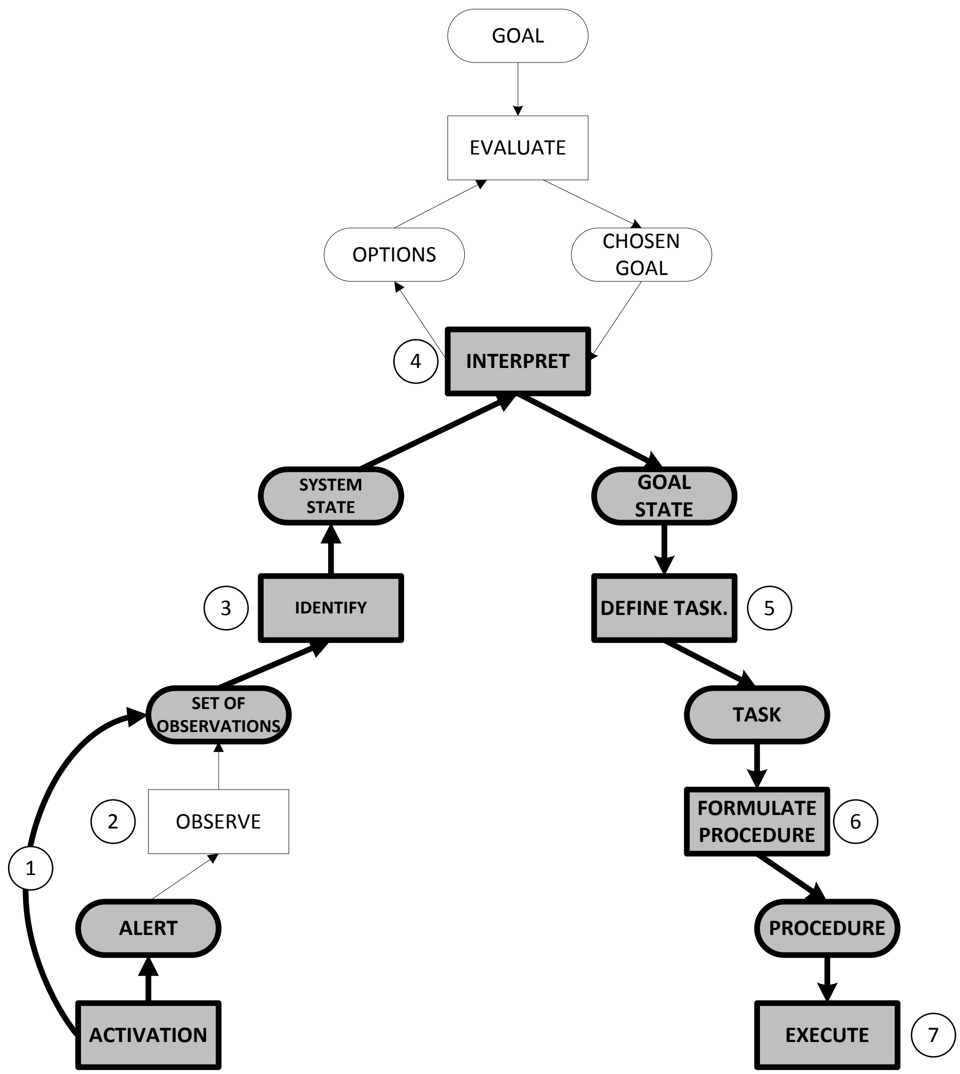


Figure 7. DL of trend following trading (high DOA, routine operation).

2.4.2.4 High DOA scenario: The unanticipated situation (Case 4)

The trend following trading system may face a disturbance and be faced with an unanticipated situation. Possible disturbances are algorithm bugs (e.g., incorrect order quantity), event risk (e.g., political event) and illiquidity. Illiquidity happens during times of low volatility when market price swings in a small range. The lack of liquidity causes a slippage, a difference between "the intended price of a trade and the price at which the trade is really executed" (Investopedia, n.d.-a). A tremendous loss of liquidity of many financial products, or systemic illiquidity, disturbs the entire market and fails most trading systems in the market (e.g., the May 6, 2010 "Flash Crash": Minotra & Burns, 2016; U.S. Commodity Futures Trading Commission & U. S. Securities & Exchange Commission, 2010). The scalping algorithm used in the trend following trading system being discussed requires a highly volatile market in order to enter and exit a trade at will in order to get a good price for the order fill. During an unanticipated situation, the trader must intervene to take a diagnosis task in order to understand the situation and try to save the system from the disturbance. The diagnosis task is presented as follows.

1. *Activation*. Just like we discussed before for the case of basket trading in an unanticipated situation, the DL begins with receiving quotes from the market exchange. The brokerage trading software summarizes market information that can contribute to the set of observations;
2. *Observe*. In an unanticipated situation, the trader must observe the collected data. The resulting observations indicate all evidence that a disturbance has happened. For example, to estimate the likelihood of a systemic liquidity risk, observations must be made on illiquidity in multiple stocks and market indexes;
3. *Identify*. The trader now identifies the current state of the trading system and confirm type and magnitude of the disturbance. The resulting system state also involves knowledge of the control law of the trading strategies;
4. *Interpret, Evaluate and Re-evaluate.* In order to generate the knowledge of goal state, it is especially important to evaluate the current system state and justify the efficacy of goal state. This stage is extremely time-consuming and may include additional data-processing activities on top of the DL. For example, when a market crash is observed, traders must be

very careful in setting up new trading strategies. The trader must decide whether an algorithm-placed order has been processed by the market exchange;

5. *System State to Goal State shortcut*. Another constraint to generate a goal state is timing. Using the same example of a market crash, sophisticated trading strategies may provide robust market disturbance tolerance (e.g., a certain market condition and trading status can trigger a precaution trading execution), and some of the extreme market conditions can be considered in algorithm design. In this case, there is a state knowledge transfer from the current system state to the goal state. The transfer is illustrated as a leap from system state to goal state;
6. *Define Task*. Because the trading system has a high DOA, there are limited options to recover the system from hazardous conditions to compensate for the disturbance. The trader may reconfigure the current trading system to compensate for the disturbance (e.g., modifying an offset setting of the SMA crossover rule). In the case of irretrievable disturbance, a stop loss task is decided by the trader;
7. *Formulate Procedure*. The scalping algorithm (action implementation automation) determines order parameters;
8. *Execute*. The brokerage trading software (action implementation automation) places new orders to the market and waits for a confirmation message from the market exchange. For example, the scalping algorithm submits a stop order to the market exchange. A confirmation message is then received from the market exchange.

It becomes apparent that in the two high DOA cases, routine operation and unanticipated situation, the trader must interrupt the automation and assume a larger scope of control. Further, because the automation is likely handling small fluctuations well, the problem at hand is likely more complex than usual, for example, a market liquidity change as discussed. Compared to the low DOA cases, the opportunities to recover are more limited as more information-processing steps are allocated to the automation. The automation can result in rapid executions that can be challenging to interrupt.

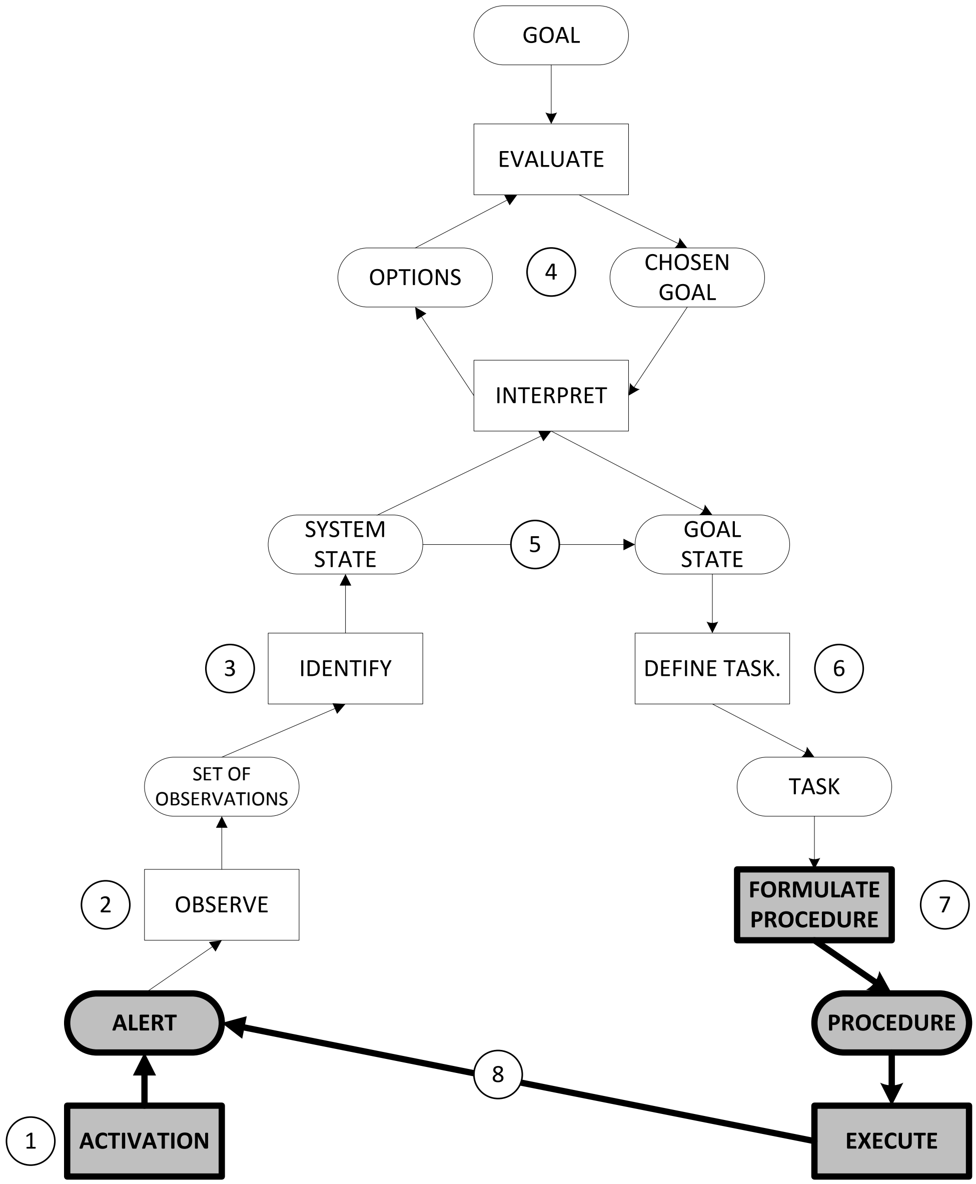


Figure 8. DL of trend following trading (high DOA, unanticipated situation).

2.5 Discussions

In this modeling exercise, we proposed a DOA layering approach for conducting two analyses, first the WDA then the ConTA. For each analysis, we first built a base model (i.e., AH or DL), then mapped function allocation in an additional layer. The base model is similar in many respects to that of an original AH or DL consistent at any DOAs. Functions in the base model can be allocated to any actor (human or automation) and represented in the DOA layer. Shared allocations could also be included in the DOA layer, though the analyst may want to differentiate between shared allocation approaches. The DOA layer adds a new dimension to the base model, showing how human and automation work collaboratively at a certain DOA. The DOA layering approach can be used for representing function allocations at the domain-level and the task-level. First, the presented base AH example shows that the physical and the functional structures are consistent in the basket trading system and the trend following trading system. The DOA layer suggests function allocations are different in the two trading systems, depending on the system's DOA. In the DL examples, we analyzed two tasks - routine operation and unanticipated situation - of the basket trading system and the trend following trading system. Second, the base DL as a template (Vicente, 1999) was augmented with four regions to show the four stages of automation. The DOA layer enriches the base DL with more features, such as automated information-processing steps, stages of knowledge and shortcuts.

The following discussions are focused on a comparison of the DOA layering approach and how automation was previously modeled with CWA, and what more design implications the DOA layering approach might have for designing support for automation than existing approaches. Lastly, we discuss implications of applying the DOA layering approach to adaptive automation.

2.5.1 Comparing DOA Layering Approach to Dual-Model Approach

In this section, we compare the DOA layering approach to how automation was represented in the CWA literature. In most cases, CWA literature explicitly focusing on modeling automation has taken a *dual-model* approach, with one model showing non-automated systems and a separate model showing automated components. We first introduce the dual-model approach and its origin, then conduct a comparative analysis of the DOA layering approach and the dual-model approach. The objective of this analysis is not to draw a conclusion on which approach is superior to the other. Instead, we suggest that either approach has its own application depending on the type of system and problem being modeled. The applicable occasions of the two approaches are discussed, suggesting when to develop the DOA layered model and when to develop the dual-model.

2.5.1.1 Dual-model approach

Typically, automation and function allocation requirements are explained in the Social and Organizational Analysis of CWA, after the WDA and the ConTA are completed (Vicente, 1999). The dual-model is a relatively new approach formally introduced by Mazaeva and Bisantz (2007) using a digital single-lens reflex camera analysis study. Mazaeva and Bisantz (2007) suggested that automation should be explicitly modeled at the WDA and the ConTA phases, using AH and DL tools. We found that the two aspects of the dual-model approach, the dual-model AH, and the dual-model DL, have somewhat different origins.

1. *Dual-model AH*. The original AH proposed by Rasmussen (1986) modeled automation as work domain components at the lower AH levels, which are Physical Function and Physical Form. It seems to be an appropriate modeling decision in various examples of automated systems given by Rasmussen. These examples include a washing machine (pump and valve function, configuration and weight, and size), manufacturing plant (physical functioning of equipment

and machinery) and computer system (electrical function of circuitry). Burns, Bisantz and Roth (2004) suggested an extension to the original AH to represent automation with WDA. Burns et. al comparied decisions made on modeling naval sensor systems. They proposed that developing dual AH models in which one AH represents the system being controlled and another AH represents the automation, may "help specify information needs for those responsible for automation monitoring and control" (Burns et al., 2004). Inspired by Burns et. al's work, Mazaeva and Bisantz (2007) provided a detailed example of dual-model AH, introducing a "camera AH" to show non-automated camera components that are being controlled, and a separate "automation AH" to show automated systems allow for automated movement and exposure control of the camera's components. Their dual models also represent interconnections between each level of the two AHs;

2. *Dual-Model DL*. To our knowledge, the first introduction of the dual/multiple model approach is Rasmussen and Goodstein's work (1987). It was pioneering work, using three DLs to represent cooperative decision making in a nuclear reactor control domain among different actors, including a human designer, an operator, and a computer. Each DL reveals a different sub-task and together they complete a control task (e.g., operator intervention during an accident). Similarly, Mazaeva and Bisantz (2007) developed three DLs, representing interrelated control tasks distributed across the automation, the photographer and the designer represented.

2.5.1.2 Occasions where the dual-model approach fits

Table 2 briefly summarizes the differences between the dual-model approach and the DOA layering approach.

Table 2. A Comparison of the Dual-Model Approach and the DOA Layering Approach.

	Dual-model approach	DOA layering approach
Basic concepts	Allocate user and automated system functions to separate AHs;	Allocate user and automated system functions to separate layers in the same AH;
	Allocate user and automated procedures to separate DLs.	Allocate user and automated procedures to separate layers in the same DL.
Deliverables	User model (AH and DL);	Base model (AH and DL);
	Automation model (AH and DL).	User layer (AH and DL);
		Automation layer (AH and DL).

The dual-model approach is a successful first attempt to explicitly represent automation within a CWA model. To understand the applicability of the dual-model approach, readers must note that Rasmussen and Goodstein's approach (1987), the origin of the dual-model approach, was initially proposed for supporting supervisory control system design. Supervisory control is associated with an intermediate DOA (Sheridan & Verplank, 1978). The system boundary is clear for analyzing a supervisory control system, whereby the automation takes a task performer's role in the closed inner loop, while the operator manipulates control parameters in an outer loop (Sheridan, 2011). Similarly, Mazaeva and Bisantz's AHs and DLs (2007) are exclusive representations of decision process allocation within the ongoing supervisory control. They looked at a digital single-lens reflex camera, a commercial product whose DOA has already been decided. In other words, the dual-model AHs and DLs are constrained by a certain DOA.

From Rasmussen's example (1986), and Mazaeva and Bisantz's example (2007), we can see that the dual-model approach is an appropriate approach to model with a fixed DOA, and to analyze and understand existing automated systems. This finding echoes with Burns et. al's work (2004). They pointed out that the automation AH is "perhaps created later in the analysis, once the levels of automation have been specified". In this case, guiding automation design is not the primary objectives of the analysis. Instead, analysts may focus on addressing multiple control tasks and strategies to represent sophisticated interactions between human and automated behaviours, at a predetermined DOA. Since the dual-model approach develops a more explicit model of the workings of that automation, the dual-model approach is a good choice where operators must diagnose or fix the automation itself.

2.5.1.3 Occasions where the DOA layering approach fits

There are certain occasions where the DOA layering approach fit better than the dual-model approach.

1. *Modeling systems with a variable DOA.* Mazaeva and Bisantz (2007) modeled a fixed DOA system, the dual models lack flexibility of analyzing domains of a variable DOA. Analysts using the dual-model approach may encounter a scalability issue while applying the dual-model approach to guide automation design in significantly more complex, personal and property safety systems (e.g., aviation, process control, finance) than a camera system. On the other hand, the automated trading system we analyzed, is an example of intelligent automated systems with a consistent physical structure and a variable DOA. The proposed DOA layering approach is a single-model approach. In the presented AH examples, a consistent physical structure is shared between the human and the automation. Likewise, in the presented DL examples, the same control task is shared between the human and the automation. The DOA layered models have simplicity in how automation is modeled. Our approach may prove useful in representing more coordinated human-automation interaction, by leaving more flexibility in modeling multiple system modes (e.g., routine operation and unanticipated situations);
2. *Incorporating stages and levels of automation model*. CWA was built for understanding complex automated systems. Automated is an important aspect of socio-technical systems

(Vicente, 1999). A considerable interest has grown up in the human factors community around the theme of how to model the DOA (e.g., Sheridan & Verplank, 1978; Parasuraman et al., 2000). Among the previous examples discussing how to explicitly model automation in CWA, the stages and levels of automation model has not been well utilized in the CWA literature. The DOA layering approach takes the first step to fill this gap, by transforming DOA knowledge to CWA models. We found that ConTA could play an important role in describing how information-processing activities are allocated to the user and the automation;

3. *Supporting automation design*. The DOA layer, layering on the base DL, is a supplement to the stages and levels of automation model. The DOA layering approach supplements the "broad-brush" description of levels of automation (Pritchett, Kim, & Feigh, 2014), by enabling two important features within DL, which are the ability to show shortcuts between information-processing steps, either 1) leaps and shunts an expert takes in the case of human information-processing, or 2) bypassing a non-automated step in the case of automated information-processing. The former feature is inherent from the original DL and is still available to analysts using the DOA-layered approach. The latter feature describes alternative routes of information-processing, implying opportunities of human operators.

2.5.2 Implications for Design

The DOA layering approach makes unique contributions to automation design, both designing automation displays and deciding stages and levels of automation. We discussed two possible design implications, designing ecological automation displays with DOA and constraint-based procedure supports, and deciding stages and levels of automation. We used the presented AHs and DLs as examples.

2.5.2.1 Implication for display design: Designing ecological automation displays

The Ecological Interface Design approach requires user interface designers to first conduct an information analysis to extract information from a completed AH (Burns & Hajdukiewicz, 2004). Next, this information should be organized as a list of variables for representational design, with constraints from the work domain.

The DOA layering approach allows user interface designers to capture variables from the base AH, and constraints from both the base AH and the DOA layer. For example, in the trend following trading AH, a functional purpose of the system is to achieve a maximum rate of revenue in trading. This function could be described by revenue run rate, a metric for predicting future financial performance based on the current financial information. The constraint of this metric is decided by the technical limitation of the trading system and can be found in the DOA layer. Allocating this functional purpose to automation means the trading system is running in real-time in a day trading setting. Therefore, a short duration (milliseconds or seconds) of this metric must be calculated and monitored by the automation, as the trader is incapable of monitoring the rate of revenue in an extremely short duration. On the other hand, at the Physical Form level, both variable cost and fixed cost functions are allocated to the trader. According to the base AH, the two types of costs are constrained by a certain currency type of the trading market. Other constraints are related to the trader only, not the automation. They are trader specific information such as the trader's personal financial status, indicating that the trader is ultimately responsible for cost control in a trend following trading system.

More variables and constraints can be seen from the means-ends relationships on the base model, as well as the DOA layer overlaid. For example, "position" at the Physical Function level of the base AH connect to "market price", "order price" and "position price" at the Physical Form level, suggesting market and portfolio are two inter-related sides, and an integrated market-portfolio display may support direct perception of information from both sides. With a high DOA layer, such a relationship remains consistent but “position” is allocated to the trader and the price-related functions are allocated to the automation. This allocation suggests that, in a trend following system, although these price-related functions are represented on the display with appropriated constraints, the trader does not take control of these functions. Therefore, additional visualizations may be provided to the trader to understand how automation processes these functions.

2.5.2.2 Implication for automation design: Determining automation stages and levels

Another implication of the DOA layering approach is this approach could fit into the framework for automation design proposed by Parasuraman et al. (2000) to help determining automation stages and levels. The stages of automation model, an important “starting point for considering what types and levels of automation should be implemented in a particular system” (Parasuraman et al., 2000), provides “a simple guide for automation design”. The framework suggested that automation design should begin with identifying what class of functions should be automated. The automation designers then apply evaluative criteria (e.g., automation reliability and situation awareness) and recommend “particular levels of automation for each of the four types of automation”.

We believe that fitting DOA layering approach to an existing automation design framework could supplement the stages and levels of automation model, rather than replacing this model. We suggest that automation designers may use the DOA layering approach at an early phase of automation design, before applying evaluative criteria, to help automation designers determine what stages and levels of automation are appropriate for the system. The base DL represents the four functional domains on a DL, providing an easy start point for automation designers to develop a conceptual design estimation. We hope trading algorithm developers may consult with the base DL in the future, to decide which algorithm to use, an intelligent algorithm (i.e., decision automation) or an order-placing script (i.e., action automation). Automation designers must also justify the use of a certain stage of automation. At a subsequent stage of automation design, automation designers must decide what level of automation should be developed within each functional domain (Parasuraman et al., 2000). The DOA-layered DL provides richer information than the stages and levels of automation taxonomy. The DL shows not only what human or automation functions should be applied within each stage (shades), but also what aspects of human interactions with automated systems should be considered (shortcuts).

The DOA layering approach could potentially help with understanding and design for modern, intelligent automation. The DOA layering approach echoes a recent suggestion by Sheridan (2017), suggesting that modern automation is hierarchical in the same way as the human work competencies. If modern automation is hierarchical, then automation competencies can be modeled by the Skill-Rule-Knowledge (SRK) taxonomy, the last phase of CWA that has only been used to model human work competencies in the literature. Sheridan gave hints for identifying an SRK for automation: 1) skills of the automation are continuous actions triggered by the laws of physics (e.g., the spinning of steam turbine) but are conditioned through commands of an automated (e.g., a programmable logic controller) or human (e.g., an operator) agent; 2) supervisory control and artificial intelligence go beyond the traditional realm of classic feedback control, and invoke the “rule” or “knowledge” level on the hierarchy of SRK. An “if *x,* then *y,* else *z*” logic forms a stored rule to invoke designated human (e.g., an action recommendation system) or automation (e.g., action automation) activities; 3) automation using the knowledge level is rare, but becomes possible in machine learning based systems such as the IBM Watson computer.

The implication of the SRK for automation is, automation may use all stages and levels of information processing. By modeling the DOAs on the AH and DL, it can clearly be seen that in the higher DOA situation, functions at the higher AH levels (e.g., Functional Purpose and Abstract Function, in the high DOA AH) and information-processing steps on top of the DL (e.g., Interpret, in the high DOA, routine operation DL) are allocated to the automation which may present knowledge-based automation (c.f., Rasmussen, 1986). The interconnections of the presented AH and DL examples and SRK for automation suggests a future extension of DOA layering, layering function allocations in other phases of CWA to support automation design.

2.5.3 Implications for Modeling Adaptive Automation

The DOA layer on the DL may help the analyst model DOA shifts, shedding some light on how to model adaptive automation in future. For example, Table 3 presents two cases of DOA shifts, a DOA-increase case and a DOA-decrease case, based on the high DOA scenario. It can be seen from the table that a DOA shift can occur at any DL step, as any DL step (box) can be shaded (i.e., functions reallocated to automation) or not shaded (i.e., functions reallocated to the human). On the other hand, DOA shifts can be frequent, as algorithm development is an extremely flexible process depending on traders' expertise and preference. It is also an iterative process, with each iteration starts from developing, back-testing to live-trading. At this stage, the DOA layering approach portrays the relationship between human and automation functions at a task level, and we hope it grows into a potentially useful approach for modeling adaptive automation.

Table 3. Example Reasons for DOA Shift per DL Step in Trend Following Trading.

DL processing steps (in financial trading terms)	Reasons for DOA shift	
	DOA increases	DOA decreases
Activation by normality and abnormality in maket or portfolio	Automated signal detection is capable (e.g. timely tick data in shorter duration; Level II data); impulse control	Technology is unavailable due to high costs or lack of work competence; distrust in technology (e.g., concerns with latency of the data); obsessive financial market monitoring
Observe the dimensions of the issue	High computing power is available for real-time pattern generation	Countervailing trading philosophy (e.g., fundamental analysis is favored over real-time technical analysis)
Identify the current state	High computing power is available for real-time pattern recognition; system state can be quantitatively modeled	Concerns with latency in pattern recognition (e.g., unavoidable delay in automated executing)
Interpret the ambiguity of historic and current states, as well as the consequences of future states; **evaluate** the current state with a goal from a higher level of abstraction	Artificial intelligence advances; no or little ambiguity in the current status the current market condition is predicted; historic market data is accessible and understandable by the trading algorithm; prediction model is reliable	Automation is not capable to interpret or is believed misinterpreted the current status; market condition is abnormal; the current status is interpretable, but the consequences of future states are not acceptable (e.g., risk of spoofing, see N.D. Ill. v. Sarao, 2015)

DL processing steps (in financial trading terms)	Reasons for DOA shift	
	DOA increases	DOA decreases
Define task in financial trading (e.g. buy or sell)	Indispensable in some high-frequency trading systems (Chan, 2009); adequate knowledge of automated trading and high-performance programming	Complexity and cost are not acceptable
Formulate procedures, in another word, generate orders with appropriate arguments (e.g. order quantity, order price and target financial product)	Indispensable in some high-frequency trading systems	
Execute an order in the market	Indispensable in some high-frequency trading systems	Lack of knowledge in high-performance programming, but Semi-automated alternative

2.6 Conclusion

Information systems should support human-automation coordination (e.g., either human or automation must seamlessly switch between responsibilities). CWA helps the development of "simple qualitative models" (Sheridan, 2017) that can be represented by graphical interfaces. An adoption of function allocation models, such as the stages and levels of automation model to CWA could provide a new design opportunity. Yet, this approach has been not well developed. We attempted to fill this gap by proposing a DOA layering approach, layering DOA on AH and DL to express domain- and task-level function allocation respectively. This paper is an extension to two earlier versions in the Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Li et al., 2015, 2016).

Automated trading, a domain rarely explored by the human factors community served as an appropriate example in this modeling exercise. Automation in financial trading is versatile in terms of the various stages and levels of automation involved, the highly-coupled relations between traders, infrastructure and trading algorithms and the unpredictable dynamics in the environment. Two scenarios of financial trading are provided in this paper, each has a unique DOA. New models in the context of automated trading were developed, using extended AH and DL with DOA layers. In each case, a base model was first created, followed by mapping two scenarios (i.e., low DOA, and high DOA) onto the base model. In particular, we correlated the stages and levels of automation model (Parasuraman et al., 2000) to the DL, whereby DL steps were organized into four stages. This paper is the first to propose a DOA layering approach, and the first to comprehensively use CWA and the stages and levels of automation model to model automated trading.

This paper provides useful insights to the debate of using a single-model approach or a dual-model approach to model automated systems. The DOA layering approach extended the flexibility of the single-model approach by representing the DOA and echoes Sheridan's recent homage (2017) to Rasmussen's frameworks (e.g., AH) for their robustness and applicability to behaviors of humans or highly intelligent automation. Future works include how to model adaptive automation using the template-layering approach. This paper also corroborates Borst's recent suggestions (2015) on providing more automation status on ecological displays to support human-automation coordination. We will further examine the design implications in an experimental study.

2.7 Acknowledgements

This project is financially supported by a Natural Sciences and Engineering Research Council of Canada grant (CRDPJ 445965-12). Special thanks to Quantica Trading Inc. for providing the venue for the participant observation. We thank Drew Coles, Roni Hoffman, Travis Felker (Quantica Trading Inc.), Dr. Rui Hu (Microsoft Corporation) and Xian Wang (Shenzhen Platinum Venture Capital Ltd.) for serving as the subject-matter experts. We also acknowledge the comments of our associate editor Dr. Emilie M. Roth and three anonymous reviewers.

2.8 Chapter Summary and Connections to Research Questions

2.8.1 Key Findings

DOA layering approach: This chapter proposed a DOA layering approach to extend the CWA. The base AH and the DL are identical to how the AH and the DL has been typically used in the CWA literature and shows the constraints and relationships for the functional working of the system without any automation applied to it. The DOA layer, as its name suggests, reflects additional constraints related to the system's DOA. The stages and levels of automation model has been adopted in the DL. The applicability of this approach has been demonstrated using automated trading examples. The examples as well as the resulting models have demonstrated how the DOA layering approach can be applied to reveal the connections between the human and the automation. The possibly most important deliverables of the models were the gaps between human and automated steps on the same information-processing template.

Design implications: This chapter demonstrated that the DOA layering approach has implications for designing ecological displays and automation in general. These implications are useful in guiding the design of an experimental study for evaluating the effectiveness of ecological displays.

Modeling adaptive automation: This chapter suggested future extensions to the DOA layering approach. The current form of the DOA layering approach described DOA shifts using a DL template and has enabled preliminary modeling of adaptive automation with the CWA. Future development in this area is warranted.

2.8.2 Connections to Research Questions

The author has three research questions for this dissertation:

Research question 1: How can we model automated trading systems with a variable DOA using CWA?

Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

This chapter examined research questions 1 in great detail. With the complexity relevant to the variable DOA in the automated trading domains explicitly modeled using the DOA layering approach, this research objective has been achieved. Further, the DOA layering approach has strong potentials in supporting automation and design, which will be demonstrated in Part C of this dissertation.

Part C

Design and Evaluation

Part C of this dissertation is formed by two chapters that give examples of designing ecological displays based on DOA-layered models as well as evaluating these displays in a financial trading simulation.

Chapter 3 presents Experiment 1 to evaluate a trend following trading system that used four scenario types (combinations of two DOA configurations and two display types). The four scenarios were moderate-conventional, moderate-ecological, high-conventional and high-ecological. Guided by the DOA-layered models presented in Part B, automation is designed with two DOA configurations that involve the identical acquisition and analysis stages but the different decision and action stages. Conventional displays that are typical in financial trading software are implemented to facilitate the basic control of the two configurations. Ecological displays are developed to add additional support to the conventional displays that are appropriate to the specific DOA. The ecological displays are derived from the base CWA models (previously presented in Part B). In this experiment, a series of hypotheses are proposed based on the research questions of this dissertation and are examined with a variety of measures in the categories of task performance, SA, eye-tracking, workflow and risk preference. Results of this experiment are reported, and connections to the research questions are discussed.

Chapter 4 presents Experiment 2 to evaluate the trend following system with four other scenario types – combinations of two DOA configurations and two display types. The first DOA configuration used automation in all information-processing stages and the second DOA configuration was an adaptive automation condition. The first DOA configuration is similar to one described in Experiment 1 with a disconnection between the earlier stages (i.e., acquisition and analysis) and the later stages (i.e., decision and action) of automation eliminated to improve ecological display support in this configuration. The second DOA configuration demonstrated a preliminary design exercise of adaptive automation in the financial trading domain, with automation being allocated dynamically during task performance. The design of Experiment 1 and 2 are identical in other aspects, sharing similar apparatus and using the same evaluation methods. Results of the evaluation are presented later in this chapter.

Chapter 3

Experiment 1: Trend Following Trading

3.1 Foreword

Starting in this chapter, AUTRASS is introduced as a simulator for evaluating concepts for designing automation as well as ecological displays. This chapter describes Experiment 1 to understand how human participants interact with the AUTRASS in a trend following trading setting which has been previously modeled in Part B of this dissertation.

First, as part of the apparatus of Experiment 1, reasons to build the simulation on the trend following trading scenario are explained. After that, the development of the AUTRASS simulator is introduced and the distinct function allocations between the trader and the automation and the design of conventional displays are described. Later, the design of two ecological displays which might provide additional support for using the two DOA configurations is explained. In Experiment 1, four scenario types denoting all combinations of the two DOA configurations and the two display types were examined. The participants performed different tasks in the two DOA configurations. The simulated market dynamics demonstrated unanticipated situations that were unfamiliar to the participants. As discussed in the literature, it is important to note that these unanticipated situations were likely where the ecological displays outperformed the conventional displays (Lau, Jamieson, et al., 2008). Later in this chapter, independent variables and the dependent variables relevant to Experiment 1 are introduced, and a series of hypotheses about the research questions of this dissertation are proposed. Lastly, experiment results and connections of these results to the research questions are discussed.

3.2 Trend Following Trading: A Revisit

For design and evaluation purposes, the trend following trading scenario ("high DOA") that has been previously discussed in the modeling stage of this dissertation was re-adopted. This particular scenario was chosen over the basket trading scenario ("low DOA") due to the following reasons:

First, the trend following trading scenario can be studied in more depth as an example of an automated system with a variable DOA. Trend following trading is a commonly applied strategy in financial trading in general and, as demonstrated in the DOA-layered models, may use all stages and levels of automation. In the presented models, the automation generates trading signals based on the calculations of technical analysis indicators of market prices and executes the trades autonomously. However, trend following trading does not always use an upper-intermediate to high DOA. In other cases not demonstrated in the presented models, the use of automation in trend following trading can be limited to just calculating the technical analysis indicators and providing the results to the trader, in which case, the trader must decide on what action to take and implement the action. Indeed, Murphy (1999) suggested that automation (or "computer" in his original words) can even be largely excluded from trend following trading. Murphy stated that "much of the work involved in technical analysis can be performed without the computer. Certain functions can be more easily performed with a simple chart and ruler than with a computer printout". Thus, it became necessary and feasible to implement and evaluate a lower DOA trend following trading system;

Second, it is more practical to develop monitoring and fault detection tasks for human participants to evaluate different DOA and display types with trend following trading. Typically, human factors researchers evaluate design concepts in a scenario-based experiment (DURESS: Vicente & Rasmussen, 1992). In trend following trading, the trader and the automation must respond instantly to both opportunities and anomalies in trading. The fault detection task (as shown in the unanticipated DL) are performed in parallel to the monitoring task (as shown in the routine operation DL) in a simulation of trend following trading and can be integrated to a scenario-based experiment. The trend following trading environment can be used to evaluate existing measures in the literature (e.g., task performance, workload, and SA) and to develop new measures relevant to the automated trading domain (e.g., traders' risk preference). The basket trading scenario, however, involves longer task phases than those in trend following trading. The longer task phases may not be completed in a short timeframe and are more expensive to simulate in a lab-control experiment;

Third, it is easier to recruit and train participants for an experiment that examines trend following trading in comparison to basket trading. As shown in the DOA-layered models, legally achieving a profitable revenue is the major mission objective of trend following trading. This mission objective was expected to be straightforward to novice participants who did not have in-depth knowledge of financial trading. Technical analysis and trading algorithm would be new concepts to most novice participants; however, with appropriate on-site training before performing designated tasks, these concepts should be understandable by those who have basic knowledge of mathematics and computer programming, which represent the typical student population of the University of Waterloo. On the other hand, evaluating basket trading requires the participants to have formal knowledge that might only be obtained through professional training. Indeed, to successfully generate a complete basket of financial products or just to understand the importance of portfolio management, the participants must have adequate knowledge of market fundamentals, portfolio and risk management, and laws and regulations. In this case, the selection of participants was extremely limited, and those who are experts in this domain (e.g., institutional traders) would be qualified.

3.3 Apparatus

3.3.1 AUTRASS: The Simulator

The author of this disseration led and was actively involved in a student development team to develop AUTRASS under the guidance of the subject-matter experts. AUTRASS was programmed in Microsoft Visual Studio 2013 development environment using C# with the Microsoft Dynamic Data Display framework (CodePlex, 2011) under an open-source Microsoft Reciprocal License (Ms-RL). In total, there were approximately 2,200 lines of code, and it took approximately 12 person-months to implement. The AUTRASS was deployed onto a 3.8Ghz quad-core desktop computer with 12 gigabytes of memory and a 27” liquid-crystal display. The display used a resolution of 1920 × 1080 pixels.

For this experiment, AUTRASS was fine tuned to provide playback of historical market data of a financial product named SPY at a real pace. SPY stands for the Standard & Poor's depositary receipts, a popular exchange-traded fund designed to track the S&P 500 stock market index. The update interval of the playback was 5 seconds and such corresponded to the interval of the market data.

A module front-end of AUTRASS was designed to allow different ways for submitting orders to the back-end that would be appropriate with various DOAs. In section 3.3.3 of this chapter, automation design will be discussed in detail. Different displays were designed and implemented on the AUTRASS front-end for evaluation purposes, and the development will be described in detail in section 3.3.5 and 3.3.6. In the back-end, AUTRASS simulated a market exchange to process orders submitted from the front-end by the participants or the automation.

It is important to note that participants who attended the experiment were explicitly told that the AUTRASS back-end would only process one trade per 5 seconds to simulate a *latency* (i.e., the time delay in the telecommunication and trade processing), which is typical in financial trading in the real world. Only one order (i.e., buying or selling) can be submitted from the front-end at any time stamp (let it be *t*). This order would be silently processed by the back-end at the current timestamp (*t*), and the results of the trade would be executed and presented on the display when the next timestamp arrived (*t* + 1). Note that AUTRASS backend did not allow for short selling (i.e., profiting by selling shares borrowed from a brokerage in a falling market). Although brokerage was captured in the CWA models, it was not simulated on AUTRASS for simplicity. All simulation data were silently recorded in log files by the AUTRASS back-end. A Gazepoint GP3 eye tracker (Figure 9) was attached to the bottom of the computer monitor. The eye tracker used infrared cameras to identify participants' scan pattern at a 60 Hz sampling rate. The eye tracker has .5° to 1° of accuracy. The participants attending this experiment were told that they can naturally move their head and eyes. Raw eye-tracking data and the simulation screen were jointly captured by a software tool provided by Gazepoint.

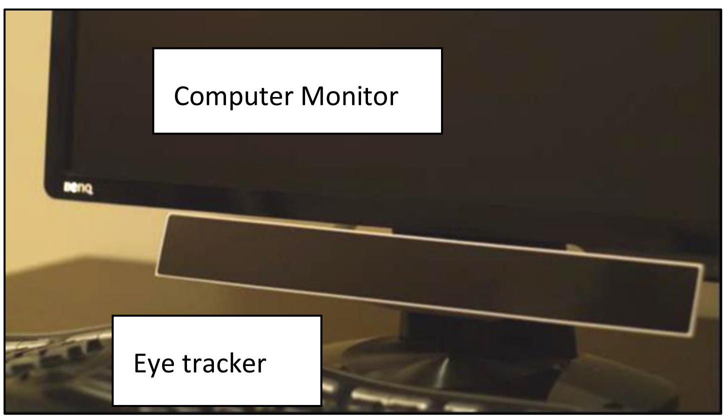


Figure 9. Eye tracker set-up.

3.3.2 Unanticipated Situations

The latency in executing trading orders simulated by AUTRASS created a *slippage* between the intended market price of a trade at which the decision would be made and the market price at which the execution would be completed. Slippage, as well as the market dynamics naturally created unanticipated situations in financial trading that would be important for evaluating the effectiveness of ecological displays. The ecological displays provide a functional representation of the physical structure of the work domain to support knowledge-based problem-solving. According to the literature, ecological displays were particularly useful in supporting problem-solving in unanticipated situations where procedural support is not always available (Lau, Jamieson, et al., 2008).

This treatment of unanticipated situations was different from how ecological displays have previously been studied in the literature. The EID literature studied both anticipated and unanticipated situations in scenario design (e.g., Burns et al., 2008). Typically, each scenario was designed to have different phases in their order of occurrence in time (e.g., the detection and mitigation phases, in Burns et al., 2008; the climb, cruise, and descent phases, in Ellerbroek, Brantegem, van Paassen, de Gelder, & Mulder, 2013). In this dissertation, unanticipated situations were presented to the participants across the entire scenario, because the slippages and the market dynamics occurred all the time.

3.3.3 Automation Design

Two DOA configurations - moderate DOA and high DOA - were designed to demonstrate how DOA layering approach (cf. section 2.5.2) can be used to imply automation design. Technical capabilities and limitations of the AUTRASS were also considered. The two DOA configurations, based on the CWA models, described distinct function allocations between the trader and the automation under similar requirements of trend following trading. Each participant attending Experiment 1 took the role of the trader that was considerably different within the two DOA configurations. Table 4 shows the multiple task phases that correspond to the Parasuraman et al. stages and levels of automation model. This particular model was selected over the other automation models for demonstrating the connections of the design to the CWA models previously presented. Both DOA configurations involve the first two stages of automation (i.e., acquisition automation and analysis automation), relating to financial trading. With the high DOA, automation has been allocated to more authorities regarding financial decision-making and execution.

The following sections explain the design considerations for each DOA configuration, with reference to the trend following trading CWA models.

Table 4. Function Allocation Between the Participant and the Automation across the Moderate DOA and the High DOA (Experiment 1).

		DOA			
		Moderate [1]		High	
Stages of Automation	Task Phases	Trader	Automation	Trader	Automation
Information acquisition	Collect and observe information from the market quotes and display it on a display (in favour of the traders)		×		×
Information analysis	Calculate a short-period SMA and a long-period SMA and plot the curves onto a display (in favour of the traders)		×		×
Decision selection	Interpret the current and predict situations to decide signals to buy and sell	×			×
Action implementation	Determine and perform a buying or a selling task	×			×

[1] Its DOA was lower but it was also derived from the trend following trading scenario described in Part B. It is however different from the basket trading scenario. The word “moderate” was used to avoid confusions.

3.3.3.1 Information acquisition

According to the DL models, acquisition automation in trend following trading involves using a real-time data feed to receive quotes from the market, storing the data in appropriate data structures, performing basic data cleaning and filtering and, in the case of moderate DOA, graphically representing the data on a display (e.g., a bar chart or a candlestick chart). For the high DOA, since later stages of automation were also being used, visualizing the market data was not necessary as the automation directly accessed the data structures. For consistency in the experimental design, this visualization has been used in both DOA configurations. The information automation in a real-world setting must also include a hardware back-end to ensure the telecommunication to the market exchange is stable. In this experiment, since the trading software and the market were being simulated on the same computer, functions related to the hardware were omitted for simplicity.

3.3.3.2 Information analysis

Analysis automation identifies a series of system states about the financial market that can be formed into a technical analysis tool for the decision maker (either the trader or the automation) to pursue the later task stages. Automation must calculate technical analysis indicators in real-time and, if necessary, presents the results in an appropriate visual form for the trader to utilize these results. For consistency in experimental design, the technical analysis indicators have been illustrated as curves that overlaid the market data visualization in both DOA configurations. In the high DOA routine-operation DL example, a long-period SMA (200-day), a short-period SMA (50-day), and an RSI indicator have been used. To simplify the training materials for this experiment, this experiment did not include the RSI indicator which was previously presented in the DL models. Further, the periods of the two SMA indicators have been shortened to 20-second and 10-second, as the original periods are too long for a lab-control experiment that has a limited timeframe.

3.3.3.3 Decision selection and action implementation

For the moderate DOA configuration, functions in the last two information-processing stages (i.e., decision selection and action implementation) have been completely allocated to the trader. Although the concept of trend following trading might have been familiar to the participants already (i.e., “buy low sell high” in a market that does not involve short-selling), the participants must be trained to utilize features provided by earlier stages of automation. For example, the participants should familiarize themselves with the characteristics of the two SMAs, and understand how SMA can help with the identification of market trends.

For the high DOA configuration, a trading algorithm was developed to trade on the crossovers of the two moving averages to replace the participants in the last two stages. The original “two moving averages” method was best described by Murphy (1999, p. 203), referred to as the Murphy method; whereas this experiment used an “inverted two moving averages” method. It can be seen in Table 5 that the Murphy method and the “inverted moving averages” method responded to similar system states about the market that has been identified in the information analysis but had inverted execution behaviours.

Table 5. Distinct Operation Logic of the Murphy Method and the Inverted Two Moving Averages Method.

System State regarding the Market	Decisions Made with the Murphy Method	Decisions Made with the Inverted Two Moving Averages Method (Experiment 1)
The short-period moving average **crosses above** the long-period moving average	Buying signal	Selling signal
The short-period moving average **crosses below** the long-period moving average	Selling signal	Buying signal

The inverted two moving averages method was chosen over the Murphy method to create a realistic simulation of a most common market condition. Ellis and Parbery pointed out that moving averages are lagging indicators and therefore, the Murphy method may only be profitable in a market that has major trends (2005). The latency simulated by the AUTRASS created a slippage between the intended market price of a trade at which the decision was made ($P_{\text{m, buy}}$ or $P_{\text{m, sell}}$) and the market price at which the execution was completed ($P_{\text{m, buy'}}$ and $P_{\text{m, sell'}}$). Indeed, if an SMA crossover occurred between two timestamps (let them be t and $t + 1$), an execution submitted using the Murphy method would be delayed and may take place at $t + 1$. In a trending market, as described hypothetically in the left portion of Figure 10, the Murphy method may be still profitable despite the slippage, because the trend is steady and strong ($P_{\text{m, buy'}} < P_{\text{m, sell'}}$). However, in a market that only fluctuates within a narrow range, as shown in the right portion of Figure 10, following small trends may not be possible with the Murphy method and the trading system may consistently perform poorly ($P_{\text{m, buy'}} > P_{\text{m, sell'}}$).

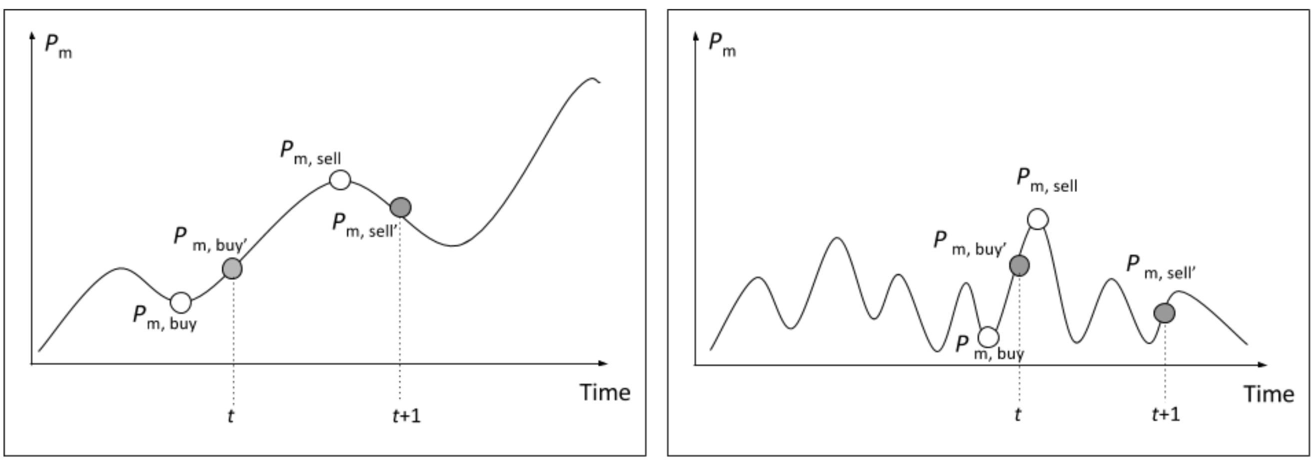


Figure 10. Trading with the Murphy method in a market with (left) or without major trends (right).

Since market trends are relatively rare in a real trading environment (Schlossberg, 2005), with limited resources for recruiting and scheduling the participants, it is more feasible for this first experiment to simulate a more common market condition that has no major trends. In this experiment, the "inverted two moving" averages method was adopted to develop a profitable trading algorithm in this simulated market condition. As shown in Figure 11, in a market that has no major trends, this trading algorithm would be able to accumulate profits over a certain period by making a selling execution on the same number of shares as soon as a rebound ended (at timestamp *t*) and a buying execution on a fixed number of shares when the next rebound began (at timestamp *t* + 1).

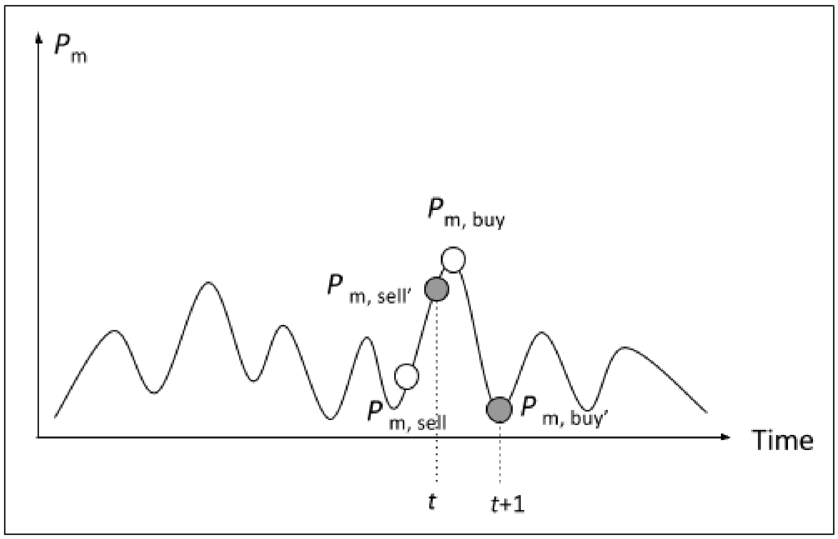


Figure 11. Trading with the inverted two moving averages method in a market without major trends.

The automation used in the high DOA configuration has been designed in such way that the automation was solely responsible for making decisions and executing trades, following a rigorous engineering standard which is typical in the use of trading algorithms (Kumiega & Van Vliet, 2012). At the same time, the participants were asked to carefully monitor the automated trading. The monitoring performance was evaluated through a fault detection task, in which the participants were asked to monitor the automation entering a simulated market and reported on the possible realized loss upon exiting the simulated market. This dissertation defines the pair of a buying execution and a corresponding selling execution that followed as a *buy-sell pair*. In short, to perform the fault detection task, the participant must be able to compare the $P_{m, buy'}$ and $P_{m, sell'}$ of the most recent buy-sell pair. It should be noted that the selling execution at time stamp *t* and the buying execution at time stamp *t* + 1 as shown in Figure 11 belong to two different buy-sell pairs. To monitor for possible automation failure at time stamp *t* + 1, the participants must wait for the next selling execution that will be made by the automation. The fault detection task will be further described in section 3.4.4.

The trading algorithm used in the high DOA configuration had a rigorous operation logic which was solely based on system states about the market dynamics. Obviously, the automation did not take any responsibility for the temporal performance of a portfolio or the eventual revenue of the trading system (i.e., a losing portfolio or a losing trading system would not change the operation logic of the trading algorithm). To ensure the participants effectively monitor the work of the automation (and in a future trading system where the participants can be assigned to more authorities in trading - deciding opportunities for intervention), the participants were trained with the two SMAs and the two system states described in Table 5 in which the trading algorithm would make trading decisions (i.e., SMA crossovers); whereas the logic of the trading algorithm for identifying which trading decision to make was not made available to the participants.

3.3.4 Financial Market Data

In the experiment, AUTRASS replayed historical market data. Each training scenario was designed to last 5 minutes. To reduce learning effects, each formal scenario was designed to last 10 or 15 minutes and used a different financial market data set.

Eight data sets were purchased from an online financial data source (Trading Physics, n.d.) which has been previously used in other studies (e.g., Cartea, Jaimungal, & Penalva, 2015). These data sets contained tick-by-tick data aggregated in 5 seconds (matches the update interval of the trading simulator). These data sets were originally recorded on eight trading days in the year of 2010. According to the suggestion of a subject-matter expert, these eight trading days were chosen because they had moderate *high-low spreads*. High-low spread, representing the total price movement over the trading day, was calculated by subtracting the lowest price of the day from the highest price. The high-low spread for each of the 252 trading days in the year of 2010 was calculated, $M = 1.519$, $SD = .955$. The 122 trading days which have the highest high-low spreads and 122 trading days which have the lowest high-low spreads were excluded from the selection. Finally, tick-by-tick data for the remaining eight trading days, which have moderate high-low spreads, were selected as the data source for this experiment, $M = 1.365$, $SD = .020$. The first half hour of each trading day (i.e., 9:30 am to 10:00 am) is arguably the most volatile time in the financial market, typically showing the highest bid-ask spread (Ahn & Cheung, 1999). In the finance industry, traders avoid trading in the first half hour of a trading day (TradingSim, n.d.). The eight data sets were subsequently adopted from the original data starting at 10:00 am on each trading day.

The two DOA configurations used distinct data sets due to consideration for the design of the tasks in the high DOA. This design consideration will be further discussed in section 3.4.4. Table 6 presents details of the eight data sets used in this experiment.

Table 6. SPY Historical Market Data Sets.

Data Set	Starting Price	Original Trading Day	Total Timestamps	Total Duration (min)	DOA
TD1	111.76	August 2, 2010	60	5	Moderate
TD2	110.63	July 26, 2010	60	5	Moderate
TD3	106.76	July 8, 2010	60	5	High
TD4	117.67	April 8, 2010	60	5	High
D1	115.75	November 30, 2010	120	10	Moderate
D2	114.11	October 18, 2010	180	15	Moderate
D3	118.37	October 7, 2010	120	10	High
D4	118.98	September 22, 2010	180	15	High

3.3.5 Conventional Displays

Conventional displays, representing financial trading industry state of art, have been designed to ensure the basic use of the automation. Two conventional displays were developed - the first one was developed to support the moderate DOA configuration and the second one supported the high DOA configuration. It can be concluded from Table 7 that most display elements were shared by the two DOA configurations. These shared elements are: 1) market panel, showing quotes and technical analysis indicators; 3) fundamental history panel, showing the fundamentals of the financial product; 3) portfolio panel, showing the size and the average price of the portfolio and trading performance; 4) trading history panel, presenting a list of buying and selling executions that have been submitted or completed during the simulation; 5) the execution panel, which had distinct views in the two DOA configurations to support different responsibilities of the trader in the decision-making and the execution stages.

Table 7. Display Elements of the Conventional Displays for the Moderate DOA and the High DOA Configurations.

Moderate DOA	High DOA
Market panel	
Fundamental history panel	
Portfolio panel	
Trading history panel	
Execution panel for moderate DOA	Execution panel for high DOA

Figure 12 and 13 provide an overview of the two conventional displays. The following subsections describe each display element in detail.

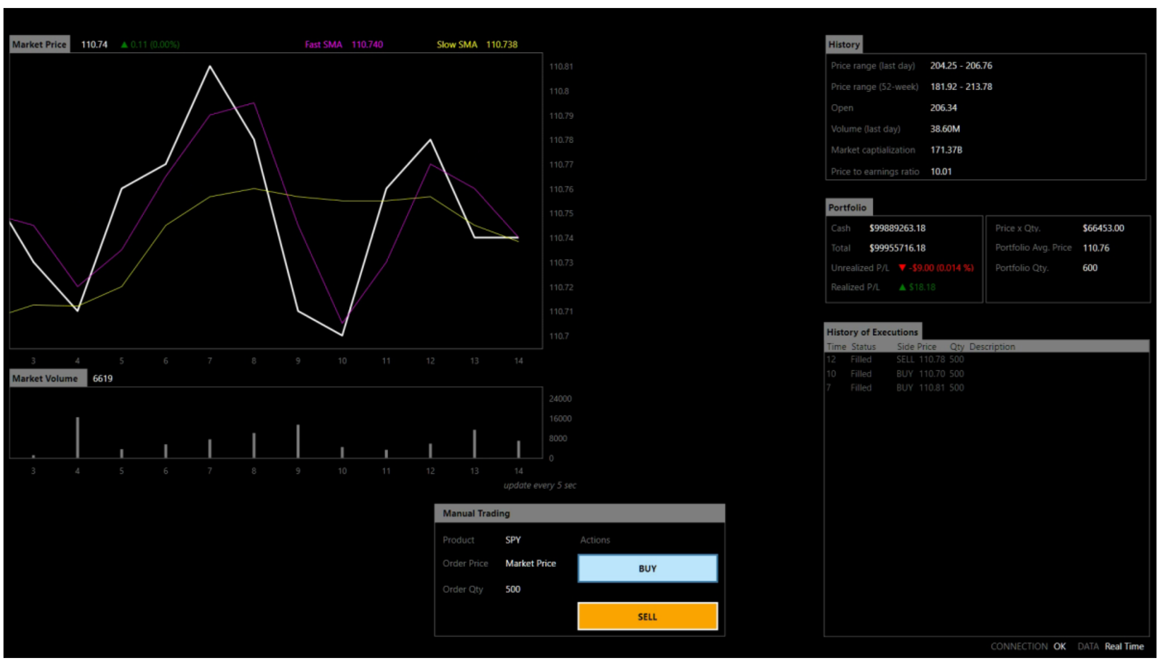


Figure 12. Conventional display for moderate DOA.

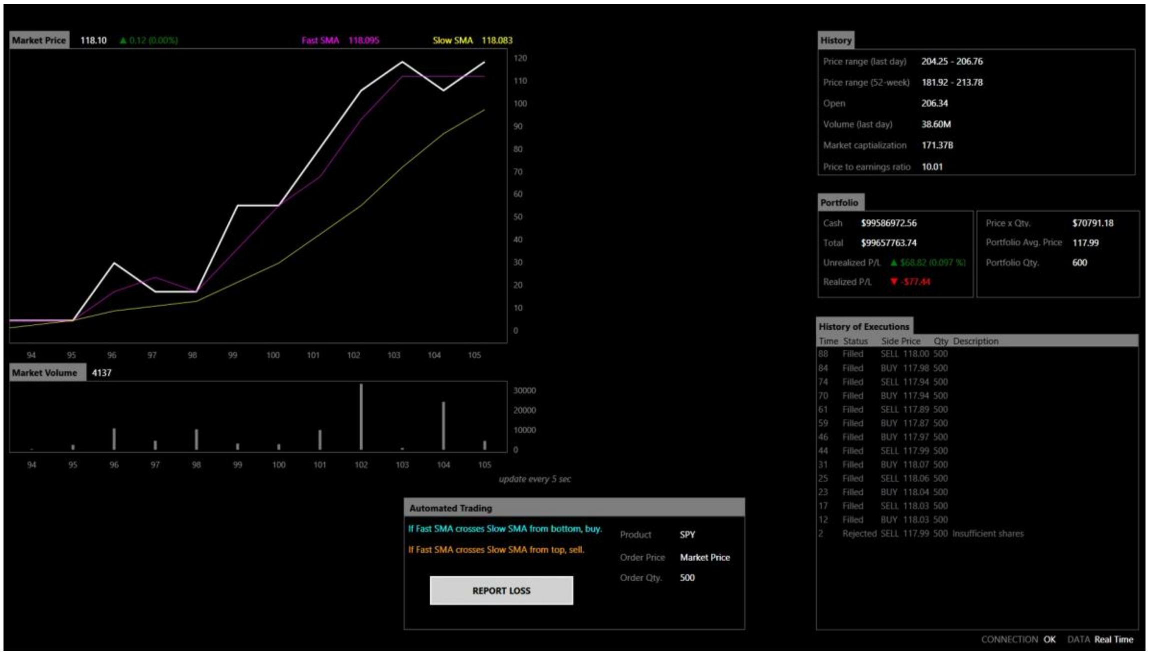


Figure 13. Conventional display for high DOA.

3.3.5.1 Market panel

In the left portion of the conventional displays there is a *market panel* as shown in Figure 14. The market panel consists of a market price chart and a market volume chart for the SPY. The market panel presents the market price of the SPY in a white line with a white numeric indicator showing its value (P_m, unit: dollar). In addition to the numeric indicator, an arrow showed the trends of the market price and was coloured to represent the direction (green: an uptrend; red: a downtrend). The market price chart is overlaid with a yellow curve and a purple curve, accompanied by a yellow numeric indicator (SMA_{slow}, unit: dollar) and a purple numeric indicator (SMA_{fast}, unit: dollar), representing the two SMAs. Although the market price and the two SMAs were graphically represented on the display, however, in a rare situation, an extremely small change in these indicators would be unlikely to recognized by the participants with normal visual acuity. The numeric values showing the same indicators were provided to the participants for dealing with this situation. Although volume is an indicator in technical analysis as it can be used to evaluate the strength of trends in the market price, it requires the participants to have a deeper knowledge of the law of supply and demand. Therefore, the market volume chart was provided just to improve the resolution of the simulation, but the participants were not expected to make decisions based on features provided on this chart.

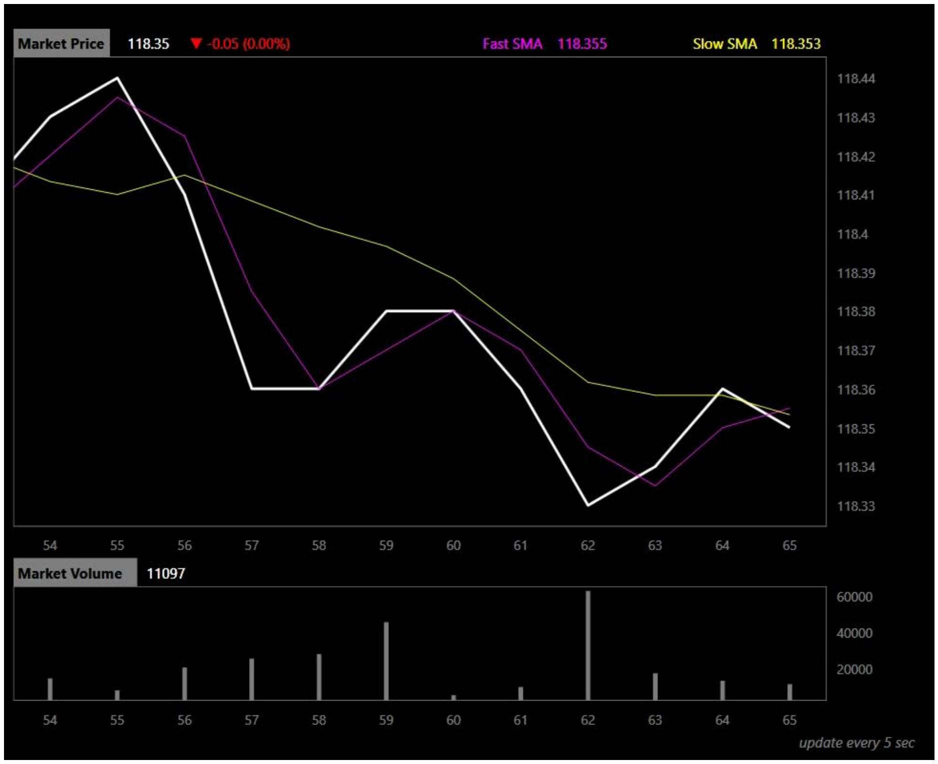


Figure 14. Market panel.

3.3.5.2 Fundamental history panel

On the top right of the conventional displays, the fundamental history panel presents features regarding the long-term fundamentals of the SPY (Figure 15). They are: 1) price range of the day before the current day, which was the day when the historical market data originally came from (unit: dollar); 2) price range of the past 52 weeks, indicating the long-term performance of the SPY (unit: dollar, Investopedia, 2003a); 3) opening price, the market price at which the SPY was first traded on the current day (unit: dollar); 4) total market volume of the day before the current day (unit: dollar); 5) market capitalization, the total value of the outstanding shares of the SPY (unit: billion dollar). Features provided on the fundamental history panel are related to the fundamental analysis and might not be useful for participants who have limited knowledge in financial trading. However, these features are typical in financial trading software and therefore, the fundamental history panel was included to provide a realistic simulation environment.

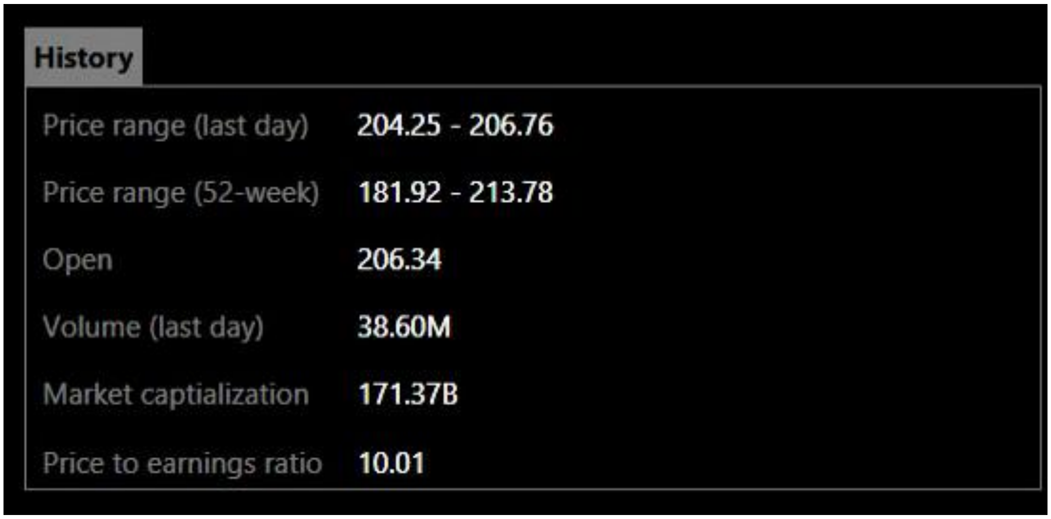


Figure 15. Fundamental history panel.

3.3.5.3 Portfolio panel

The *portfolio panel* presents several key portfolio indicators: 1) portfolio's average price (P_p, unit: dollar); 2) portfolio's size (S_p, unit: shares); 3) unrealized profit and loss of the current portfolio (*UPL*, unit: dollar); 4) realized profit and loss (*RPL*, unit: dollar), accumulating through each scenario of the experiment; 5) current asset of the financial product (unit: dollar), which can be calculated by multiplying the portfolio's average price (P_p) by the portfolio's size (S_p); 6) cash, including the initially provided cash for trading and money that has been cashed out from a winning portfolio (unit: dollar). For simplicity in the experimental design, the initially provided cash was 100 million dollars which mean the buying power was always unlimited within the full duration of each scenario; 7) total asset, including cash and the current asset of the financial product. The unrealized and the realized profit and loss were illustrated in numeric forms, with distinct arrows and colours provided to indicate the directions of the changes. The portfolio panel is shown in Figure 16.

Figure 16. Portfolio panel.

3.3.5.4 Trading history panel

The *trading history* panel (Figure 17) displayed properties of trades previously made by the trader in moderate DOA or the automation in in high DOA. These properties include: 1) timestamp at which the trade was made; 2) status (submitted: when a trade was submitted to the AUTRASS back-end; filled: when a trade was completed); 3) side (buy or sell); 4) market price (P_m) when the trade was submitted, equals to the market price when the trade was completed, as each trade was proceeded by the AUTRASS back-end at the current timestamp; and 5) size, the number of shares that have been bought or sold in the trade.

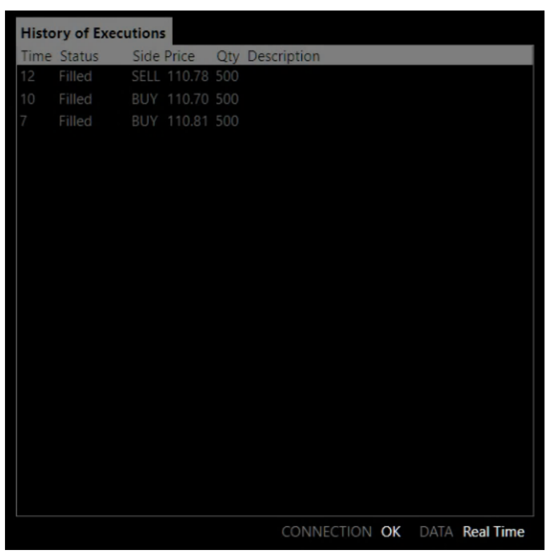


Figure 17. Trading history panel.

3.3.5.5 Execution panel for moderate DOA

The bottom-most panel, referred to as the *execution* panel, showed distinct information in the two DOA configurations. In moderate DOA, the execution panel contained execution buttons that can be used to perform buying or selling executions according to the limitation of the AUTRASS. At each timestamp, participants may choose to buy (at any time) or sell (if the portfolio's size permitted) 500 shares of the SPY at the current market price which was presented on the market panel (Figure 18), or do not act (holding to the current portfolio). If a buying or selling execution were made, both execution buttons would be disabled until the next timestamp. The AUTRASS front-end generated and submitted an order to the back-end. The order appeared on the trading history panel as a new row added to the top of the list.

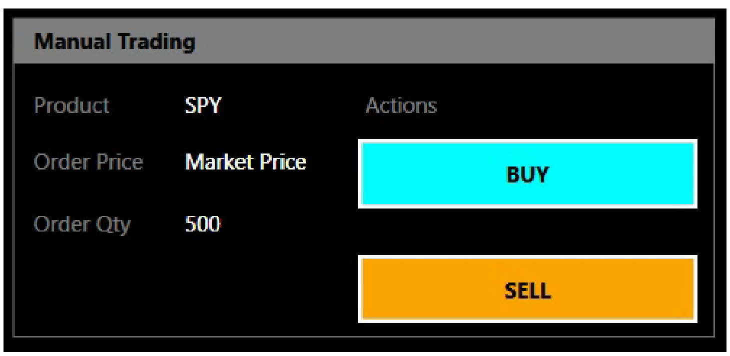


Figure 18. Execution panel for moderate DOA.

3.3.5.6 Execution panel for high DOA

To support the monitoring task in the high DOA configuration, an alternative execution panel illustrated the logic of the Murphy method that could also be concluded from the analysis automation (Figure 19). Although the automation used a rebound trading algorithm which did not have the exact behaviour as with this method, the conditional logic of the algorithm was not explicitly presented on the execution panel. The Murphy method was provided for training on how to interpret the two-SMA information. At the bottom of this panel, a report loss button was provided for participants to complete the fault detection task which will be further introduced section 3.4.4.

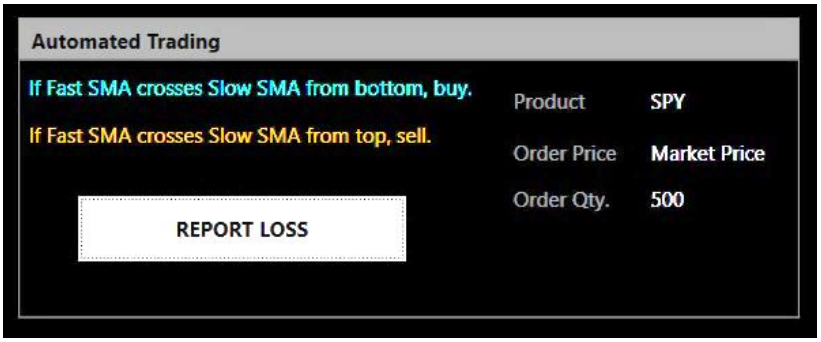


Figure 19. Execution panel for high DOA.

3.3.6 Ecological Displays

Similar to the design of the conventional displays, two ecological displays were designed and implemented to support the two DOA configurations. In the following subsections, the author first presents two design concepts for designing visualizations that can provide additional help for each of the two DAO configurations. Further, the author organized the ecological displays by adding these visualizations to the conventional displays.

3.3.6.1 Design concept inspired by a base model to support both moderate DOA and high DOA

The first design concept was based on the base AH previously presented in section 2.3.1 and therefore, this design concept was independent of DOA. This design concept is supported by the theoretical foundation of EID to provide knowledge-based support for participants to cope with unanticipated situations in both the moderate DOA and the high DOA scenarios. Based on the base AH previously modeled in this dissertation, a causal relationship between the financial market, the portfolio and the buying and selling executions to achieve a maximum rate of revenue in trading has been identified. In this dissertation, this relationship is referred to as the *market-portfolio-execution* relationship. This relationship has been shown on the base AH where multiple levels are logically connected, including the generalized function (where the execution functions take place), the physical function (where some market and portfolio and functions can be found), and the physical form (where other market functions and portfolio functions lie). Although the allocation of these functions to traders and automation varies through the DOA, as shown in the different DOA layers previously modeled, the market-portfolio-execution relationship remains consistent. Therefore, it is possible that this design concept can derive an ecological display that may support both moderate DOA and two which differ in their DOAs.

The next step, according to the EID principles, is to take information out of this relationship, quantify the information, and generate a list of variables that can be graphically represented on a display (Burns & Hajdukiewicz, 2004). The causal relationship can be quantified in equation (1), with several constraints extracted from the base AH.

$$UPL = (P_m - P_p) \times S_p \qquad (1)$$

UPL is the portfolio's unrealized profit and loss of a financial product (unit: dollar). P_m is the market price of the financial product (unit: dollar). P_p is the portfolio's average price of the same financial product (unit: dollar). S_p is the portfolio's size (unit: share). P_p and S_p were determined by two trading executions that are available to the traders: 1) if the trader makes a buying execution, S_p will increase, and P_p will be updated accordingly; 2) if the trader sells shares, S_p will decrease, and P_p will remain unchanged. Together, P_m, S_p, and P_p determined *UPL*. The consequences of the trading executions were decided by *UPL* and related to realized profit and loss (*RPL*, unit: dollar), which is the definite revenue, of the trading system. When *UPL* > 0, there is a *winning portfolio*. The traders may sell shares to convert the profiting portfolio into cash, in other words, realized a profit. When *UPL* < 0, there is a *losing portfolio*. If the traders decide to sell their shares at this point, the unrealized loss becomes a cost, in other words, realized loss.

Market-portfolio-execution visualization

Equation (1) and the accompanying constraints (*UPL*, P_m, $P_{p,}$ and S_p) can be further designed into a visualization that may support the hierarchical behaviour of the trader, inspired by Rasmussen's SRK taxonomy (Burns & Hajdukiewicz, 2004). The skill-based and rule-based behaviour would be given priority over the knowledge-based behaviour in the representational design. Burns and Hajdukiewicz pointed out that operators should be able to take actions directly from the perceptual information (skill-based behaviour), or at least, there is a one-to-one mapping of the visualizations onto the tasks that should be performed (rule-based behaviour). Further, showing multivariate constraints in one visualization makes complex relationships easy to understand by operators. With these suggestions in mind, a multivariable visualization was developed for this design concept. The mechanism of this visualization is graphically presented in Figure 20. This visualization was implemented in AUTRASS as shown in Figure 21. The multivariable visualization was presented when the trading system has 1) a profiting portfolio (i.e., unrealized profit, $UPL > 0$) or 2) a losing portfolio (i.e., unrealized loss, $UPL < 0$). In each case, there were two stacked boxes having the same width. The width of the two boxes represented the portfolio's size in shares (S_p). The height (P_p) of the box with a thick black border represented the average price in the portfolio. The market price (P_m) was plotted on the Y-axis as well. The shaded portion (green) on top of the box showed the unrealized profit in the portfolio. The shaded portion (red) inside the box showed the unrealized loss in the portfolio.

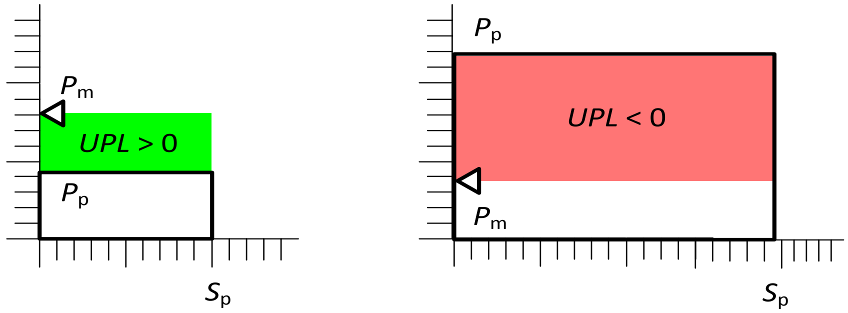


Figure 20. Mechanism of the market-portfolio-execution visualization: A profiting portfolio (left) and a losing portfolio (right).

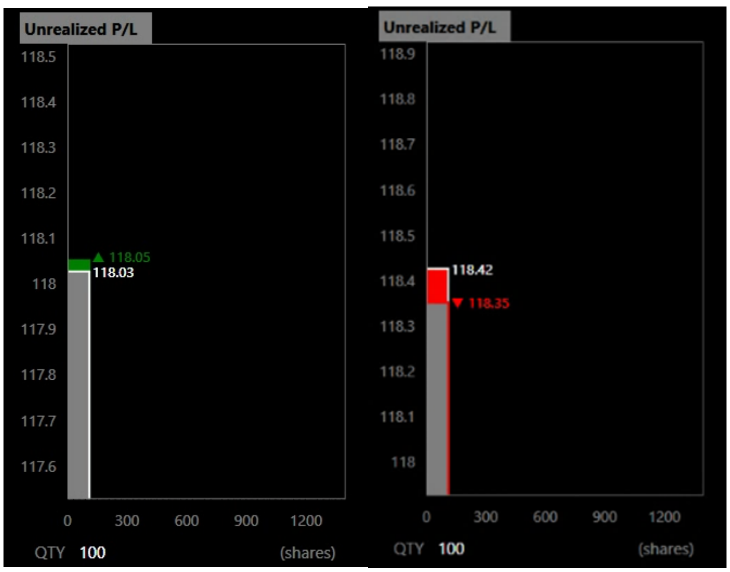


Figure 21. Implementation of the market-portfolio-execution visualization: A profiting portfolio (left) and a losing portfolio (right).

The viewport of the market-portfolio-execution visualization is generally stable throughout the experiment to ensure the information was perceived consistently. That being said, the Y axis of this visualization was generally limited to a $1 range, and the market price generally fluctuated within this range. On the X axis, the maximum limit was initially set to 1200 shares and would be doubled only when the current portfolio's size was close to the limit. For example, the maximum limit would be rescaled to 2,000 shares if the participants held a 1,300 shares portfolio (initial 100 shares plus extra shares - gained by performing two buying executions, 2 × 500 shares). In the case of high DOA, the X axis would not change because the trading algorithm would not hold more than 600 shares.

In summary, the first design concept was based on the base AH and derived the market-portfolio-execution visualization to make clear a causal relationship in the work domain. The causal relationship would be consistent in all DOAs that would be possible with trend following trading. The market-portfolio-execution visualization is similar to how ecological displays are typically designed. It explicitly represented the information structure of the work domain and is expected to support problem-solving in unanticipated market situations.

3.3.6.2 Design concepts inspired by DOA-layered models to support high DOA

The high DOA configuration may be more vulnerable to the disturbances caused by the slippage, as the trading algorithm followed a rigorous logic and the participants were unable to intervene. However, the market-portfolio-execution visualization provided neither a) problem-solving support for understanding the automated decision-making process and b) procedural support for detecting automation failures. As previously discussed in section 2.5.2.1, the DOA layer may be particularly useful for adding more variables and constraints that are not visible on the base models, suggesting opportunities to develop additional visualizations for specific DOAs.

States-task visualization

To develop a supplement to the market-portfolio-execution visualization that can specifically support the different task requirement in the high DOA configuration, new design concepts were inspried by consulting with the previously made DL models for trend following trading in the routine operation (Figure 7) and in the unanticipated operation (Figure 8). The AH for modeling trend following trading (Figure 3) was also used as a reference. Rationale for the design concepts was to understand how the DOA has been layered in four regions that corresponded to the stages of automation and how to conclude a causal relationship concerning the AH model. As the result of this comparison, several functions were extracted from the models and the DOA layers to develop problem-solving support and procedural support. After that, the two types of support were integrated into a states-task visualization.

Problem-solving support. Three DL functions were captured from multiple stages of automation and were used to develop knowledge-based support for participants in coping with unanticipated situations. These functions used different function allocations in the routine operation DL and the unanticipated DL. They are “system state” (analysis automation), “goal state” (decision automation) and “task” (action automation). A comparison of the different function allocations and the connections to the trend following trading AH are presented as follows.

The system state represented the calculations of SMA relevant to two Abstract Functions, the “flow model of capital” and the “flow model of market information”. The goal state described that to profit from the market; the automation must identify a trading signal when a crossover of the two moving averages was achieved (Functional Purpose: “to achieve a maximum rate of profitable revenue”). The task specified the buying and the selling executions of the automation and indicated that the performance of the executions should be evaluated by the participants (Generalized Functions: “to buy”, “to sell”, and underlying Physical Functions and Physical Forms). The trend following trading AH (Figure 3) showed that most of the corresponding AH functions were allocated to the automation, suggesting that the automation had high authority in the high DOA configuration.

The routine operation DL demonstrated that all three DL functions would be successfully established by the automation if no slippage occurred in this environment (Figure 7: Step 3, 4, and 5). In this case, the relationship between these functions was anticipated and no knowledge-based reasoning would be needed. In contrast, in unanticipated situations, this relationship was more complex due to the influence of the slippage. Indeed, the unanticipated DL showed that all three functions must be understood by the participants (Figure 8: Step 3, 4, and 6), suggesting that a common monitoring strategy would not be sufficient in this situation. The slippage was attributed to the latency simulated by the AUTRASS (Physical Form: “latency”) and was coupled with both system state and task. The coupled relationship could not be directly perceived by the participants through the conventional displays. System state and goal state were represented on the market panel as SMA curves and crossovers. The participants may observe the trading history panel for information related to the execution completion time but could not directly perceive the latency. The conventional displays contained heuristic cues of the system state, the goal state, and the task but supported neither the interpretation of the slippage nor the representation of their relationship.

Graphically representing this relationship on the ecological displays should support participants in coping with all situations, including unanticipated situations where a more complex relationship existed. Visualizing this relationship in the unanticipated situations would support knowledge-based reasoning activities that were identified on top of the unanticipated DL (Figure 8: Step 4).

Figure 22 demonstrates how a graphical representation of this relationship could work. The dashed curve describes the system state by graphically representing the difference between the long-period and the short-period SMAs. The solid-horizontal line has zero vertical distance and inherently matches the timeline of the simulation. Whenever the dashed curve reached the solid-horizontal line, meaning a crossover of the two SMA curves occurred on the market panel, the desired goal state would be achieved. If the trading algorithm decides to execute on this goal state, a circle would appear on the dashed line, representing the type of the execution. The latency caused by the slippage could be identified by visually scanning the position difference between the SMA crossover and the circle on the dashed curve. Following the EID principles (Vicente & Rasmussen, 1992), knowledge-based problem solving is supported with the entire problem space demonstrated in this representation to help participants understand factors that complicated the relationship.

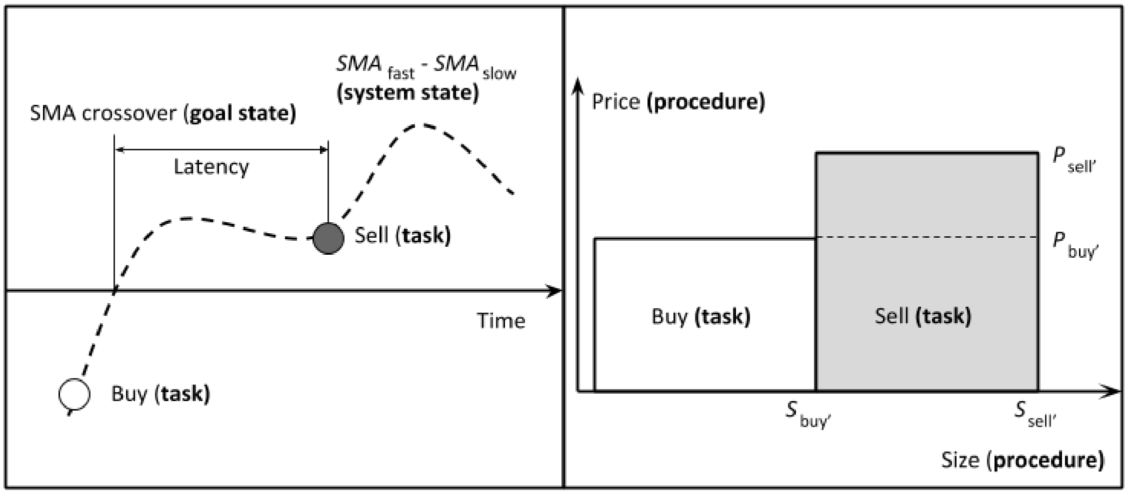


Figure 22. Mechanism of the states-task visualization for high DOA: Problem-solving support (knowledge-based, left) and procedural support (skill- and rule-based, right).

Procedural support. Procedural support was provided for detecting automation failures at the skill- and rule-based levels. The trading algorithm ran iteratively to capture a buying signal that was associated with a lower market price and the following selling signal that was associated with a higher market price. Although the trading algorithm was fine tuned to perform well in the simulated market, it was still possible that the automation experienced a temporal loss which should be detected by the participants (see section 3.4.4 for the task descriptions). The temporal performance of the automation in the most recent iteration relative to the current timestamp can be described using the realized profit and loss that was gained through the buying and selling executions:

$$RPL_{(\text{buy'},\, \text{sell'})} = P_{\text{sell'}} \times S_{\text{sell'}} - P_{\text{buy'}} \times S_{\text{buy'}} \qquad (2)$$

The relationship described in equation (2) was visualized in the right portion of Figure 22. Since the trading algorithm always traded with the same number of shares, $S_{sell'}$ equals to $S_{buy'}$. The participants may directly observe the temporal performance of the automation by comparing the heights of the two boxes which essentially represented $P_{buy'}$ and $P_{sell'}$. The boxes include parameters related to the "task" function that has been identified on the DL models. This design concept utilized the procedure that was formulated and executed by the automation and should support rule- and skill-based processing. The temporal performance of the trading algorithm can be characterized by patterns of the height difference of $P_{buy'}$ and $P_{sell'}$ representing a winning or losing buy-sell pair. Following the EID principles (Vicente & Rasmussen, 1992), a consistent skill- and rule-based mapping was developed between the pattern and the algorithm performance.

The design concepts for the states-task visualization was polished to include a variety of visual cues. Distinct colours were used to illustrate functions that had different meanings. For example, for the circle representing the executions on the left portion of the visualization, cyan was used to represent a buying execution and amber was used to represent a selling execution. The box representing the parameters of the most recent buying execution was consistently assigned a gray colour, showing that the buying execution alone did not decide the trading performance. If the corresponding selling execution was associated with a market price that was lower than the buying price (i.e., $P_{sell'} < P_{buy'}$), the box representing the parameters of the selling execution would be coloured green, showing that the trading performance was failing. As shown in Figure 23, the execution panel for the high DOA which has been used in the conventional displays has been modified to incorporate the states-task visualization.

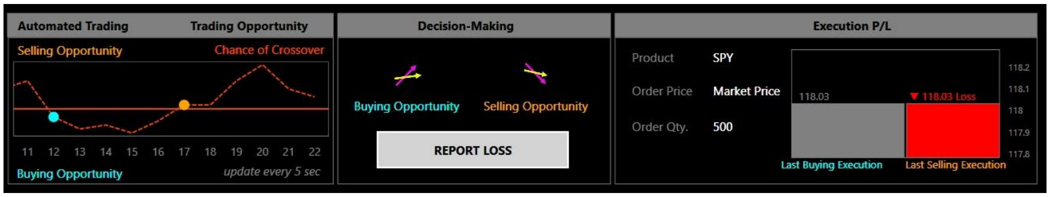


Figure 23. Integrating the states-task visualization to the execution panel (high DOA).

In summary, while the market-portfolio-execution visualization may be still useful to assist the traders in maintaining some aspects of their awareness, the problem-solving support (left portion, Figure 22 and 23) should help the participants develop knowledge about how the automation selected decisions and whether the slippage caused automation failures. On the other hand, the procedural support (right portion, Figure 22 and 23) would not overcome the difficulty associated with the knowledge-based processing but could provide skill- and rule-based support to fault detection.

3.3.6.3 Putting it all together: Designing ecological displays

The ecological displays were built on the conventional displays and included the additional visualizations that have been described in the previous sections, as presented in Table 8. Since the market-portfolio-execution visualization was based on the base AH, it was adopted in the design of both the low DOA and the high DOA configurations. The states-task visualization was designed to specifically support the monitoring for the high DOA configuration; therefore, this visualization was incorporated to the high DOA execution panel.

Table 8. Display Elements of the Ecological Displays for the Moderate DOA and the High DOA Configurations.

<table>
<tr><th>Moderate DOA</th><th>High DOA</th></tr>
<tr><td colspan="2">Market panel</td></tr>
<tr><td colspan="2">Fundamental history panel</td></tr>
<tr><td colspan="2">Portfolio panel</td></tr>
<tr><td colspan="2">Execution history panel</td></tr>
<tr><td colspan="2">Market-portfolio-execution visualization*</td></tr>
<tr><td>Execution panel for moderate DOA</td><td>Execution panel for high DOA and the states-task visualization*</td></tr>
</table>

* New in ecological displays.

Figure 24 and 25 show distinct ecological displays that have been used in the moderate DOA and the high DOA configurations respectively.

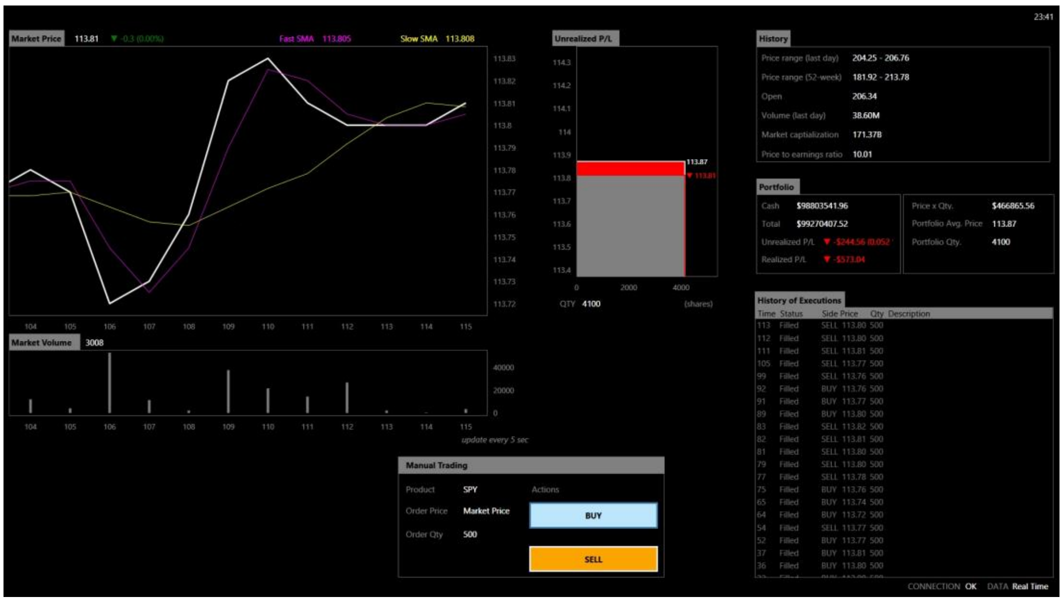


Figure 24. Ecological display for moderate DOA.

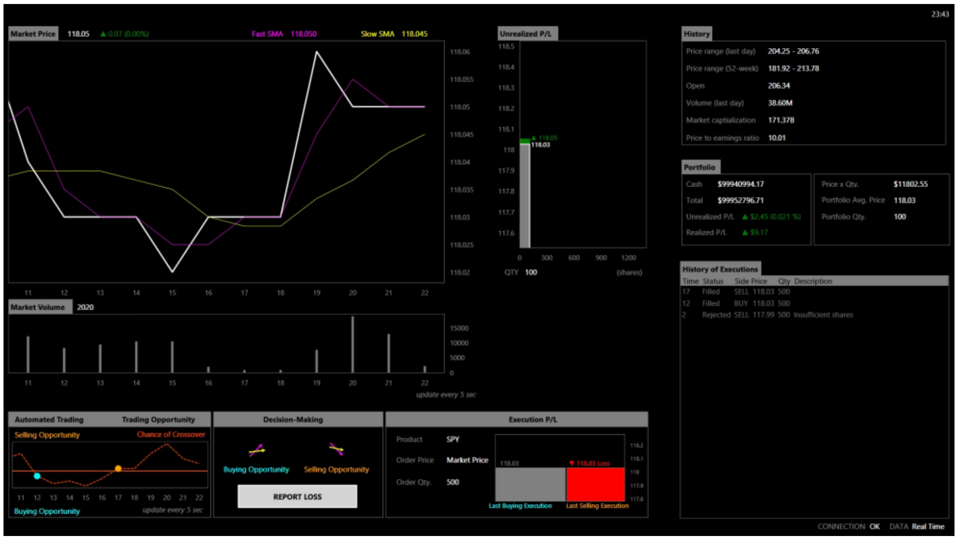


Figure 25. Ecological display for high DOA.

3.4 Method

3.4.1 Experimental Design

This dissertation defines "scenario type" in equation 3 according to the individual levels of the DOA and the display type:

$$Scenario\ type = DOA\ \text{-}\ display\ type \qquad (3)$$

Each participant completed four scenarios denoting combinations of the two DOAs and the two display types. Therefore, the experiment was generally a one-way, within-subject design (scenario type: moderate-conventional, moderate-ecological, high-conventional and high-ecological). Scenarios that were considered in the data analysis (referred to as the *measurement scenarios*, not including the training scenarios) were completely counterbalanced for each participant to reduce learning effects. Other than scenario types, there were additional within-subject independent variables that were used with certain dependent variables. These independent variables were SA level (used with the SA dependent variable), area of interest (used with the eye-tracking dependent variable) and workload subscale (used with the perceived workload dependent variable).

The following sections introduce procedure, demographics of the participants, descriptions of the different tasks performed by the participants in the moderate and the high DOA configurations, independent variables and dependent variables.

3.4.2 Procedure

Experiment 1 received initial ethics clearance by a University of Waterloo research ethics committee on November 30, 2015 (ORE #: 21061). All participants were recruited from undergraduate and graduate student applicants at the University of Waterloo. During the recruitment process, each applicant was asked to complete a screening questionnaire to report their age, visual acuity, and colour vision, and provided information regarding their degree program and their minor or option at the university if there was any. On the screening questionnaire, the applicant was also requested to state whether they have successfully completed at least one computer programming course, and how comfortable they would be with using spreadsheet software (e.g., Microsoft Office Excel) and information graphics software (e.g., creating a chart from a data set) each on a customized five-point scale (1 for strongly uncomfortable to 5 for strongly comfortable). A copy of the recruitment letter and an example screening questionnaire are provided in Appendix B and Appendix C of this dissertation respectively. The answers to the screening questionnaire were reviewed. Each selected participant was invited to attend an experimental session that will be described in the following paragraphs. The experimental session lasted approximately two hours, but there were individual differences (typically within ± 20 minutes).

1. *Consent*. The participant was asked to read the information letter and provided their consent to participate in the experiment. Examples of the information letter and the consent form can be found in Appendix D and Appendix E respectively;
2. *Demographic questionnaire*. The participant completed a demographic questionnaire on their gender, experience using computers, and experience with financial trading. A copy of the demographic questionnaire is provided in Appendix F;
3. *Training slides and training scenarios.* The participant was provided with training slides introducing them to the basic concepts of financial trading (including most content in section 2.2 and 2.3 of this dissertation) and instructions on how to use the automation provided in each scenario and how to perform the designated tasks. The participant was first introduced with moderate-conventional, as the moderate DOA is a more familiar situation than the high DOA regarding the participant's responsibility in trading, and the conventional displays are less complex than the ecological displays. After that, the participant experienced the first training scenario on the same scenario type simulated with financial data set TD1. This training scenario included 60 time stamps and lasted 5 minutes. To avoid biases in trading decision-making, each scenario started with a neutral portfolio (i.e., *UPL* = 0). The starting size of the portfolio was 100 shares, and the average price was identical with the starting market price. The *RPL* started at zero as well and would be accumulated throughout the scenario. During the training scenario, the simulation was paused at a random timestamp to evaluate the participant's SA for the moment before the pause using designated SA queries. The SA queries were presented in a full-screen mode so that the participant was unable to revisit the simulation screen for clues. The SA queries were customized to describe the financial trading context and will be described in detail in section 3.4.6. The training scenario ended with feedback on the participant's risk preference (Appendix G) and perceived

workload (Appendix H) for that scenario. Similarly, the participant was presented with training slides and scenarios for other scenario types, which are moderate-ecological, high-conventional and high-ecological using financial data sets TD2, TD3, and TD4 respectively. It is important to note that all training materials did not include descriptions of the algorithm logic;

4. *Measurement scenarios.* The participant experienced four measurement scenarios. To balance the trade-off between the limited time and the quality of the training, the measurement scenarios were similar to the training scenarios regarding the designated tasks and the measurements during and after the simulation but had four major differences. First, the presentation order of the measurement scenarios was counterbalanced to reduce learning effects, whereas the training scenarios were presented in a fixed sequence to ensure the participant was appropriately trained. Second, the measurement scenarios used financial data sets D1, D2, D3, and D4 which led to a significantly longer simulation than those data sets used by the training scenarios, allowing the participant's performance and risk preference to be better measured. Third, in each measurement scenario the simulation was paused twice to facilitate more measurements of the participant's SA. Lastly, the participant's scan pattern was measured in the measurement scenarios but not in the training scenarios to reduce the duration of the experimental session. As a result, before the first measurement scenario started, the participant was asked to complete a 9-point eye tracker calibration task using a software tool provided by Gazepoint. According to the guidance of Gazepoint, the participant was required to complete at least 80% calibration points to proceed in this experiment. Should the participant cannot complete the eye tracker calibration within 20 minutes, they would be asked to withdraw from the experiment but would still receive their participation payment in full;

5. *Debriefing.* The participant was encouraged to comment on the experimental design and was debriefed about the objective of this experiment. After that, the participant was remunerated for their participation.

3.4.3 Participants

To robustly estimate the minimum number of participants that would be required to reasonably detect an effect in the data analysis, a prior power analysis was performed using G*Power 3.0 (Faul, Erdfelder, Lang, & Buchner, 2007). As suggested by Cohen (1988), the alpha was set as 95% (two-tail) and the desired power was set as at least .80. For a one-way (scenario type: moderate-conventional, moderate-ecological, high-conventional or high-ecological), within-subject experimental design, the minimum number of participants suggested to detect a medium effect (f = .25) is 24.

Eight females and sixteen males voluntarily participated in this study, and each was remunerated 30 Canadian dollars for their participation. The average age of all participants was 25.1 years (*SD* = 3.256). All participants were undergraduate and graduate students registered at the University of Waterloo. All participants reported they have a normal or corrected normal visual acuity (i.e., wearing glasses or contact lenses) and normal colour vision, and they would be comfortable interacting with numeric and colour visualizations that are commonly used in financial trading displays (rated at least 3 in the 5-point scale in both cases). All participants stated that they have successfully completed at least one computer programming course, indicating that the participant had a basic understanding of computer algorithms. Regarding trading experience, seventeen participants claimed no previous trading experience. Seven participants identified themselves as having previous trading experience, of which four participants had performed personal investment, one participant had taken academic courses related to financial markets. Only one participant claimed to have received professional training in financial trading by completing an internship in that industry.

3.4.4 Task Descriptions

3.4.4.1 Moderate DOA: Flexible trading task

The participant's task in the moderate DOA scenarios was flexible, which was to monitor the simulation screen for trading opportunities with automation support and perform trend following trading. The participants were explicitly told during the training stage to follow four restrictions set by the AUTRASS. These restrictions helped with the control of confounding in the experimental design but may limit the use of certain trading strategies. Some of these restrictions have been described previously. A summary of all four restrictions is being provided as follows.

1. Only one execution (buying or selling) could be submitted between two timestamps;
2. To simulate latency (or slippage) in financial trading, executions (buying or selling) may be filled to generate a successful trade at the next timestamp after the execution was submitted to the simulator;
3. The participant was provided with unlimited buying power (i.e., cash). Perceptually, there was no limit on the maximum portfolio's size the participant could achieved. However, because of the limit on the duration of each scenario and restriction (1), the participant could buy as many as 60,000 shares in the case of the 120-timestamp scenario or 90,000 shares in the case of the 180-timestamp scenario;
4. Any execution (buying or selling) must be performed at the current market price of the SPY with 500 shares. The participant was unable to specify the price at which order would be submitted.

All restrictions simplified the task and were expected to reduce individual differences that may confound the experimental design. Restriction (1) had a known drawback: participants were unable to immediately sell off a position larger than 500 shares. Restriction (4) reduced individual differences in typing in the number of shares to be traded. This restriction helped to reduce the noise caused by human errors in the resulting data. For example, slips and lapses could happen in the highly routinized task of specifying the parameters needed to submit an execution, as identified by Leaver and Reader (2015).

3.4.4.2 High DOA: Fault detection task

In the high DOA configuration, the trading algorithm performed a selling or buying execution of 500 shares at the current market price whenever a rebound trading opportunity occurred. The participant performed a fault detection task in parallel to the automated trading. The fault detection task was designed for the participants to detect temporal automation failures at the stages of decision selection and action implementation in the high DOA situation. Since the automation traded in a series of buy-sell pairs (t_1-$t_{1'}$, t_2-$t_{2'}$, t_3-$t_{3'}$, as shown in Figure 26), the participants must effectively monitor the simulation screen for realized profit and loss made by the automation through each pair. It can be seen in Figure 26 that the automation has achieved realized profits through the first two pairs, as the market price at which the buying execution was made was lower than that at which the selling execution was made ($P_{m, x, buy} < P_{m, x, sell}$, $x = 1$ or 2). A temporal automation failure occurred at the third pair ($P_{m, 3, buy} < P_{m, 3, sell}$), in which case, the participant should click on the "report loss" button as soon as possible to demonstrate they have detected the failure. The report loss button was located at the execution panel used in the case of high DOA configuration.

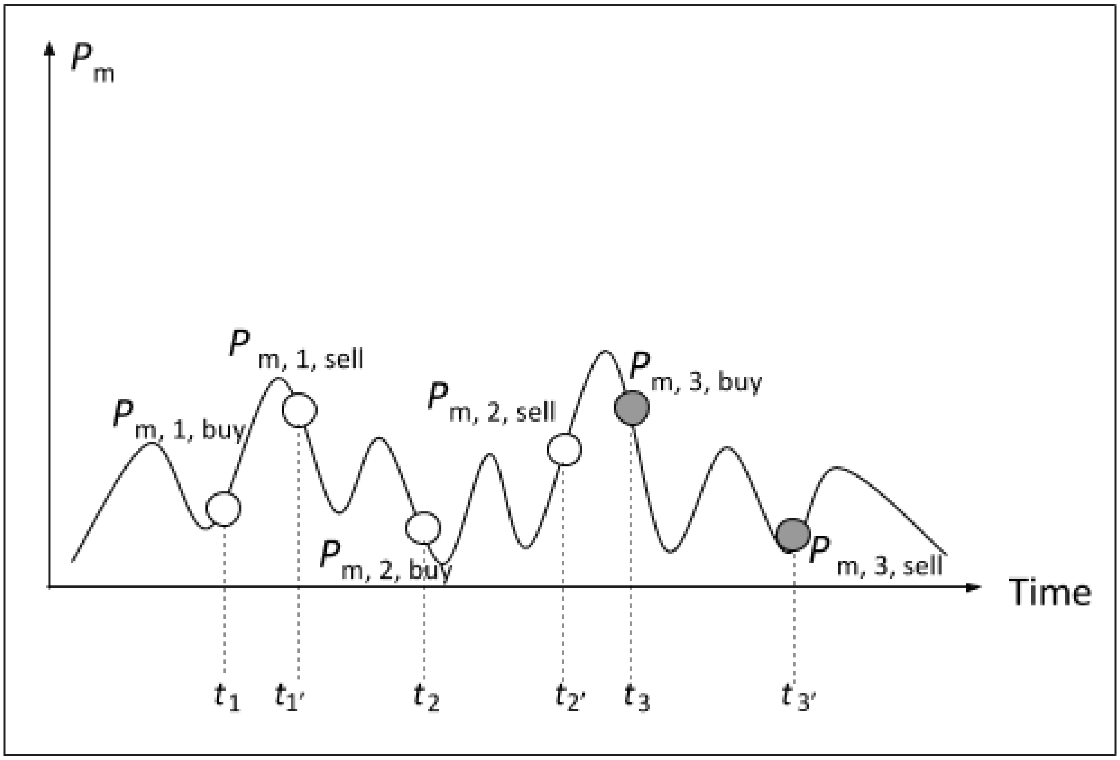


Figure 26. P_m overlaid with winning and losing buy-sell pairs (high DOA).

During the training stage of this experiment, the participants were explicitly told that constant temporal failures of the automation would accumulate a significant realized loss and caused a system failure. The system failure could be directly observed through the “realized profit and loss” indicator on the portfolio panel (i.e., $RPL < 0$). However, the temporal automation failure which would be reported in the fault detection task was only related to the trading algorithm and should be differentiated with a system failure. Details of the system failure will be described in section 3.4.5.3.

As have been explained previously, each participant experienced two counter-balanced scenarios with the high DOA. In each scenario, either financial market data set D3 or D4 was used. The presentation order of the two data sets was also counter-balanced for a strict experimental design. The reason for choosing these data sets for the high DOA scenarios is they triggered similar numbers of losing buy-sell pairs that should be reported by the participants. There were four losing buy-sell pairs occurred within D3 and five in the case of D4. Using D3 and D4 as the financial market data sets was an important design consideration in establishing a consistent baseline reference for evaluating the fault detection task performance across different display types within high DOA.

3.4.5 Independent Variables

All independent variables are listed in Table 9. The following sections explain how each independent variable was manipulated in this experiment.

Table 9. Summary of the Independent Variables (Experiment 1).

Independent Variable	Type	Level	Used with
Scenario type	Within-subject	Moderate-conventional, moderate-ecological, high-conventional or high-ecological	All dependent variables
Financial market data	Within-subject	D1, 2, 3 or 4	Manipulated in the experimental design (D1 and 2 for moderate DOA and D3 and 4 for high DOA) but was not included in data analysis
System state	Within-subject	Profiting, neutral or losing	Included in the evaluation of mean position size portfolio and decision preference in a guaranteed profiting situation
Area Of Interest (AOI)	Within-subject	Market, portfolio, trading history, market-portfolio-execution or states-task	Included in the evaluation of all eye-tracking measures: 1) the market, portfolio and trading history AOIS were evaluated with all scenario types; 2) the market-portfolio-execution AOI was involved in the evaluation of conventional display scenarios; 3) the states-task AOI was evaluated within the high DOA scenarios only
SA Level	Within-subject	SA Level 1 and 2, or 3.	Included in the evaluation

Independent Variable	Type	Level	Used with
			of SA rating
NASA TLX subscale	Within-subject	Mental, physical, temporal demand, performance, effort or frustration	Included in the evaluation of perceived workload

3.4.5.1 Scenario type

Scenario type was manipulated in the different scenarios. Simulation data for each measurement scenario was logged in a separate file to facilitate data analysis.

3.4.5.2 Financial market data

Different financial market data sets were counterbalanced by scenario for consistency and statistical validity. This independent variable was considered as a covariable and therefore not analyzed.

3.4.5.3 System state

System state has been previously introduced as a DL component. During the modeling stage (section 2.4.2.3 and 2.4.2.4) and the development stage of the automation for the high DOA (section 3.3.3), system state was described similarly but was limited to the state of the financial market. In moderate DOA, the participants not only interpreted the market dynamics but also managed the state of the portfolio and the performance of each execution, adding up to the eventual revenue of the trading system. That being said, system state must be evaluated in a broader sense than in the previous modeling to describe the overall state of the trading system. To evaluate this effect, the data analysis of this experiment defines the *RPL* that accumulated through every trading execution as the system state (profiting: $RPL > 0$; neutral: $RPL = 0$; losing: $RPL < 0$). System state reflected realized profits and losses and was independent of the unrealized profit and loss (e.g., a trader may have a realized loss and yet hold an unrealized profit in the portfolio). At any time during the simulation, the trading system may either encounter an anticipated routine situation (in the case of a profiting state or a neutral state) or an unanticipated failure (in the case of a losing state).

System state was not rigorously manipulated in the experimental design. Instead, it reflected differences in trading performance that was attributable to 1) financial market data, 2) slippage (i.e., unanticipated situations), 3) scenario type (i.e., DOA configuration and display type). It is noteworthy to mention that the impact of financial market data on the distribution of system states has been controlled by counterbalancing the financial market data sets. In moderate DOA, since the participants may use different strategies for the conventional and the ecological displays, it might be necessary to examine the differences in the distribution of the system states between participants and displays. In this case, system state was analyzed as an independent variable. For high DOA, however, the system state could not be evaluated as the trading algorithm was well calibrated to achieve a consistent performance with selected financial market data sets (i.e., D3 and D4) and would neither be influenced by participant nor display. In this case, system state was not analyzed.

It is also important to know that only a small number of dependent variables could be evaluated with system state. For example, it has been previously introduced that the participants' SA was only measured twice within each measurement scenario. Therefore, the SA ratings reflected the overall SA per scenario, not the individual SA for each system state. For dependent variables that could be evaluated per system state, the system state would be determined through a data preparation process, after the completion of the experiment. *RPL* data for each scenario were divided into some groups that reflect different system states. For a hypothetical example of this data preparation process, a participant completed a scenario that contains *t* timestamps. The author first captured the *RPL* reading for each timestamp. Let the readings be RPL_1, RPL_2, RPL_3, …, RPL_t. The author then categorized all *RPL*s into three groups by system state. Table 10 shows the grouping of the first 15 time stamps using hypothetical data. It should be noted that RPL_1 was flagged as an outlier in all scenario data and excluded from the experimental data because each scenario started with a neutral portfolio and the starting *RPL* was set as zero. The grouping result shows at which time stamps the simulator was in profiting, neutral or losing system states. Using these timestamps as references, the effect of the system state on the designated dependent variables could be examined.

Table 10. *RPL* Grouping According to System State Using Hypothetical Data.

Profiting	Neutral	Losing
RPL_2	RPL_1*	RPL_7
RPL_3	RPL_5	RPL_8
RPL_4	RPL_6	RPL_9
RPL_{13}	…	RPL_{10}
RPL_{14}		RPL_{11}
RPL_{15}		RPL_{12}
…		…

* RPL_1 was consistently zero because each scenario would start with a neutral portfolio and was excluded from the analysis.

3.4.5.4 Other independent variables

Other independent variables, such as the area of interest (AOI), SA level and NASA TLX subscales, were tied to specific dependent variables and will be further described in the next section.

3.4.6 Dependent Variables

Some dependent variables existed in this experiment, and they were categorized according to the research questions they attempted to answer (Table 11). First, four categories of measures (task performance, SA, eye-tracking, and workload) were used to determine the effectiveness of scenario type on financial trading performance. Second, the influence of scenario type on trader's risk preference was examined using a paper-based questionnaire inspired by McAndrew and Gore's findings (2013). For moderate DOA scenarios, additional quantitative measures were used to examine what risk-related strategies the participants possessed in different system states.

Table 11. Summary of the Dependent Variables (Experiment 1).

Research Question to Answer	Category	Dependent Variables	Type	Independent Variables (Number of Levels)	Data Collection Method
Research question 2: "Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?"	Task performance	End of scenario *RPL*	Ratio	Moderate DOA only: scenario type (2)	Simulation data
		Fault detection accuracy	Ratio	High DOA only: scenario type (2)	Simulation data
	Situation awareness	Mean SA score	Ratio	Scenario type (4)	Computer-administrated questionnaire
	Eye tracking	Total dwell time	Ratio (log-transformed)	Scenario type (4), AOI (6)	Eye tracking data
	Workload	NASA TLX rating	Ratio	Scenario type (4), NASA TLX subscale (6)	Paper-based questionnaire
Research question 3: "Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?"	Risk preference	Choice of options (as part of the fourfold pattern of preferences)	Binary (nominal)	Scenario type (4), fourfold situations (4)	Paper-based questionnaire
		Mean portfolio's size	Ratio	Moderate DOA only: scenario type (2), system state (3)	Simulation data
		Decision preference in a guaranteed profiting situation	Ratio	Moderate DOA only: scenario type (2), system state (3)	Simulation data
		Decision preference in a guaranteed losing situation	Ratio	Moderate DOA only: scenario type (2), system state (3)	Simulation data

3.4.6.1 Task performance measures

End of scenario RPL (moderate DOA)

Since achieving profitable revenues has been identified as the primary goal of financial trading and the participants had relatively high flexibility in making decisions with moderate DOA, *end of scenario RPL* was collected at the end of each scenario and was used in the data analysis to evaluate the task performance in this environment. To explore the ceiling of end of scenario *RPL* (i.e., the maximum revenue the participants could possibly achieve given the financial market data were known), a quantitative researcher was consulted for generating an optimized result. The optimization was performed with the same constraints experienced by the participants during the flexible trading task (i.e., maximum one buying or selling execution per timestamp, with unlimited buying power). Results of the optimization showed that it was possible to achieve up to $810 with D1 (120-timestamp data) and $1745 with D2 (180-timestamp data). These results were later compared to participants' performance to determine whether a ceiling has been achieved.

Fault detection accuracy (high DOA)

Since the operation logic of the trading algorithm and the financial market data sets used in the high DOA scenarios were both predetermined, mean *RPL* would neither be influenced by the individual differences nor the displays; therefore, mean *RPL* should not continue to be used as the task performance measure for the high DOA.

In the high DOA configuration, the participants performed a monitoring task, monitoring the behaviour of the automation and the status of the trading system. *Fault detection accuracy* would be an appropriate task performance measure in this setting. This dissertation defines fault detection accuracy as the number of losing buy-sell pair correctly reported by the participant divided by the total number of losing buy-sell pairs in each high DOA scenario. The denominator was the total number of losing buy-sell pairs that should be correctly reported. This denominator, as was previously described, was only determined by the financial market data set and would serve as an appropriate analysis baseline. There were in total four losing buy-sell pairs with D3 and five losing buy-sell pairs with D4.

3.4.6.2 Situation awareness measure

Situation Awareness Global Assessment Technique (SAGAT: Endsley, 1988, 1995) was used as a measure of the participants' SA. SAGAT was originally developed to provide a measure of SA in mission scenarios in simulated aircraft systems, and has become a popular method in the automation literature (e.g., Endsley, 2015; Kaber & Endsley, 2004) and the EID literature (e.g., Burns et al., 2008). Using SAGAT, participants are requested to self-report their perceived SA by answering a series of questions (randomly selected and categorized in separate SA levels) during random pauses of the simulation without accessing the flight status displays. After the completion of the simulation, SA should be scored by comparing the participants' answers to the real situations that are typically described in the simulation data. Subject-domain experts, such as professional pilots, are usually involved in the data analysis process to help to decide the scoring rubrics. As previously introduced, in each measurement scenario of this experiment, the AUTRASS simulation was paused at two random points in time with the displays blanked. During each pause, the participants were requested to provide answers to six questions based on their understanding of the last seen situation of the AUTRASS simulation. AUTRASS randomly drew each question from a predetermined pool of sixteen SA queries. For statistical stability, two questions were drawn for each SA level, and no duplicate SA queries were used during both pauses of each measurement scenario, to ensure that all aspects of the participants' SA were equally evaluated. The SA queries were framed in a way that encouraged the participants to search for the required information in three different situations:

First, five level 1 SA queries were designed to elicit the participants' perception of the present situation (i.e., the 5-second window before the pause during which SA was measured). These questions primarily asked about the market aspects of the simulation related to the performance of the acquisition and analysis automation which was included in both DOA configurations. The market aspects included market price (related to the acquisition automation), SMAs (related to the acquisition automation) and various system states (e.g., SMA crossovers, related to the analysis automation). The last level 1 SA query targeted at awareness of the most recent trading execution relevant to the decision selection and action implementation stages. The information required to report the participants' immediate perception of this aspect could be retrieved from the information presented in the tabular form on the trading history panel which was used in all DOA and display scenarios, whereas the ecological displays may provide more explicit support;

Second, the level 2 SA queries also reflected the present situation but were designed to elicit the deeper level of comprehension. There were five level 2 SA queries. Two of these questions were specific to the identification of the decisions on trading signals which were made by the participants in the case of moderate DOA or the automation in the case of high DOA. Three other level 2 SA queries required the participants to understand the probable cause of the present situation of the portfolio. Although many aspects of the portfolio have been illustrated on the portfolio panel of the conventional displays, the present situation of the portfolio depended on other factors, including the market dynamics and the trading executions. Since the causal relationship between the market, the portfolio and the executions have not been explicitly represented on the conventional display, to reason from the situation of the portfolio, the participants must also obtain a good awareness of the market and the trading executions, particularly in the case of high DOA where the automation did not take any responsibility for the portfolio. It is reasonable to expect that the market-portfolio-execution visualization presented on the ecological displays could provide additional help;

Lastly, six level 3 SA queries were framed to encourage the participants to predict most aspects of the AUTRASS in the next 5 seconds from the occurrence of the simulation pause. The 5-second window was set for the participants to recall their working memory, rather than their long-term memory (e.g., from the beginning of the scenario). Regarding the answer format, all level 1 SA queries were multiple choice questions which were appropriate to the specific level of understanding they tried to evaluate. Level 2 and 3 SA queries required the participants to provide more detailed, open-ended answers. A list of all questions in the pool is provided in Table 12.

Table 12. SA Query Pool (Experiment 1).

SA Level	SA Query	Answer Choice
1	"In the last 5 seconds, the market close price has gone:"	Up, down or flat
	"In the last 5 seconds, the slower moving average curve (yellow) has gone:"	Up, down or flat
	"In the last 5 seconds, the faster moving average curve (purple) has gone:"	Up, down or flat
	"In the last 5 seconds, was there a crossover of the two moving average curves?"	Yes or no
	"In the last 5 seconds, what was the most recent trade?"	Buy, sell
2	"What is happening with the unrealized profit and loss in your portfolio? Why?"	(Open-ended)
	"Is there a buying opportunity? Why?"	
	"Is there a selling opportunity? Why?"	
	"What is happening with the quantity in your portfolio? Why?"	
	"Has the most recent trading execution made any profit? Why?"	
3	"What will happen to the market close price in the next 5 seconds? Why?"	(Open-ended)
	"Do you think there will be a buying opportunity in the next 5 seconds? Why?"	

SA Level	SA Query	Answer Choice
	"Do you think there will be a selling opportunity in the next 5 seconds? Why?"	
	"Do you think there will be a buying execution in the next 5 seconds? Why?"	
	"Do you think there will be a selling execution in the next 5 seconds? Why?"	
	"What will be the status of the quantity in your portfolio in the next 5 seconds? Why?"	

3.4.6.3 Eye-tracking measure

Eye-tracking data were collected continuously through a software tool provided by Gazepoint and were analyzed to support the interpretation of the SA results. This dissertation defines a dwell as some fixations that can be determined as a visit in an area of interest (AOI) on the AUTRASS simulation screen. According to Holmqvist, Nyström, Andersson, Dewhurst, Jarodzka and Van de Weijer (2011), a higher dwell time (calculated per dwell) indicates that the participants might have poor SA, have difficulties in capturing information from an AOI, or need more time to gather the information. The total dwell time, defined as the sum of all dwell times on an AOI, has similar features in interpreting the SA results and should be used to describe long-term cognitive processes. Eye-tracking data for the 5-second window before each SA pause in the measurement scenario were used to determine the frequency and duration that the participants spent monitoring each AOI.

Figure 27 shows that the three primary AOIs were defined on the top portion of the market panel (which contained the price chart), the portfolio panel and the trading history panel. These AOIs represented the most important aspects of the simulator screen and were consistent with all scenario types (using all combinations of DOA and display type). The market-portfolio-execution AOI was defined on the market-portfolio-execution panel for the moderate-ecological scenarios and the high-ecological scenarios. The states-task AOI was defined on the execution panel of the high-conventional scenarios and the high-ecological scenarios.

Figure 27. AOI layout.

3.4.6.4 Workload measure

The original NASA TLX approach required supplementary paired comparisons of six subscales to determine the weight or importance of each subscale. Each comparison is being made using a 0-5 scale. In this experiment, perceived workload ratings were calculated based on an unweighted approach recommended by Nygren (1991). Nygren argued that with the original approach, one of the weights could be incorrectly determined as 0.0 and causes the final score to reflect only five components or subscales. On the other hand, the unweighted NASA TLX approach reduces the time consumption of taking the NASA TLX measure, and might improve the validity of workload measures. A copy of the NASA TLX questionnaire is provided in Appendix H.

3.4.6.5 Risk preference measures

Fourfold pattern of preferences

A paper-based risk preference questionnaire was developed based on McAndrew and Gore's findings (2013) on professional traders' risk preference. McAndrew and Gore described four situations that were different in the outcomes (prospect gains or losses) and the probabilities of achieving these outcomes (small-probability and medium- and large-probability). Further, as shown in Table 13, risk preference can be described as either choosing a risk-seeking decision or a risk-aversion decision, while facing certain outcomes and probabilities.

McAndrew and Gore compared the findings of prospect theory, Hertwig and Erev's findings (2009) on a decision from experience and their observations on professional traders and found different fourfold patterns of risk preference. The prospect theory suggested that people would typically perceive rare events as having more weight if people are told to make decisions from descriptions of those probabilities; therefore, people are risk-seeking in small-probability gains and medium and large-probability losses. On the contrary, if people are allowed to experience the outcomes and their probabilities through observations of events in the environment, rare events would be underweighted. Indeed, it has been found in Hertwig and Erev's experiment that people may be risk-aversion for small-probability gains. As McAndrew and Gore pointed out, it is not clear that how people make those decisions for medium- and large-probability gains and losses (shown as question marks in Table 13), though a speculation suggested that the same fourfold pattern with that described in prospect theory may exist.

Lastly and most importantly, McAndrew and Gore interviewed professional traders and documented their preferences on market situations that were associated with similar outcomes and probabilities (2013, pp. 189–191). To study traders' risk preference in experience-based choice, McAndrew and Gore conducted an Applied Cognitive Task Analysis (ACTA: Militello & Hutton, 1998) and synthesized interview data collected from the traders and developed a cognitive demands table. The cognitive demands table provided an overview of the difficult cognitive elements and the strategies used by expert traders to overcome the difficulties. McAndrew and Gore then mapped the cognitive demands to the fourfold pattern of prospect theory and characterized the traders' risk preference as risk-seeking or risk-aversion. Their work not only suggested that learning from a professional source might lead to a different fourfold pattern of preferences in comparison to those patterns derived from the case of descriptions of outcomes and probabilities and the case of experience through learning, but also provided a template of various market situations and choices that could be re-evaluated in a different experimental setting.

Table 13. Fourfold Pattern of Preferences for Description- and Experience-Based Choices.

	Description (prospect theory)		Experience through learning (Hertwig & Erev, 2009)		Experience through professional training (McAndrew & Gore, 2013)	
	Gains	Losses	Gains	Losses	Gains	Losses
Small probability	Risk seeking	Risk aversion	Risk aversion	Risk seeking	Risk aversion	Risk aversion
Medium and large probability	Risk aversion	Risk seeking	Risk aversion?	Risk seeking?	Risk seeking	Risk aversion

A subject-domain expert on automated trading supported the development of this risk preference questionnaire, by framing the market situations and decisions summarized by McAndrew and Gore into questions and options. One market situation about the medium and large probability gains described anticipating the effect of changes in market fundamentals which cannot be simulated with AUTRASS. This market situation was not included in the questionnaire design of the questions.

The questionnaire consisted of four questions. Each question described a market situation that was associated with an outcome and a probability of occurrence. There were two options for each question, and each option represented a decision. After the completion of each scenario, the participant provided an answer to each question by choosing their most likely decision from the provided two options. Table 14 contains annotations in an italic font that described the probability of each question (fourfold) and the outcome of each provide an answer (risk-seeking or risk-aversion). Readers should know that these annotations were not presented to the participants.

Table 14. Fourfold Pattern of Preferences Framed onto Questions and Options.

Situations	Question	Choice (a)	Choice (b)
Identifying emerging trends (**medium and large probability gains**)	"If there is a micro trend that the market will move to one direction, I would more likely:"	"Immediately place the position (buy) for the maximum profitability" (**risk seeking**)	"Wait until the market direction is clear" (**risk aversion**)
Responding to trend reversals (**small probability losses**)	"If there is a change in the market direction after a position has been placed (buy), I would more likely:"	"Believe this reverse trend is only momentary" (**risk seeking**)	"Immediately close the position (sell) to minimize losses" (**risk aversion**)
Detecting regime shifts (**small probability gains**)	"If the market has been in a shock (regime shifts) for quite a while (5 minutes, in the context of this experiment), I would more likely:"	"Immediately place the position (buy), as I believe the market will break the shock and the space for uptrend has been opened up" (**risk seeking**)	"Wait longer, as I still believe the market is in the shock" (**risk aversion**)
Taking actions following sudden interruption to supply (**medium and large probability losses**)	"If the market is collapsing (e.g., market crash), I would more likely:"	"Immediately close out all or most positions, or do nothing" (**risk seeking**)	"Promptly buy back the same financial product to lower the average portfolio price" (**risk aversion**)

Mean portfolio's size (moderate DOA)

The fourfold pattern of preferences was measured after the participant experienced each scenario and it was a subjective and robust measure. As a supplement to the qualitative, risk preference questionnaire, quantitative measures were developed based on the objective simulation data to characterize the participants' risk preference with higher fidelity. Since in moderate DOA scenarios the participants had the authority to trade, their strategies may be influenced by the DOA and the display type and possibly reflected their risk preference. For example, mean position size of the portfolio was used to measure the participants' portfolio management strategy. The participants did not directly perform trading executions in the high DOA configuration. Therefore, no quantitative risk preference was developed in that case.

The participants performed a flexible trading task with moderate DOA and the buying power was unlimited. The only hard limit on the maximum number of shares one could obtain was subject to the duration of the scenario (i.e., 90,000 shares with 180-timestamp data and 60,000 shares with 120-timestamp data). The optimization results to achieve a relatively good end of scenario *RPL* were previously discussed. With the optimized result with 180-timestamp data, the maximum portfolio's size would be 8,000 shares. In the case of 120-timestamp data, this number would be 2,000 shares.

Decision preference in a guaranteed profiting situation (moderate DOA)

The participants' risk preference could be further studied by understanding how they balanced risk against performance in each execution. This dissertation defines a *guaranteed profiting situation* as $UPL > 0$, which can be achieved when $S_p > 500$ (the author has previously described that for simplicity the minimum number of shares the participant could buy or sell was 500) and $P_p < P_m$. Guaranteed profiting provided a common ground for making comparisons and it naturally fit into the fourfold pattern of preferences as a high probability prospect gain.

Three decisions (or *risk levels*) were available to the participant during guaranteed profiting situations - either selling off the existing shares, holding the portfolio or buying more shares:

1. Selling the financial product (i.e., SPY) in a guaranteed profiting situation was associated with no risk because unrealized profit would be immediately converted to realized profit. In other words, a selling execution was a *risk-averse* execution;

2. A holding execution in a guaranteed profiting situation can be interpreted as the participants were taking a moderate level of risk. The market price could fall right after this timestamp, cutting their unrealized profit. Unless it was a hard fall, the participants were still confident about their ability to sell off the financial product before the unrealized profit reduced or turned to an unrealized loss. It can be concluded that a holding execution was a *moderate risk-seeking* execution;

3. Buying more shares to increase the position size when $UPL > 0$ was interpreted as having high risks, due to the polarized results it might cause similar to a large portfolio size. If the market moved up after the buying execution, a large portfolio would have more unrealized profit. If the market moved down, a large portfolio would lead to more losses. Therefore, a buying execution can be interpreted as a *high risk-seeking* execution.

Decision preference in a guaranteed losing situation (moderate DOA)

This current measure is similar to the previous measure. A *guaranteed losing situation* was defined as $UPL < 0$ (i.e., $S_p > 500$ and $P_p > P_m$) and the participants should be facing a high probability prospect loss. Similar to a guaranteed profiting situation, the participants were able to perform either a selling execution, a holding execution or a buying execution.

1. A selling execution means that the unrealized loss would be immediately realized. A selling execution was a *risk-averse* execution;

2. Holding to a losing portfolio was a *moderate risk-seeking* execution;

Buying more shares with a losing portfolio suggested that the participants were looking to take a higher level of risk. Therefore, a buying execution at this time can be interpreted as a *high risk-seeking* execution.

3.5 Research Hypotheses

This dissertation examines three research questions:

Research question 1: How to model automated trading systems with a variable DOA using CWA?

Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

The following two sets of hypotheses were developed to examine research question 2 and 3 respectively.

3.5.1 Research Hypotheses for Examining Performance

H1: The participants should achieve better performance with ecological displays in comparison to conventional displays.

Hypothesis 1 will be examined in several aspects:

H1a: *The participants should perform better with moderate-ecological than with moderate-conventional, and should perform better with high-ecological than with high-conventional*. Hypothesis 1a examined what demand each DOA configuration placed upon the participants. The participants faced unanticipated situations in both DOA configurations. Therefore, the ecological displays were expected to improve the task performance relative to the conventional displays similar to Lau et al.'s findings (2008). Since the participants performed distinct tasks in moderate DOA and high DOA scenarios and the ecological displays were designed to provide different support for specific DOAs, this hypothesis would not generalize the overall difference between display type and DOA. Instead, two comparisons were made between moderate-ecological and moderate-ecological, and high-ecological and high-ecological.

H1b: The participants' SA would be higher with moderate-ecological than with moderate-conventional. The participants' SA would be higher with high-ecological than with high-conventional. There might also be some evidence in the eye-tracking measure data that can support the SA results. According to Burns et al. (2008), ecological displays should improve the participants' SA, particularly in unanticipated situations.

H1c: The participants would neither perceive higher workload with moderate-ecological than with moderate-conventional. The participants would neither perceive higher workload with high-ecological than with high-conventional. According to the literature, the ecological displays should achieve performance advantages without imposing more workload (Lau, Jamieson, et al., 2008).

3.5.2 Research Hypotheses for Examining Risk Preference

H2: The participants could have different risk preferences with ecological displays in comparison to conventional displays.

A breakdown of hypothesis 2 is presented as follows:

H2a: The participants' fourfold experience-based choice in an automated trading environment would be different from that as identified in McAndrew and Gore's observations (2013), under the influence of scenario type (moderate-conventional, moderate-ecological, high-conventional or high-

ecological). If hypothesis 2a is supported, the pattern of choice in this experiment would be compared to those in McAndrew and Gore's study, Hertwig and Erev's study (2009) and the prospect theory.

H2b: With moderate DOA, mean portfolio's size (S_p) with the ecological displays would be larger with that with the conventional displays.

H2c: In one or more system state, the participants would be leaning towards riskier actions in a guaranteed profiting situation with the ecological displays with that with the conventional displays.

H2d: In one or more system state, the participants would be leaning towards riskier actions in a guaranteed losing situation with the ecological displays with that with the conventional displays.

If hypothesis 2b, 2c and 2d can be supported, it may be possible to elaborate the finding of Borst et al. (2015) that operators occasionally make risky decisions with ecological displays. A warranted conclusion would be required.

3.6 Results

3.6.1 Conventions

The data analysis process followed several conventions:

First, all results were set to three decimal places. A *statistically significant* result is a result with $p < .05$. A result with $p \geqslant .05$ is *not a statistically significant* result.

Second, according to Cohen's (1992), In paired t tests, the effect size (d) used .2, .5 and .8 for *small*, *medium* and *large* effects respectively. In repeated measures analysis of variance (ANOVA) tests, the effect size (η^2) used .01, .06 and .14 for *small*, *medium* and *large* effects respectively (Cohen, 1992). For Wilcoxon signed rank tests, the effect size (r) was calculated by dividing z by the square root of the number of subjects (Field, 2005). The calculations were then compared to the r thresholds of .1, .3 and .5 for *small*, *medium* and *large* effects respectively. There is no generally agreed effect size measure for Friedman's tests. Effect sizes were calculated in post hoc tests (using Wilcoxon Signed Rank tests for pairwise comparisons).

Third, for repeated measures ANOVA test and paired *t* test, the arithmetic mean (*M*) was used to determine whether a statistically significant difference existed between the means of multiple independent groups. In the case of non-parametric tests, including Friedman's test and Wilcoxon signed rank test, median (*Mdn*) was used to determine the statistical significance. Skewed distributions would be common in financial data and these distributions, medians provide a measure that is more robust to outlier values than arithmetic means.

3.6.2 Data Analysis Script

Customized scripts for cleaning and statistically analyzing the data were developed with R 3.3.2 and RStudio. There were approximately 25,000 lines of code, and it took approximately eight person-months to develop the scripts.

3.6.3 Summary of Results

In the following subsections, the author presents the data analysis results which are sorted based on the types of the dependent variables.

3.6.3.1 Task performance

End of scenario RPL (moderate DOA)

End of scenario *RPL* was collected at the end of each moderate DOA scenario the participant experienced. One participant experienced technical difficulties and their data were subsequently excluded from the analysis. There were in total 46 end of scenario *RPLs* and the data were divided into two groups denoting the two scenario types (moderate-conventional and moderate-ecological), $N = 23$. The assumption of normality was violated, $ps < .05$ and a non-parametric test was performed instead.

Results of a Wilcoxon signed-rank test showed that the scenario type did not significantly affect the participants' trading performance in the moderate-conventional and moderate-ecological scenarios, $p > .05$. This result aligned with the data collected through the training scenarios (moderate-conventional: *Mdn* = \$108; moderate-ecological: *Mdn* = \$51.5). No statistically significant difference, $p > .05$).

Table 15 summarizes the descriptive statistics. Empirically, the participants achieved a slightly less end of scenario *RPL* in the moderate-ecological scenarios than in the moderate-conventional scenarios.

Table 15. Summary of End of Scenario *RPLs* (Experiment 1).

Scenario Type	End of Scenario *RPL*			
	Mdn (Middle Most)	*M* (Arithmetic Mean)	*Mo* (Most Frequent)	*SD* (Deviation)
Moderate-conventional	\$35	\$39.0	\$113	132.4
Moderate-ecological	\$30.5	\$7.1	\$83	210.6

Since the moderate DOA task was generally flexible, it might be influenced by mediating factors that were used to facilitate the within-subject design. For example, a follow-up analysis was performed to understand whether the different characteristics of the two financial market data sets could influence the scenario type effect on end of scenario *RPL*. The data were divided into two groups by financial market data (D1: 120 timestamps; D2: 180 timestamps). The analysis was only performed empirically due to the unequal sample sizes of the two groups. Results showed that with D1, end of scenario *RPL* was lower with the moderate-conventional scenario (*Mdn* = -\$18) than with the moderate-ecological scenario (*Mdn* = \$78). However, in the case of D2, a reverse pattern was found (moderate-conventional: *Mdn* = \$48; moderate-ecological: *Mdn* = -\$80) that aligned with the initial analysis results. Further analysis included only the first 120 timestamps of D2 in order to control the duration. The reverse pattern was still consistent (moderate-conventional: *Mdn* = \$35.5; moderate-ecological: *Mdn* = -\$89.3), indicating that trend or volatility difference may have contributed to the empirically different patterns with the two financial market data sets. Thus, a comparison between of the characteristics of D1 and D2 was included as follows.

Several indicators were used for comparing the financial market characteristics. Note that open-close spread is similar to high-low spread which was initially used to determine what financial market data to be used in this experiment. Open-close spread describes the overall market trend (up-trend or down-trend) of a given time period, whereas high-low spread may be related to the volatility of the market. It can be seen from Table 16 that D1 and D2 have similar volatilities, indicating by their similar high-low spreads and *SD*s. D1 has a up-trend, whereas D2 has a down-trend during the first 180 timestamps followed by a recovery in the last 60 time stamps. The different trends are evident in Figure 28 and 29, which show the trends of D1 and D2 (with a line dividing the first 120 timestamps and the last 60 timestamps, for clarity).

Table 16. Characteristics of Financial Market Data (Experiment 1).

Indicator	D1 (120 timestamps)	D2	
		(first 120 timestamps)	(180 timestamps)
Open-close spread ($)	-.220 (up-trend)	.240 (down-trend)	.100 (down-trend then recover)
High-low spread ($)	.330	.410	.410
SD	.095	.096	.096

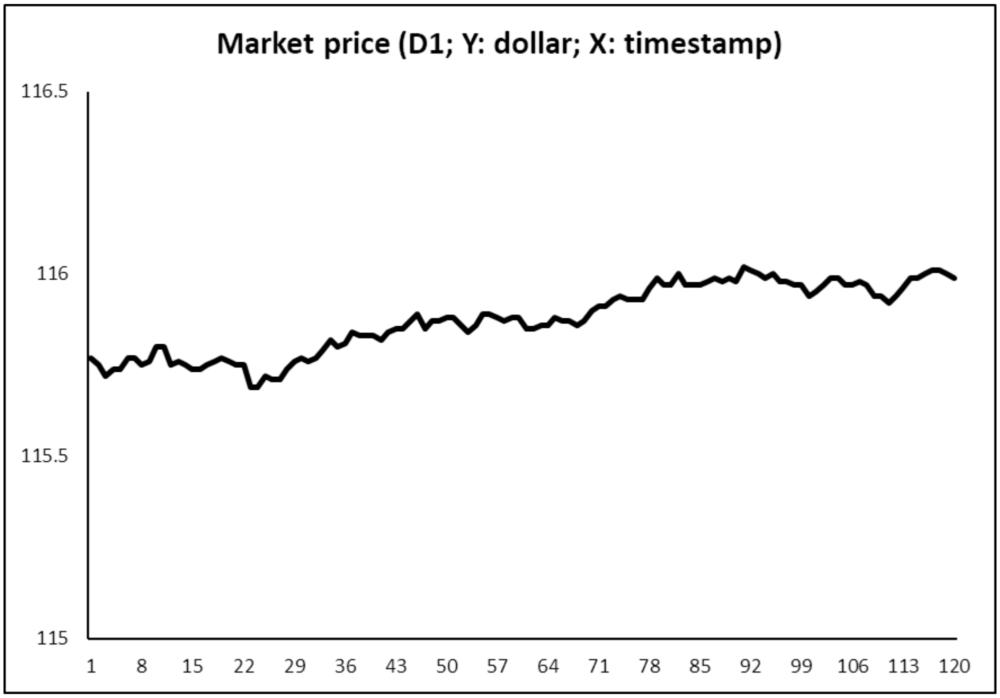


Figure 28. Market price (with D1, Experiment 1).

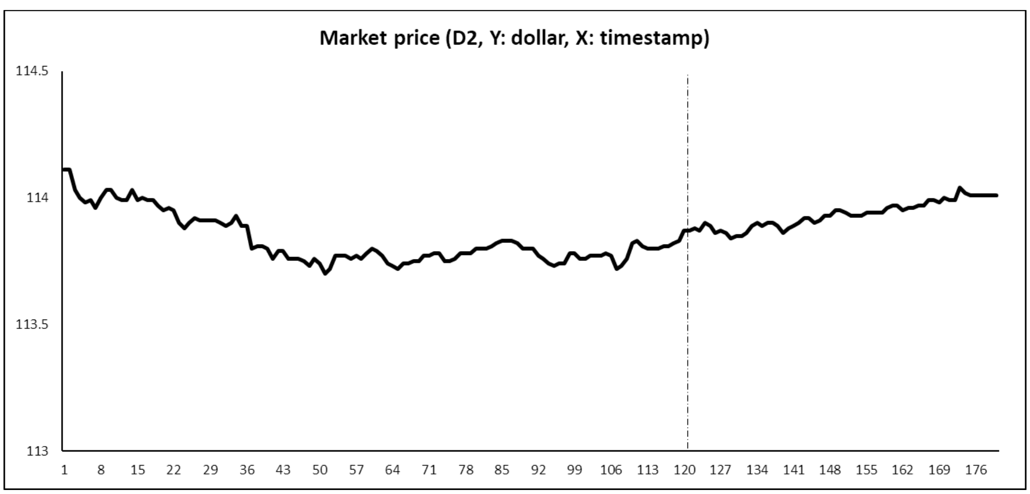


Figure 29. Market price (with D2, Experiment 1).

Fault detection accuracy (high DOA)

All participants completed the fault detection task while experiencing the high-conventional and the high-ecological scenarios. Fault detection accuracy was calculated per scenario (high-conventional or high-ecological) for each participant and all results were subsequently compared between the two scenario types, $N = 24$. There were in total 48 fault detection accuracies. The assumption of normality was violated, $ps < .05$. Wilcoxon signed rank test results suggested that the scenario type effect was not significant in the fault detection performance measure, $p > .05$.

Non-parametric tests used the median to compare multiple within-subject groups. To provide a full pattern in the data, Table 17 presents a multiple central tendency measures and provides a full picture of the simulation data. It can be seen in this table that the participants performed the fault detection task equally well with both the conventional display and the ecological display.

Table 17. Summary of Fault Detection Accuracies (Experiment 1).

Scenario Type	Fault Detection Accuracy			
	Mdn (Middle Most)	*M* (Arithmetic Mean)	*Mo* (Most Frequent)	*SD* (Deviation)
Moderate-conventional	75.0%	73.2%	100%	.289
Moderate-ecological	100.0%	87.2%	100%	.213

3.6.3.2 Situation awareness

The participants' responses to the SA queries were compared to the actual situations as simulated by AUTRASS. Due to limited expert support in this analysis stage, the actual situations were reconstructed with the quantitative simulation data recorded in log files, not with the video recordings of the simulation screen. The grading of the responses was binary (0 or 1).

The participants' responses to one SA query ("in the last 5 seconds, what was the most recent trade?") were excluded from the data analysis, because the answer choices did not include a third option for the participants to report no trading execution which may happen during the window of the past 5 seconds. Special cases in the participants' responses were treated separately in the scoring process. For example, if the participant expressed no opinion toward in response (e.g., "not sure"), the response would be assigned score zero. If the participant also expressed a guess in the response (e.g., "not sure but I guess there will be a sell trade"), the response would be scored according to the participant's guess.

Table 18 provides a list of all SA queries, answer choices and the rubrics for scoring.

Table 18. Summary of SA Queries and Scoring Rubrics (Experiment 1).

SA Level	SA Query	Answer Choice	To Achieve a Score of 1, the Participant Must Have
1	"In the last 5 seconds, the market close price has gone:"	Up, down or flat	Correctly compared P_m at t - 1 to t.
	"In the last 5 seconds, the slower moving average curve (yellow) has gone:"	Up, down or flat	Correctly compare SMA_{slow} at $t - 1$ to t.
	"In the last 5 seconds, the faster moving average curve (purple) has gone:"	Up, down or flat	Correctly compared SMA_{fast} at $t - 1$ to t.
	"In the last 5 seconds, was there a crossover of the two moving average curves?"	Yes or no	Correctly identify whether SMA_{slow} crossed over SMA_{fast} during t - 1 to t.
	"In the last 5 seconds, what was the most recent trade?*"	Buy, sell	Correctly identified the trade (buying, selling or neither) during t - 1 to t.
2	"What is happening with the unrealized profit and loss in your portfolio? Why?"	(Open-ended)	Correctly identify whether UPL was positive or negative at t and the direction of its movement from $t - 1$ to t.
	"Is there a buying opportunity? Why?"		Correctly identified a buying opportunity if $P_p > P_m$ at the timestamp t.
	"Is there a selling opportunity? Why?"		Correctly identified a selling opportunity if $P_p < P_m$ and $Q_p > 500$ at the timestamp t.
	"What is happening with the quantity in your portfolio? Why?"		Correctly compared Q_p at $t - 1$ with Q_p at t.
	"Has the most recent trading execution made any profit? Why?"	(Open-ended)	Correctly compared P_m with P_p, if there was a selling trade between $t - 1$ and t, or explained there was a buying trade or no trade.
3	"What will happen to the market close price in the next 5 seconds? Why?"	(Open-ended)	Correctly compared P_m at t and P_m at $t + 1$ (predicted).

SA Level	SA Query	Answer Choice	To Achieve a Score of 1, the Participant Must Have
	"Do you think there will be a buying opportunity in the next 5 seconds? Why?"		Correctly predicted a buying opportunity if $P_m < P_p$ at $t + 1$, or correctly explained why there would be no buying opportunity in other cases.
	"Do you think there will be a selling opportunity in the next 5 seconds? Why?"		Correctly predicted a selling opportunity if $P_m > P_p$ and $Q_p > 500$ at $t + 1$, or correctly explained why there would be no buying execution otherwise.
	"Do you think there will be a buying execution in the next 5 seconds? Why?"		Correctly predicted a buying execution if Q_p increased during t and $t + 1$, or correctly explained why there would be no buying execution otherwise.
	"Do you think there will be a selling execution in the next 5 seconds? Why?"		Correctly predicted a buying execution if Q_p decreased during t and $t + 1$, or explained there would be no selling execution otherwise.
	"What will be the status of the quantity in your portfolio in the next 5 seconds? Why?"		Correctly compared Q_p at t and Q_p at $t + 1$.

* Excluded from the data analysis due to flawed question design.

Twenty-three participants successfully completed the SA queries. One participant encountered technical difficulties, and their responses were not correctly logged. For each measurement scenario, the mean score for each SA level (1, 2 or 3) was calculated by averaging the SA ratings for all queries on that level at both pauses. Responses to the excluded level 1 SA query were flagged as unpresentable and the mean score for SA level 1 was decided by the responses to the other three level 1 SA queries presented during the two pauses. The mean scores were normalized to the 0 to 1 range. There were in total 276 mean scores. The mean scores were then divided into 4 (scenario type: moderate conventional, moderate-ecological, high-conventional or high-ecological) × 3 (SA level: 1, 2 or 3) groups, $N = 23$. The assumption of normality was violated in most groups, *ps* < .05. Non-parametric tests were performed to evaluate the individual effects of scenario type and SA level.

Results of the Friedman's test showed that there was a significant scenario type effect, χ^2 = 9.560, *p* = .023. Figure 30 is a box and whisker plot that shows this effect. It is evident that the participants had higher overall SA when they were experiencing the ecological in moderate scenario than the ecological in high scenario, *p* = .043. The participants also had higher SA in the conventional in moderate scenario than in the ecological in high scenario, *p* = .040.

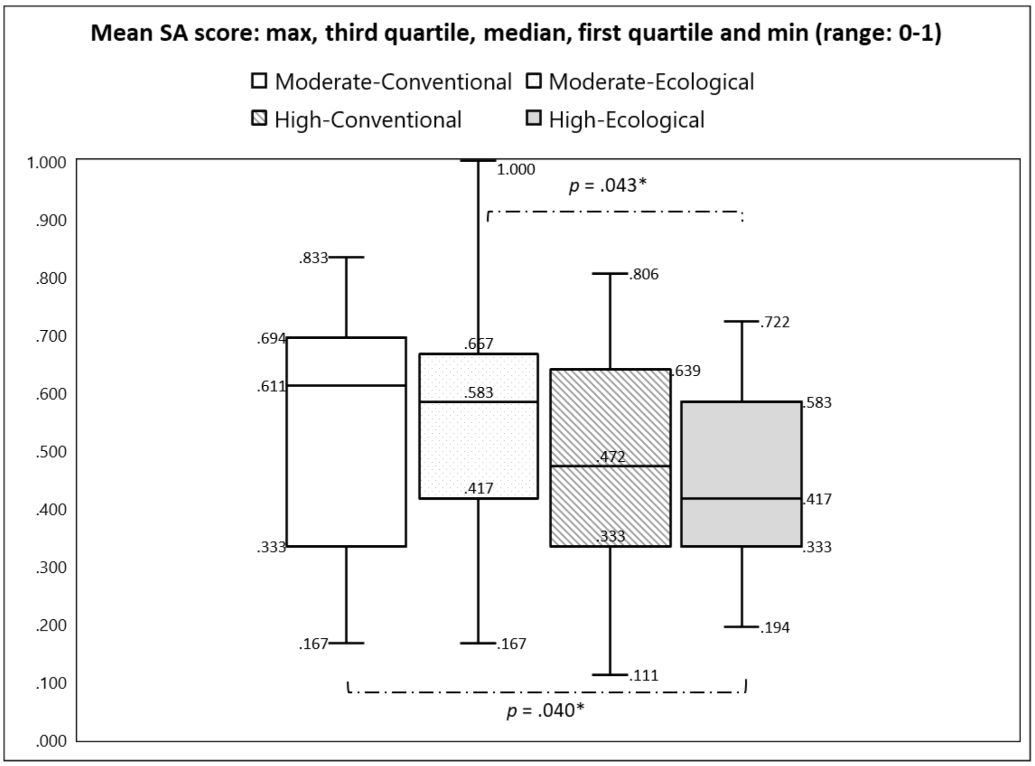


Figure 30. Mean SA score: DOA effect (Experiment 1).

The SA level effect was not significant, *ps* > .05. These results suggested that while the participants' SA was generally moderate (*Mdn* =.500), the ecological displays did not further improve the participants' SA in the current experimental setting. Median statistics per scenario type and per SA level are summarized in Table 19 for references.

Table 19. Summary of Mean SA Scores (Experiment 1).

Scenario Type	Per Scenario Type (*Mdn*)	Per SA Level (*Mdn*)		
		1	2	3
Moderate-conventional	.611	.667	.667	.533
Moderate-ecological	.583	.667	.500	.609
High-conventional	.472	.667	.333	.435
High-ecological	.417	.333	.333	.413

3.6.3.3 Eye tracking

Eye-tracking data were analyzed to help with the interpretation of the SA findings. Total dwell time data for eight participants were missing in some scenarios due to technical reasons and were subsequently excluded from the analysis.

Total dwell time (market price AOI, portfolio AOI and trading history DOA)

Total dwell time data for AOIs that were consistent with the type of display were divided to 4 (scenario type: moderate-conventional, moderate-ecological, high-conventional or high-ecological) × 3 (AOI: market, portfolio, trading history) groups. The assumption of normality was violated in several groups, $ps < .05$, $N = 16$. The total dwell time data were applied a log 10 transformation with zero data handled in a way similar to Bartlett's log (x+1) approach (1947) to normalize the distribution and were consequently submitted to a 4 × 3 repeated measures ANOVA. The assumption of Sphericity was violated with all main and interaction effects. A Greenhouse-Geisser correction was applied. There was a significant scenario type simple main effect, $F(2.111, 45) = 3.839$, $p = .023$, $\eta^2 = .038$ (medium effect) and a significant AOI simple main effect, $F(1.661, 30) = 7.666$, $p = .003$, $\eta^2 = .124$ (large effect).All other effects were not significant, $p > .05$. Post hoc tests were performed to analyze the two simple main effects.

For the market price AOI and the trading history AOI, there was no statistical significant difference between any two scenario types, $p > .05$. For the portfolio AOI, as highlighted in Figure 31, the participants spent significantly longer total dwell time when they experienced the moderate-conventional scenario than the moderate-ecological scenario, $p = .046$. A longer dwell time may correlate with a poorer SA or higher informativeness of a specific AOI (Holmqvist et al., 2011). The moderate-ecological display contained more information pertaining to the relationship between the market, the portfolio and the executions (i.e., market-portfolio-execution visualization), which was instrumental to portfolio management. As a result, the portfolio AOI has become less informative with moderate-ecological than with moderate-conventional. It is also possible that SA specific to portfolio management in the moderate DOA scenarios was improved while using the ecological displays over the conventional displays.

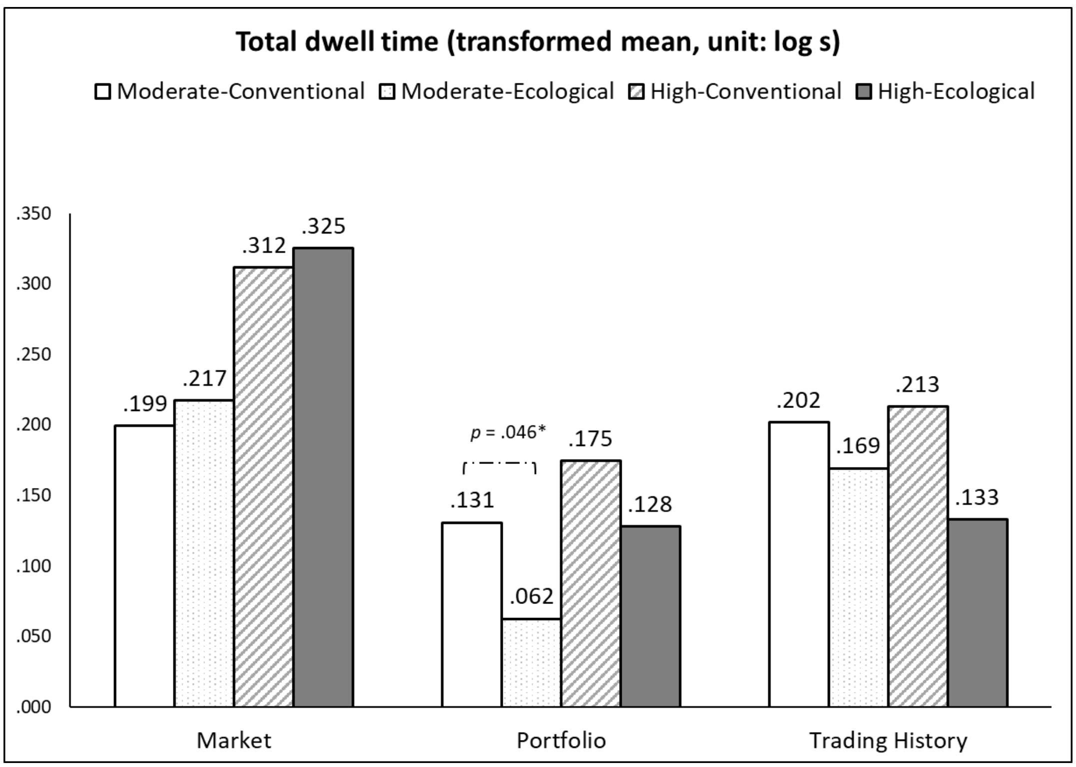


Figure 31. Total dwell time in common AOIs: Scenario type effect (Experiment 1).

For the AOI effect, total dwell time was significantly longer in the market AOI than in the portfolio AOI in the case of moderate-ecological (p = .011) and high-ecological (p = .016). Total dwell time was significantly longer in the portfolio AOI than in the trading history AOI during the moderate-ecological scenario (p = .041). Total dwell time was significantly longer in the market AOI than in the trading history AOI only with the high-ecological scenario (p = .032). The two ecological display scenarios included the market-portfolio-execution visualization which could facilitate more effective portfolio management and allow the participants to focus on the market panel, which was expected to be an important source of uncertainties in this simulated environment. This conclusion is supported by the empirically longer total dwell time between moderate-ecological and moderate-conventional and between high-ecological and high-conventional. The empirical results also showed that total dwell time on the market AOI might be longer in the high DOA scenarios than in the moderate DOA scenarios, and such may be associated with the different tasks performed in the two cases. Total dwell time on the trading history AOI was also lower with the high-ecological scenario, suggesting that executions made by the trading algorithm were provided by the states-task visualization, leading to less attention allocation to the trading history panel. A summary of total dwell time for the common AOIs is presented in Table 20.

Table 20. Summary of Mean Total Dwell Time for the Common AOIs (Experiment 1).

Scenario Type	Area of Interest (*M*, *SD*, unit: log s)		
	Market	Portfolio	Trading History
Moderate-conventional	.199 (.176)	.131 (.114)	.202 (.146)
Moderate-ecological	.217 (.167)	.062 (.081)	.169 (.166)
High-conventional	.312 (.186)	.175 (.176)	.213 (.180)
High-ecological	.325 (.199)	.128 (.118)	.133 (.132)

Total dwell time (market-portfolio-execution AOI)

Total dwell time on the market-portfolio-execution AOI with the ecological displays were log transformed and submitted to a paired *t* test for examining the scenario effect (scenario type: moderate-ecological or high-ecological), $N = 20$. No statistically significant difference in the total dwell time was found between the two configurations, $p < .05$, as shown in Table 21. The empirical difference was likely to be a result of different task requirements with moderate DOA and with high DOA. The market-portfolio-execution visualization might be less informative to the monitoring and automation fault detection task.

Table 21. Summary of Mean Total Dwell Time for the Market-Portfolio-Execution AOI (Experiment 1).

Scenario Type	Area of Interest (*M*, *SD*, unit: log s)
	Market-Portfolio-Execution
Moderate-ecological	.157 (.156)
High-ecological	.093 (.102)

Total dwell time (states-task AOI)

Total dwell time on the states-task AOI in the high DOA scenarios was log transformed and examined using a paired *t* test on the expected effect of the states-task visualization (scenario type: high conventional or high ecological), $N = 20$. This effect was not significant, $p > .05$, as shown in Table 22. Empirically, the states-task AOI drew more participants' attentions due to the more information provided by the visualization displayed in this area.

Table 22. Summary of Mean Total Dwell Time for the States-Task AOI (Experiment 1).

Scenario Type	Area of Interest (*M*, *SD*, unit: log s)
	States-Task
High-conventional	.169 (.213)
High-ecological	.208 (.155)

3.6.3.4 Workload

Arithmetic means of all participants' unweighted NASA TLX ratings on all subscales after they completed each of the four scenarios were calculated and subsequently submitted to a 4 (scenario type: moderate conventional, moderate ecological, high conventional or high ecological) × 6 (NASA TLX subscale: mental workload, physical workload, temporal workload, performance, frustration or effort) repeated measures ANOVA. This repeated measures ANOVA provided a robust estimation as the assumption of normality violated in several groups, *ps* < .05. The assumption of sphericity was violated with the NASA TLX subscale, $p < .05$. A Greenhouse-Geisser correction was applied. There was a significant effect of scenario type, $F(2.142, 69) = 3.390$, $p = .023$, $\eta^2 = .032$ (small effect). Post hoc tests were performed to make pairwise comparisons using Bonferroni corrections. However, none of the pairwise difference was significant, *ps* > .05. Emprical results presented in Figure 32 suggested that the unweighted ratings for the two moderate DOA scenarios were greater. Compared to the high DOA scenarios, the participants possibly endured higher workload in the moderate DOA scenarios in which they were requested to perform a more demanding task. The NASA TLX subscale effect was also significant, $F(2.960, 115) = 9.002$, $p < .001$, $\eta^2 = 0.126$ (medium effect). All other effects were not significant, *ps* > .05.

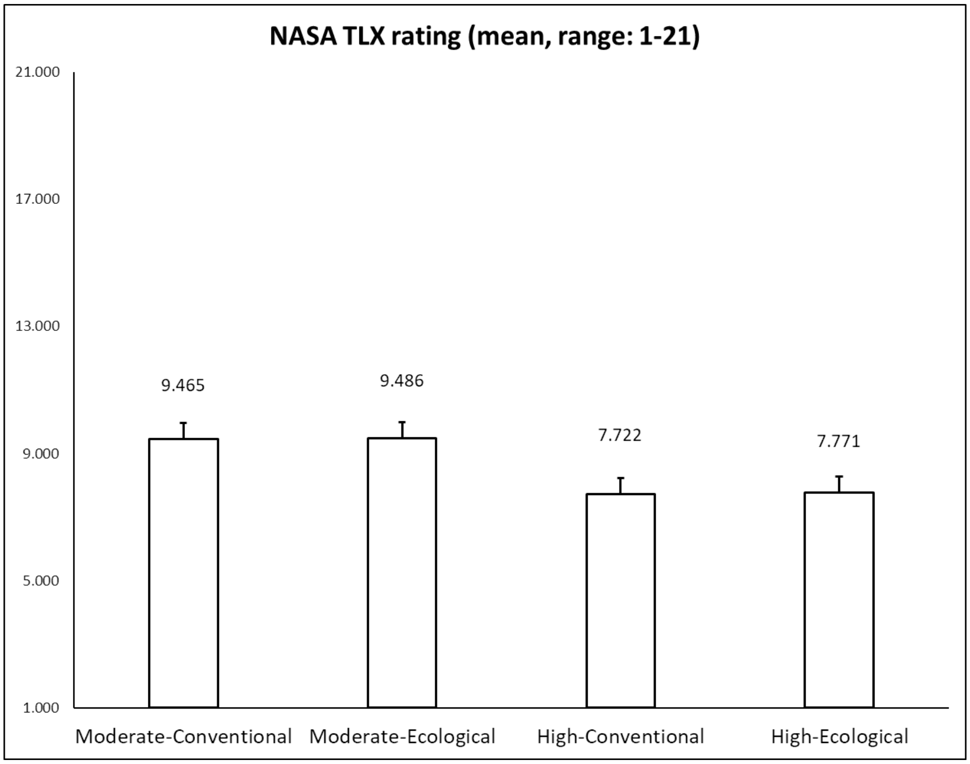


Figure 32. NASA TLX rating: Scenario Type effect (Experiment 1).

3.6.4 Risk Preference

3.6.4.1 Fourfold pattern of preferences

Each participant provided one response to each of the four questions on the risk preference questionnaire, after the completion of each scenario. All 24 participants' responses were recorded. There were in total 384 responses. The number of risk-seeking choices was compared to the number of risk-aversion choices for each scenario type (moderate-conventional, moderate-ecological, high-conventional or high-ecological), N = 24. Four McNemar's exact tests were performed on each question to examine whether the DOA and the display could affect the preference choices (i.e., comparing the responses between the two DOA configurations while controlling the display independent variable, and vice versa). The McNemar's test, similar to the Pearson's chi-squared test, can be used to evaluate the consistency in participants' responses across two variables in a within-subject experimental design.

Neither the DOA effect nor the display effect was significant, *ps* > .05, suggesting that there was a consensus on the preference choices in each question among all DOA and display groups. It can be concluded from Table 23 that the participants were risk-seeking with small-probability gains and losses, and risk-aversion with medium-and large-probability losses. For medium- and large-probability losses, there was no consistent pattern in the participant's choices.

Table 23. Fourfold Patterns of Preferences: Display and DOA Effects (Experiment 1).

Medium- and large-probability gains

	Moderate-Conventional	Moderate-Ecological	High-Conventional	High-Ecological
Risk seeking	<u>14</u>	<u>14</u>	<u>17</u>	<u>16</u>
Risk aversion	10	10	7	8

Small-probability losses

	Moderate-Conventional	Moderate-Ecological	High-Conventional	High-Ecological
Risk seeking	<u>15</u>	<u>17</u>	<u>19</u>	<u>18</u>
Risk aversion	9	7	5	6

Small-probability gains

	Moderate-Conventional	Moderate-Ecological	High-Conventional	High-Ecological
Risk seeking	8	6	6	9
Risk aversion	<u>16</u>	<u>18</u>	<u>18</u>	<u>15</u>

Medium- and large-probability losses

	Moderate-Conventional	Moderate-Ecological	High-Conventional	High-Ecological
Risk seeking	11	12	8	<u>14</u>
Risk aversion	<u>13</u>	12	<u>16</u>	10

3.6.4.2 Mean portfolio's size (moderate DOA)

Mean portfolio's size was calculated by averaging the portfolio's size of all timestamps for each scenario (moderate-conventional or moderate-ecological). The data were divided to 2 (scenario type: moderate-conventional or moderate-ecological) × 3 (system state: profiting, neutral or losing) groups for analyzing the mean position sizes, $N = 23$. The assumption of normality was violated in all groups, $ps < .05$. The data were subsequently analyzed using a non-parametric test. Results of the Wilcoxon signed rank test showed that the system state effect was not significant, $p > .05$ (profiting: *Mdn* = 625.3 shares; neutral: *Mdn* = 707.5 shares; losing: *Mdn* = 1,034.6 shares). There was a significant scenario type effect, $z = 2.464$, $p = .014$, $r = .363$ (medium effect), suggesting that the participants held a significantly larger portfolio in the moderate-ecological scenario (*Mdn* = 904.3 shares) than in the moderate-conventional scenario (*Mdn* = 618.0 shares). The median difference and other data attributes, including max, min and quartiles, are presented in a box and whisker plot as in Figure 33. So far, it is not clear whether the largest data point in the ecological display data (12,092.5 shares) should be flagged as an outlier. The participants were provided with unlimited buying power so that building a large portfolio was technically possible. To be sure, a follow-up Wilcoxon signed test was performed that excluded this data point. The difference between the moderate ecological scenario and the moderate conventional scenario was still significant, $z = 2.256$, $p = .024$, $r = .333$ (medium effect).

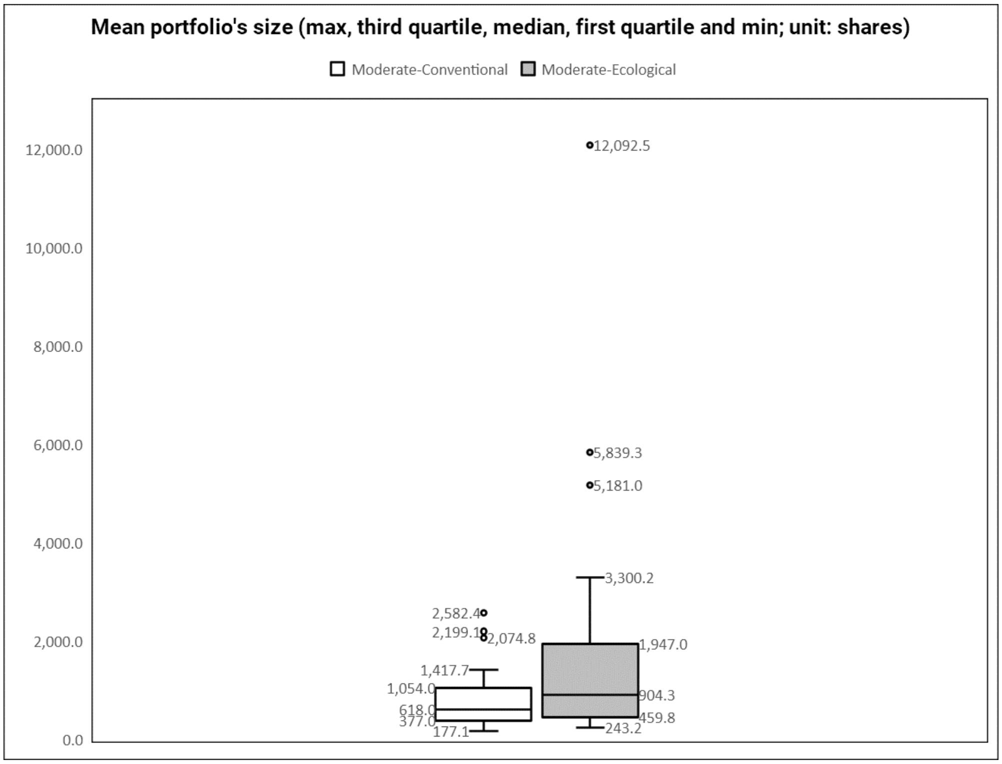
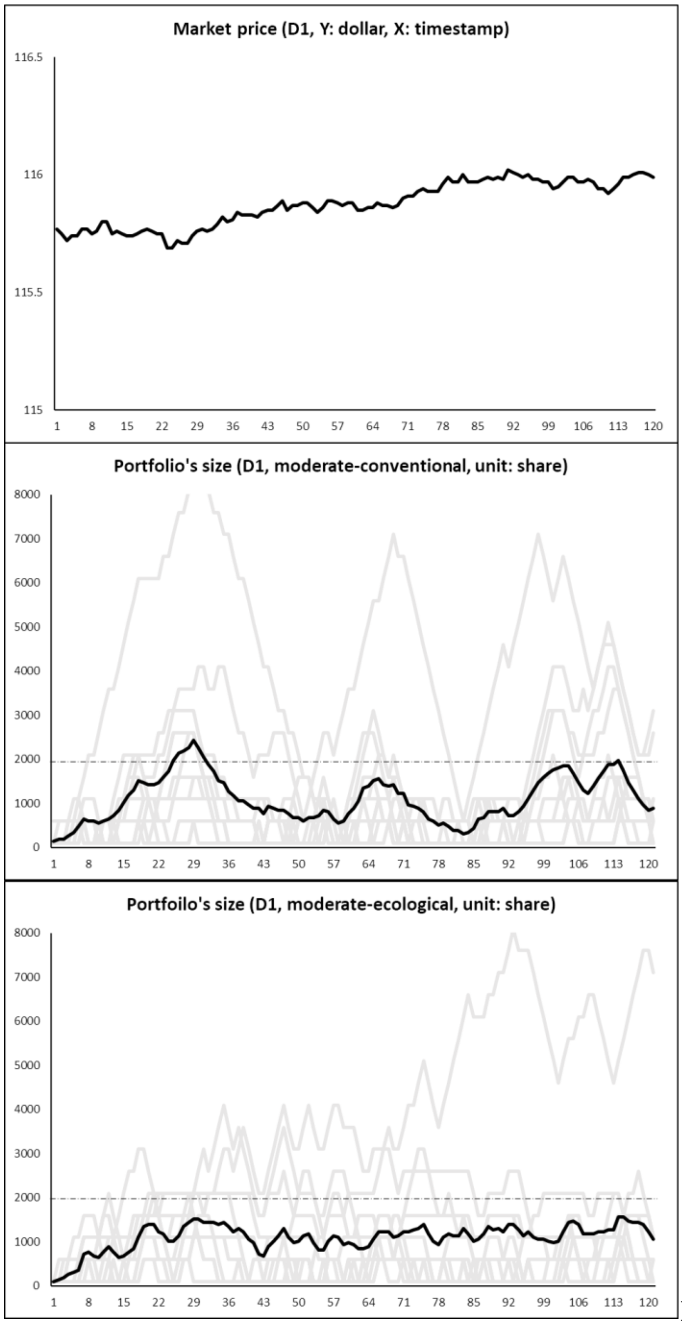


Figure 33. Mean portfolio's size: Scenario type effect (Experiment 1).

Since the two financial market data sets had distinct characteristics, a further analysis was conducted to understand how the participants developed and managed their portfolio in the two simulated financial markets. Note that this analysis was empirical because the two financial market sets were randomly (unequally) assigned to participants and scenarios. Figure 34 and 35 demonstrated the different portfolio management strategies the participants used with the moderate-conventional scenario and the moderate-ecological scenario. The two financial market data sets, D1 and D2, showed distinct portfolio management behaviours. For D1 (Figure 34), with the market moving up-trend, the participants abruptly sold off their shares with moderate-conventional. On the other hand, with moderate-ecological, the participants adopted a more consistent portfolio management strategy and performed less "panic" sell-offs. For D2 (Figure 35), the market moved down-trend during the first 120 timestamps. The behaviour patterns between moderate-conventional and moderate-ecological were very different. With moderate-conventional, the participants generally held a more risk-averse portfolio management strategy, whereas with moderate-ecological a much larger portfolio has been maintained throughout the scenario. These results are empirical, and more work is needed.



D

Figure 34. Portfolio's size: moderate-conventional versus moderate-ecological (with D1, Experiment 1).

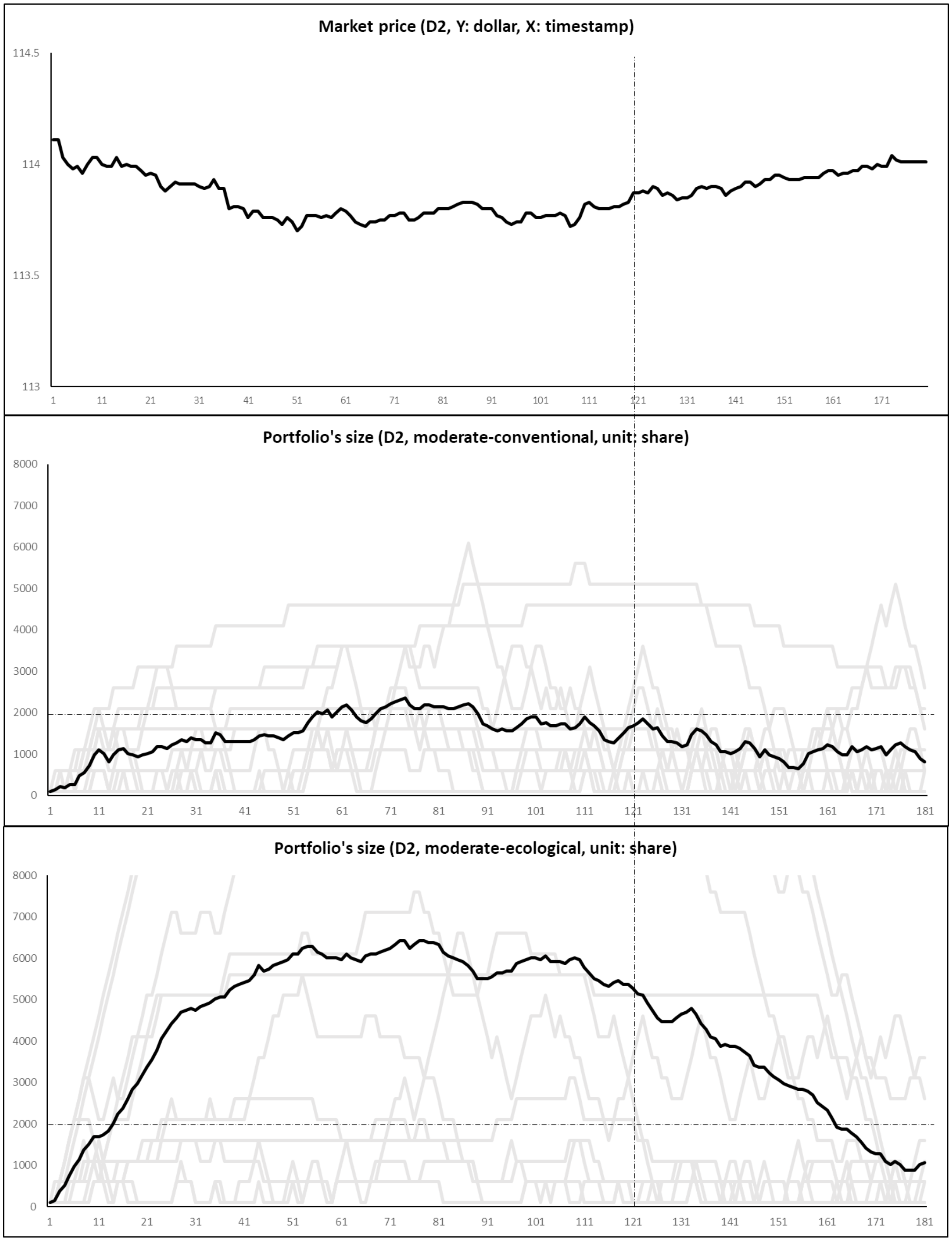


Figure 35. Portfolio’s size: moderate-conventional versus moderate-ecological (with D2, Experiment 1).

3.6.4.3 Decision preference in a guaranteed profiting situation (Low DOA)

The percentage of performing each execution (selling, holding or buying) in a guaranteed profiting situation reflects the likelihood of being risk-aversion, moderate risk-seeking or high risk-seeking when the participants faced a high probability prospect gain. It can be calculated by counting the number of each execution made then dividing the sum by the total number of guaranteed profiting situations.

First, the total number of guaranteed profiting situations were calculated for each moderate conventional scenario and high ecological scenario, $N = 23$. The author then counted the number of each execution (risk level: risk-aversion, moderate risk-seeking or high risk-seeking). The percentage of execution that was associated with each risk level was calculated by dividing the number of each execution by the total number of guaranteed profiting. The statistical model for analyzing the percentage of execution was a 2 (scenario type: moderate conventional or moderate ecological) × 3 (system state: profiting, neutral or losing) × 3 (risk level: risk-averse, moderate risk-seeking or high risk-seeking) repeated measures design. It has been observed in some scenario the participant did not experienced any guaranteed profiting opportunities during one or more system states, given the system states was dynamically distributed. In these cases, the percentage of execution at each risk level was manually set as 1/3, representing a neutral risk preference among the three risk levels. The assumption of normality was violated in all 2 × 3 × 3 groups, $ps < .05$. As a result, non-parametric tests were performed. The scenario type effect was not significant, $p > .05$ (moderate conventional: *Mdn* = 33.3%; moderate ecological: *Mdn* = 33.3%). The system state effect was also not significant, $p > .05$ (profiting: *Mdn* = 33.3%; neutral: *Mdn* = 33.3%; losing: *Mdn* = 33.3%). Results of the Friedman's test showed that the risk level effect was significant, $\chi^2 = 9.478$, $p = .009$ (risk-aversion: *Mdn* = 22.5%; moderate risk-seeking: *Mdn* = 53.7%; high risk-seeking: *Mdn* = 22.0%), suggesting that the participants achieved different percentages for each execution, as shown in Figure 36. Post hoc tests indicated that the participants were more likely to make a moderate risk-seeking decision than a risk-averse decision (z = 2.677, p = .007, r = .395, a median effect) and a high risk-seeking decision (z = 3.346, p < .001, r = .493, a medium effect). All other effects were not significant, p > .05.

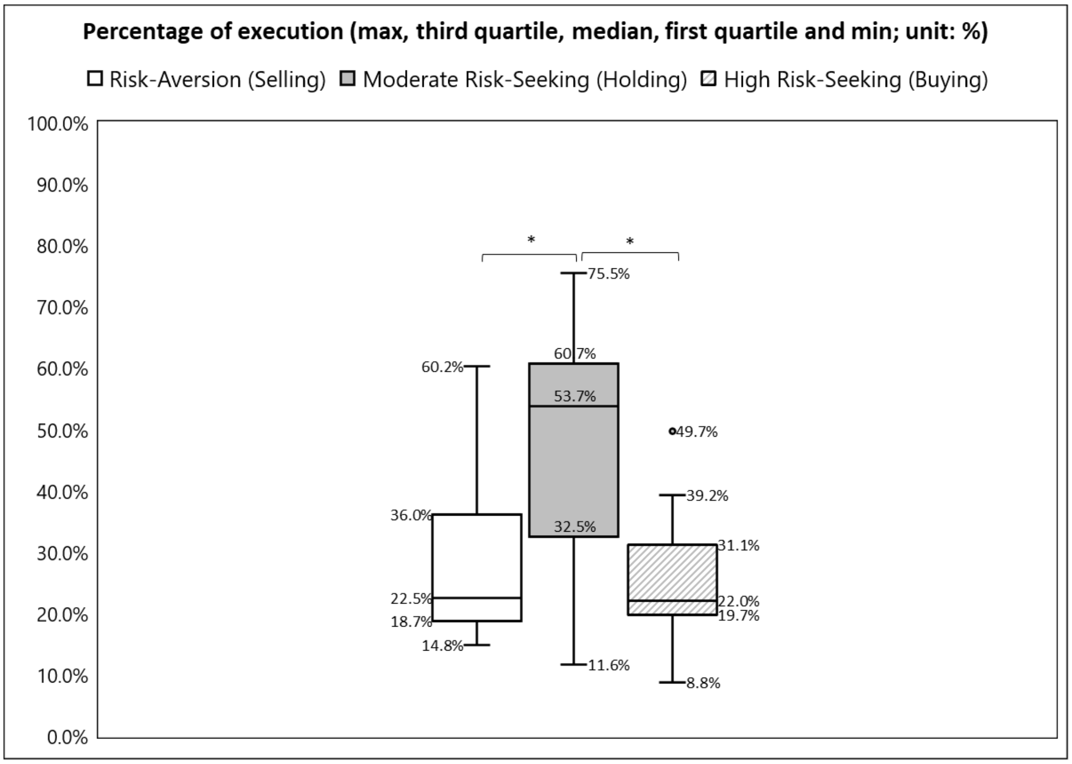


Figure 36. Percentage of execution in a guaranteed profiting situation: Risk level effect (Experiment 1).

Since the risk level × scenario type interaction cannot be examined through a non-parametric test, the simulation data were collapsed into categories representing the three risk levels, with each category further examined with the same statistical model. The statistical model was fine tuned to evaluate whether the system state could influence how the participants chose between the three types of executions. The statistical model used a 2 (scenario type: moderate conventional or moderate ecological)

× 3 (system states: profiting, neutral or losing) with-subject design, $N = 23$. The assumption of normality was violated in all groups, $ps < .05$. Non-parametric tests wereperformed for each risk level:

1. For the risk-averse risk level, the percentage of selling executions was analyzed. The scenario type effect was not significant, $p > .05$ (moderate-conventional: *Mdn* = 24.7%; moderate-ecological: *Mdn* = 24.9%). The system state effect was also not significant, $p > .05$ (profiting: *Mdn* = 22.8%; neutral: *Mdn* = 25.0%; losing: *Mdn* = 24.4%);

2. For the moderate-risk seeking risk level, the percentage of holding executions made by the participants was analyzed. The scenario type effect was not significant, $p > .05$ (moderate-conventional: *Mdn* = 49.1%; moderate-ecological: *Mdn* = 49.2%). The system state effect was also not significant, $p > .05$ (profiting: *Mdn* = 55.6%; neutral: *Mdn* = 33.3%; losing: *Mdn* = 44.2%). For the high risk-seeking risk level, the percentage of buying executions made by the participants was analyzed. The scenario type effect was not significant, $p > .05$ (moderate conventional: *Mdn* = 24.9%; moderate ecological: *Mdn* = 23.4%). The significant system state effect was also not significant, $p > .05$ (profiting: *Mdn* = 21.7%; neutral: *Mdn* = 33.3%; losing: *Mdn* = 19.6%).

Table 24 summarizes all median percentages of execution that have been calculated.

Table 24. Summary of Percentages of Executions in a Guaranteed Profiting Situation (Experiment 1).

Scenario Type	System State	Percentage of Execution (*Mdn*)		
		Risk-Aversion	Moderate Risk-Seeking	High Risk-Seeking
Moderate-conventional	Profiting	28.0%	56.2%	20.6%
	Neutral	33.3%	33.3%	33.3%
	Losing	33.3%	33.3%	33.3%
Moderate-ecological	Profiting	32.1%	56.2%	20.6%
	Neutral	33.3%	33.3%	33.3%
	Losing	30.0%	33.3%	33.3%

3.6.4.4 Decision preference in a guaranteed losing situation (Low DOA)

Similar to the previous analysis on decision preference in guaranteed profiting situations, decision preference in a guaranteed losing situation by calculating the total number of guaranteed losing situations were calculated for each moderate conventional scenario and high ecological scenario, $N = 23$. The percentage of execution that was associated with each risk level was calculated using a 2 (scenario type: moderate conventional or moderate ecological) × 3 (system state: profiting, neutral or losing) × 3 (risk level: risk-aversion, moderate risk-seeking or high risk-seeking) within-subject statistical model. The scenario type effect was not significant, $p > .05$ (moderate conventional: $Mdn = 33.3\%$; moderate ecological: $Mdn = 33.3\%$). The system state effect was also not significant, $p > .05$ (profiting: $Mdn = 33.3\%$; neutral: $Mdn = 33.3\%$; losing: $Mdn = 33.3\%$). Results of the Friedman's test showed that the risk level effect was significant, $\chi^2 = 29.826$, $p < .001$ (risk-aversion: $Mdn = 18.2\%$; moderate risk-seeking: $Mdn = 60.7\%$; high risk-seeking: $Mdn = 21.3\%$). Post hoc test results showed the participants were more likely to make a moderate risk-seeking decision than a high risk-seeking decision ($z = 4.106$, $p < .001$, $r = .605$, a large effect) and a risk-averse decision ($z = 4.197$, $p < .001$, $r = .619$, a large effect). The difference between high risk-seeking and risk-averse was also significant ($z = 2.007$, $p = .045$, $r = .296$, a small effect). The risk level effect is portrayed in Figure 37.

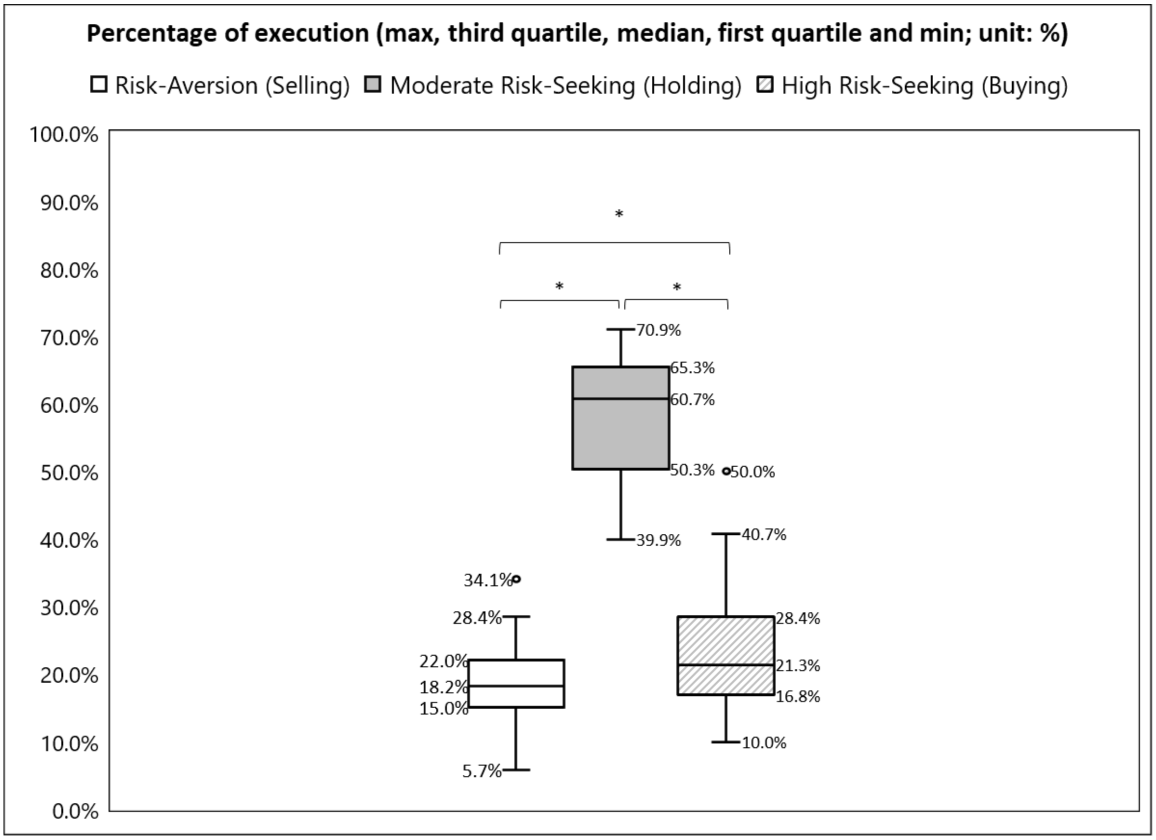


Figure 37. Percentage of execution in a guaranteed losing situation: Risk level effect (Experiment 1).

The rest of the analysis aimed at understanding how scenario type (moderate-conventional or moderate-ecological) could influence decision preference within each risk level (risk-aversion, moderate risk-seeking or high risk-seeking):

1. For the risk-averse risk level, the scenario type effect was not significant, $p > .05$ (moderate-conventional: *Mdn* = 17.6%; moderate-ecological: *Mdn* = 18.0%). The system state effect was significant, $\chi^2 = 21.356$, $p < .001$ (profiting: *Mdn* = 33.3%; neutral: *Mdn* = 8.3%; losing: *Mdn* = 19.9%). Post hoc test results showed that the difference between the profiting state and the neutral state ($z = 3.848$, $p < .001$, $r = .567$) and the difference between the losing state and the neutral state were significant ($z = 3.102$, $p = .002$, $r = .457$). The difference between the profiting state and the losing state, however, was not significant, $p > .05$;

2. For the moderate risk-seeking risk level, the scenario type effect was not significant, $p > .05$ (moderate-conventional: *Mdn* = 63.2%; moderate-ecological: *Mdn* = 60.0%). The system state effect was significant, $\chi^2 = 13.604$, $p = .001$ (profiting: *Mdn* = 41.7%; neutral: *Mdn* = 77.1%; losing: *Mdn* = 60.5%). Pairwise comparison results showed that the difference between the profiting and the losing state ($z = 2.601$, $p = .009$, $r = .383$), the difference between the neutral state and the profiting state ($z = 3.650$, $p < .001$, $r = .538$) and the difference between the neutral state and the losing state ($z = 2.068$, $p = .039$, $r = .305$) were all significant;

3. For the high risk-seeking risk level, the participants made significantly more buying executions in the moderate-ecological scenarios than in the moderate-conventional scenarios, $z = 2.403$, $p = .016$, $r = .354$ (moderate-conventional: *Mdn* = 20.6%; moderate-ecological: *Mdn* = 24.1%), as demonstrated in Figure 38. This effect is a medium effect. The system state effect was significant, $\chi^2 = 10.352$, $p = .005$ (profiting: *Mdn* = 33.3%; neutral: *Mdn* = 13.2%; losing: *Mdn* = 20.0%). The difference between the profiting and the losing state ($z = 3.088$, $p = .002$, $r = .455$) and the difference between the neutral state and the profiting state ($z = 2.129$, $p = .033$, $r = .314$) were significant. The difference between the neutral state and the losing state was not significant, $p > .05$.

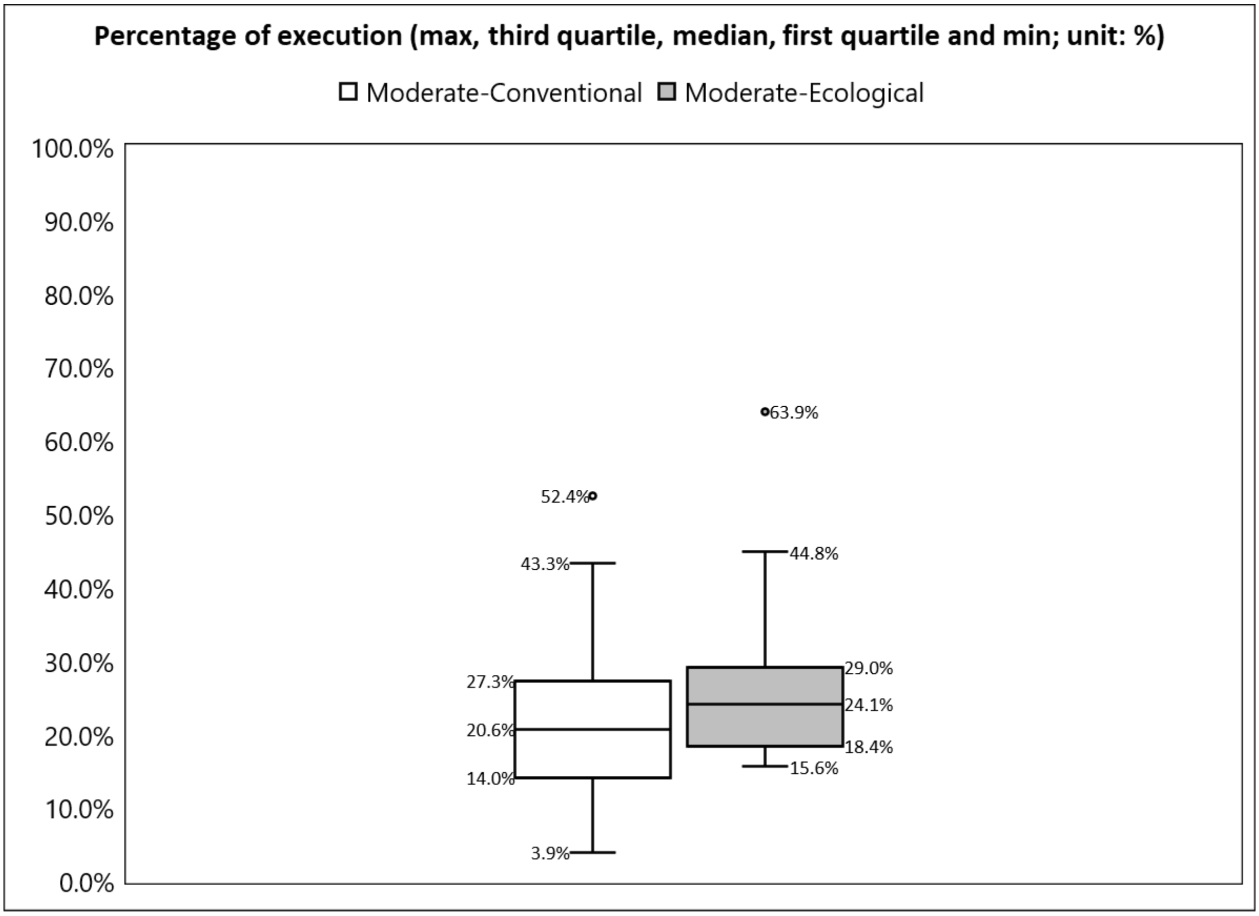


Figure 38. Percentage of buying execution in a guaranteed losing situation: Scenario type effect (Experiment 1).

The ecological display being used in the moderate DOA scenarios has made the participants more likely to take high risk-seeking executions in guaranteed losing situations than with the conventional display, a similar result demonstrated in the mean portfolio's size data. Table 25 summarizes all median percentages of execution that have been calculated.

Table 25. Summary of Percentages of Executions in a Guaranteed Losing Situation (Experiment 1).

Scenario Type	System State	Percentage of Execution (*Mdn*)		
		Risk-Aversion	Moderate Risk-Seeking	High Risk-Seeking
Moderate-conventional	Profiting	33.3%	33.3%	33.3%
	Neutral	5.6%	80.0%	7.6%
	Losing	23.1%	54.8%	23.5%
Moderate-ecological	Profiting	33.3%	33.3%	33.3%
	Neutral	2.9%	80.0%	7.6%
	Losing	16.7%	54.8%	23.5%

3.7 Discussion

Experiment 1 simulated trend following trading with a higher DOA configuration similar to the original, high DOA trend following trading scenario provided in Part B of this dissertation and a moderate DOA configuration that reverted later stages of automation to human operation. The moderate-ecological display was designed to visualize a clear market-portfolio-execution relationship for supporting problem-solving behaviour with the moderate DOA. The high-ecological display again addressed the market-portfolio-execution relationship and provided problem-solving and procedural support pertaining to detecting automation failure with a states-task visualization for the high DOA configuration. Task performance, SA, eye tracking pattern, workload and risk preference were evaluated in Experiment 1. The participants monitored unanticipated situations in the AUTRASS simulation, and their flexible trading performance was expected to be best supported in the moderate-ecological scenario and their automation failure detection performance was expected to best support in the high-ecological scenario.

3.7.1 Performance

3.7.1.1 Task performance

According to hypothesis 1a, the participants should achieve better performance with the two scenario types that used ecological displays. That being said, the participants should achieve better flexible trading performance with moderate-ecological than with moderate-conventional and better automation fault detection performance with high-ecological than with high-conventional. However, neither case was observed. Hypothesis 1a is rejected.

End of scenario RPL: A detection-mitigation confound

The participants did not achieve a better end of scenario *RPL* with the moderate-ecological scenario type. A probable explanation to this result is that end of scenario *RPL* not only measures detection performance (i.e., how well the participant detected market disturbances), but also describes mitigation performance (i.e., how well the participant mitigated the disturbances to improve the profitability of the trading system). The participants were responsible for detecting and mitigating the disturbances in the financial trading while the ecological display only provided support in the detection phase. The performance of the mitigation in financial trading may be largely influenced by individual differences. With a considerably large amount of human operation involved in the moderate conventional scenario and the moderate ecological scenario, detection and mitigation could be confounded, and the performance difference between moderate-conventional and moderate-ecological in detection would be diminished. This explanation was directly supported by Lau et al. (2008), who observed no direct performance improvement with their ecological display during the mitigation of disturbances, as their ecological display did not provide relevant cues to support the mitigation.

Providing cues on an ecological display to support mitigation could be challenging in a financial trading setting. In the real world, mitigation performance typically reflected both the risk and the reward of financial trading. For example, *Sharpe Ratio* (1994) is a real-world measure of risk-adjusted return for evaluating a portfolio's performance. It can be calculated by subtracting the rate of return of a risk-free financial product (e.g., a debt instrument issued by a government) from the rate of return of the portfolio, then dividing that result by the standard deviation of the portfolio's return. As a real-world measure, Sharpe Ratio suggests that the mitigation performance in financial trading is attributable to both the effectiveness of the trading strategy (including strategy used by the trading algorithm and strategy manually executed by the trader) and the volatility of holding a highly risky financial product (decided by the market and is probably beyond the control of most traders). The effectiveness of the trading strategy may be determined by the support for problem solving provided by the ecological displays as well as many other factors, including expertise and training which was observed by Lau et al. (2008). Obviously, a professionally trained trader is more likely to consistently profit in trading than a typical university student similar to the participants attending this experiment. In a broader sense, financial trading is ultimately a probability game where every trading system has its probability of success (Treynor, 1981). The probability of success is subject to many factors and, not surprisingly, only a small set of these factors have been simulated in this experiment.

Fault detection accuracy: Ceiling effect, stages of automation misconnection, and inadequate training

It has been found that both the conventional display used in the high conventional scenario and the ecological display used in the high ecological scenario effectively supported fault detection for unanticipated situations, but the ecological display did not provide better support over the conventional display. This result was largely unexpected. Since the system did not deviate from its present stage whether an automation failure was detected (i.e., the trading algorithm strictly followed its rule no matter if the participants clicked on the "report loss" button or not), fault detection accuracy only described detection performance and did not affect the mitigation performance. Therefore, there was unlikely a detection-mitigation compound. Here the author offers three explanations for the absence of a statistical difference in the fault detection accuracy in the type of display.

The first explanation for the missing difference between the high conventional display and the high ecological display is a ceiling effect, given many central tendency measures indicated that the participants performed the task extremely well. In particular, the most frequent value in the simulation data was 100% with both the conventional display and the ecological display, as demonstrated by the mode measure. The ceiling effect was possibly caused by the low incidence rate of the automation failure (2.67%: four total losing buy-sell pairs divided by 180 timestamps with D3; 4.17%: five total losing buy-sell pairs divided by 120 timestamps with D4). The low incidence rate was mainly attributable to the "inverted two moving average methods" used by the algorithm logic, which generally performed well in the simulated market. Previous EID studies have found ecological displays have a larger effect on task performance in unanticipated situations and generally no difference in performance in anticipated situations. Although the slippage still caused latency issues in automated trading, the simulated environment may be more anticipated than it was intended to be.

The second explanation is the misconnection between the earlier stages of automation (i.e., perceptual understanding of the market trend through the two SMA curves presented on the market panel) and later stages of automation (i.e., the operation logic of the trading algorithm). The participants may perceptually develop a consistent rule-based mapping between the cues provided by the two SMA curves on the market panel and the actions described by the Murphy method. In contrast, the ecological display provided problem-solving support based on the "inverted two moving average" logic and was inconsistent with this rule-based mapping. The complexity of the work domain unexpectedly increased with the misconnection between the perceptual support provided by the earlier stages of automation and the inconsistent automation behaviour determined by the later stages of automation. With limited support provided by the states-task visualization to deal with the increased complexity, no mental model that would support knowledge-based problem solving was developed with the ecological display.

The third explanation is inadequate training. The algorithm logic was unfamiliar to the participants and might have a negative impact on the training of the ecological displays. As Borst et al. (2015) clarified, operators required certain information in the deep structure of the work domain to think productively, and EID makes such information transparent to the operators but cannot eliminate the need for training. Although the states-task visualization also provided procedural support to develop a consistent mapping from the heuristic cues (i.e., height of the $P_{buy'}$ box and height of the $P_{sell'}$ box) to the algorithm performance (i.e., winning or losing buy-sell pairs) and the procedural supported should require less training, it seemed that this rule-based mapping did not replace the missing support of knowledge-based problem solving, which was, by the definition of the ecological displays, most important to the improvement of detection performance over the conventional display in unanticipated situations.

3.7.1.2 Situation awareness and eye tracking

Participants attending this experiment had an intermediate SA (*Mdn* = .500) as measured by SAGAT. It has been found that the higher DOA configuration degraded SA. This finding is in general agreement with the automation trade-off which suggested that increasing DOA resulted in a loss of SA. The meta-analysis of automation trade-off has forecasted that ecological displays could modify or even reverse the automation trade-off (Onnasch et al., 2014). However, in this experiment, the conventional and the ecological displays resulted in no difference in SA. Therefore, hypothesis 1b is rejected. Three explanations for this finding is provided as follows.

First, the participants may have not found the ecological displays adequately useful in eliciting answers to certain SA queries, especially the level 3 SA queries. Since SA queries were randomly picked from the pool, a participant may experience two queries describing the same situation through opposite angles (e.g., “do you think there will be a buying opportunity in the next 5 seconds? why?” and “do you think there will be a selling opportunity in the next 5 seconds? why?”). Conflicting responses to these questions were frequently documented in the SA data. For example, some participants have answered “yes” to both questions, showing low confidence in their ability to make predictions. An improved SA query design, however, may describe the financial trading context in more precise way.

Second, the conventional displays may already have some features that can improve the awareness of the trends of simulation that paralleled the ecological displays. For example, the arrow displayed to the right of the market price indicator had useful sensory features - coloured in green when the market moved up or red when the market moved down.

Third, it is possible that the ecological displays could provide some benefits to retaining SA, but the benefits were limited to portfolio management in the moderate DOA scenarios. This explanation was not directly supported by the SA rating data, as these data were collected by aggregating many situational aspects. Indeed, there was no statistically significant SA difference between moderate-conventional and moderate-ecological, and high-conventional and high-ecological. This explanation, however, is supported by the significant shorter total dwell time on portfolio AOI with moderate-ecological than with moderate-conventional. Eye tracking measure indicates attention allocation and has been closely related to level 1 SA, as discussed in the literature. Indeed, Endsley reviewed a variety of physiological measures of SA during the development of SAGAT (1988) and included eye-tracking measures. Gugerty reviewed eye-tracking measures as a category of online SA measures which did not require the simulator to be paused and suggested that the most commonly used eye-tracking measure in the driving research is dwell time (2011). Van de Merwe, van Dijk and Zon (2012) pointed out that dwell time measures the relative importance of the display and therefore, it can be used as a predictor of performance and SA. Total dwell time captures long-term cognitive processes (Holmqvist et al., 2011) that are related to slower decision-making in comparison to dwell time per dwell. Indeed, the participants had empirically similar level 1 SA in the moderate-conventional scenario and the moderate-ecological scenario.

3.7.1.3 Workload

The pattern of NASA TLX rating data in the response data is clear. The perceived workload in Experiment 1 was moderate. The main effect of DOA was expected given that the high DOA configuration only required monitoring the automation while moderate DOA required both participant intervention and monitoring. Ecological displays did not impose higher workload as predicted in hypothesis 1c.

3.7.2 Risk Preference

Overall risk preference

The fourfold pattern of preferences drawn across the participants was somewhat different from those observed in previous studies on description- and experience-based choices. No scenario type effect was observed. As the result, hypothesis 2a was partially supported. A comparison of the different patterns observed in description-based choice research (i.e., prospect theory), experience-based choice research (i.e., experience through learning and experience through professional training) and this experiment is presented in Table 26.

Table 26. Empirical Choice Pattern Extended to DOAs and Displays (Experiment 1).

	Description (prospect theory)		Experience through learning (Hertwig & Erev, 2009)		Experience through professional training (McAndrew & Gore, 2013)		Experiment 1	
	Gains	Losses	Gains	Losses	Gains	Losses	Gains	Losses
Small probability	Risk seeking	Risk aversion	Risk aversion	Risk seeking	Risk aversion	Risk aversion	Risk aversion	Risk seeking
Medium and large probability	Risk aversion	Risk seeking	Risk aversion?	Risk seeking?	Risk seeking	Risk aversion	Risk seeking	(No consensus)

McAndrew and Gore's findings were derived from a real-world financial trading setting, and similarly, the current experiment presented in this dissertation described a pseudo real-world setting of trading. As the results of the current experiment show, risk aversion for small-probability gains (67 out of 96 choices) and risk seeking for small-probability losses (69 out of 96 choices) oppose the McAndrew and Gore's findings but are consistent with the patterns of experience through learning (Hertwig & Erev, 2009). It is possible that the participants made decisions from experience gained through AUTRASS, as they were not as professionally well-trained as in McAndrew and Gore's case. However, their experience was not obtained from an abstract learning environment similar to Hertwig and Erev's setting either. With AUTRASS, the participants were guided by visual cues and task requirements from the physical environment and were not required to conclude a clear structure of the statistical probabilities, as opposed to Hertwig and Erev's participants who learned the "probability structure over outcomes through trial-by-trial feedback" (2009). This finding expands the understanding of how people make decisions "with incomplete and uncertain information 'in the wild'" (Hertwig & Erev, 2009). For novice participants, gaining experience through learning in a simulated environment may foster a similar pattern of risk preference in comparison to learning with statistical probabilities unfolded.

Participants were generally risk seeking for medium- and large-probability gains (61 out of 96). This result was also supported by the quantitative results that the participants were generally moderate risk-seeking (i.e., holding the portfolio) while facing a large-probability prospect gain (i.e., a guaranteed profiting opportunity). It has only been speculated in the literature that risk aversion for medium- and large-probability gains and risk seeking for medium- and large-probability losses for experience-based choices (Rakow & Newell, 2010). This current experiment is in a position similar to McAndrew and Gore's work and generally opposed this speculation. A possible explanation for this result is that preference in medium- and large-probability prospect gains is more likely attributable to the naturalistic trading environment than the professional expertise of the trader. No consensus was found in participants' risk preference for medium- and large-probability losses. Since no statistical significance was observed, the difference may be the result of the small sample size, the DOA or display effects or the individual differences.

Ecological displays and risk-seeking actions

The participants held a significantly larger portfolio size when they were presented with the ecological display, which generally follows the same pattern in the aviation domain as described by Borst et al. (2015). In this sense, hypothesis 2b was supported. While most EID studies in the literature were focused on addressing how ecological displays could improve performance and SA for monitoring unanticipated events, this finding is relatively new.

Borst et al. (2015) recently reported that aircraft pilots were sometimes more risk-seeking in taking actions with ecological displays, which was likely attributable to the more salient physical structure (e.g., limits in flight control) than the intentional structure in the work domain (e.g., aviation safety regulations) on their ecological displays. Indeed, in Experiment 1, a rich physical structure has been visualized with the market-portfolio-execution visualization with the goal of maximizing the profit.

Changes in the market (e.g., price movement) or in the portfolio (e.g., buy or sell) caused the unrealized profit and loss (*UPL*) to change, and was well reflected on the shaded portion of the market-portfolio-execution visualization. The shaded portion used different colours to represent the *UPL*, indicating a clear boundary of the system performance and becoming a strong draw for the participants who were novice in financial trading. Although described in the CWA models, intentional constraints pertaining to the laws and regulations of risk management were not graphically represented on the ecological displays, as they would more likely be utilized by more expert participants.

The market-portfolio-execution visualization made the limits in the market-portfolio-execution relationship clear to the participants, improving their confidence in understanding the performance of the trading system. It seems that Borst et al.'s comment on the aviation domain was also applicable to the automated trading domain. Participants attending the automated trading experiment may be prone to "maneuvering themselves in narrow control spaces that leave little room for error" (2015). Indeed, portfolio size is similar to flight envelope in the sense that there is a trade-off between maneuverability and safety. Ecological displays could influence risk tolerance and strategies in financial trading as well. It might be worth developing new approaches to exploring how an ecological display could be structured to invoke certain risk-related strategies similar to Hilliard and Jamieson's approach (2014) and how to graphically represent intentional constraints in this work domain. However, it can be foreseen that the new displays containing the intentional constraints may require more expert knowledge to use, which was another suggestion of Borst et al.'s (2015).

System states and risk-seeking actions

The participants were generally moderate risk-seeking as demonstrated by their decisions in guaranteed profiting situations which were associated with large-probabilities gains. This result elaborates on the finding of the qualitative measure. Further, it is evident that the participants were generally moderate risk-seeking when they experienced guaranteed losing situations (i.e., high probability prospect losses). No consensus has been made with the qualitative measure. Thus, hypothesis 2c was partially supported.

The distinct moderate risk-seeking behaviors noted during the profiting and the neutral system states suggest there may be unique control tasks in financial trading, and may be analogous to process control where it is important to examine operator behavior under normal or fault scenarios. This finding generally supports McAndrew and Gore's interview results that professionally-trained traders have a risk-seeking choice behavior for medium- and high-probability gains. In our study, a guaranteed profit opportunity at which the participants made a sell, hold or buy decision has a high-probability prospective gain. While the participants in general preferred to take moderate risk seeking actions (i.e., holding the portfolio), they were more likely to take such actions while the system is making a profit than while the system is in a neutral state, suggesting the choice behavior may also be influenced by system state.

Although the overall pattern of risk preference in the ecological display and the system states is clear, the interaction effect of scenario type × system state could not be examined with a non-parametric test. The difficulty of teasing out the noise and individual differences in participants' responses was well aware of, from a study conducted with a small sample size and a relatively large degree of freedom in financial trading. Apparently, this is an important question that should be investigated in the future.

3.8 Chapter Summary and Connections to Research Questions

3.8.1 Key Findings

DOA-independent and DOA-specific displays: This chapter proposes new ways of developing ecological displays given the market-portfolio-execution visualization was independent of the DOAs, and the states-task visualization supported a specific DOA. This new design approach is consistent with the findings of the DOA-layered modelling work. Experiment 1 is a preliminary exploration of this approach, and the author hopes it can grow into a potentially useful approach for designing automation displays in the future.

Automated trading microworld: The AUTRASS simulation in which the participants experienced has unique features for novice traders to learn trend following trading. AUTRASS has potential value to be used as an equivalent microworld for examining ecological displays in the financial trading domain compared to DURESS in the process engineering domain. Further, being able to evaluate varying DOAs, a feature that has not been enabled with DURESS, is instrumental in developing new research programs that can align EID research with the human-automation interaction research.

Automation trade-off in financial trading: Results of this chapter show that a typical automation trade-off in the financial trading domain: with a higher DOA, traders' perceived workload is lower, but the SA is degraded. This finding inspires more human-automation interaction research in the financial trading field.

Mixed-method approach to risk preference: This chapter provides qualitative and quantitative benchmarks for measuring traders' risk preference, and this make a unique contribution.

Risk-seeking behaviours with EID: Results of this chapter support Borst et al.'s observation that human operators perform tasks in riskier ways with ecological displays.

3.8.2 Connections to Research Questions

The author has three research questions for this dissertation:

Research question 1: How can we model automated trading systems with a variable DOA using CWA?

Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

This chapter examined research question 2 and 3. Research question 3 is supported and directions to further examine research question 2 are identified. Experiment 1 is generally exploratory but has suggested directions for future works, some of which are being examined in a follow-up experiment.

Chapter 4

Experiment 2: Trend Following Trading and Adaptive Automation

4.1 Foreword

Empirical results of Experiment 1 showed there are two sides to the use of automated trading: with a higher DOA, the participants' perceived workload was lower, but their SA was degraded. This finding is on par with the results of human factors experiments in other domains as documented in the literature and have been predicted as the automation trade-off. Ecological displays, however, did not seem to alter this trade-off. In this chapter, the author extends the simulation of trend following trading to include the adaptive configuration, which was an adaptive automation condition. As the author introduced earlier in this dissertation, adaptive automation is a context-sensitive approach to manipulate the DOA to balance the benefits and costs of automation, especially in high DOA situations. Adaptive automation has become popular in a variety of domains, including aviation (e.g., Parasuraman, Bahri, Deaton, Morrison, & Barnes, 1992), aerospace engineering (e.g., Li, Sarter, Wickens, & Sebok, 2013), process control (e.g., Moray, Inagaki, & Itoh, 2000) and the supervision of unmanned vehicles (e.g., Parasuraman, Cosenzo, & De Visser, 2009). Inspired by these examples, the author developed a simulation of adaptive automation for trend following trading.

4.1.1 Adaptive or Adaptable Automation

Adaptive automation, in its broad sense, including adaptable automation, is an emerging opportunity (Bailey, Scerbo, Freeman, Mikulka, & Scott, 2006). Adaptable automation refers to function allocations based on human control, whereas adaptive automation is achieved by automation authority. There is generally little guidance on how to design adaptive automation in a human-centered way. Kaber et al (2001) suggested that direct-manipulation interfaces may be used to buffer the performance costs associated with the changes in system stages with adaptive automation (e.g., increased workload).

Developed with the complexity of socio-technical systems in mind, EID extended the benefits of direct-manipulation interfaces to situations that are unfamiliar to both designers and operators (Vicente & Rasmussen, 1992) and in theory, should further support problem solving under the dynamics and uncertainties brought by adaptive automation.

This chapter reports Experiment 2 as a follow-up to the first experiment to further examine the effectiveness of the ecological displays in unanticipated situations. Experiment 2 included the same conventional and ecological displays but used two different automation configurations. The AUTRASS continued to simulate the latency in trading to make the environment unanticipated to the participants, in which the ecological displays were expected to outperform the conventional displays. An *improved-high DOA* configuration with better automation design was similar to the high DOA configuration originally introduced in the first experiment (referred to as the *original high DOA* configuration in the rest of the dissertation) and was compared to the adaptive configuration using adaptive automation. Four scenario types were developed to make the comparisons possible. These scenario types included improved-high-conventional, improved-high-ecological, adaptive-conventional and adaptive-ecological.

4.2 Improved-High DOA Configuration

For the original, high DOA configuration described in Experiment 1, the author has discussed that the absence of a statistical difference of the task performance in the type of display could be attributable to the low incidence rate of automation failure, the misconnection between the stages of automation and inadequate training. To mitigate these design flaws, the automation design was modified and as a result, an improve high DOA configuration was adopted in this experiment. To do that, the algorithm logic was revised to use the Murphy method and therefore matched the perceptual understanding of the SMA curves presented on the market panel.

With the modified automation design, the incidence rate of automation failure would be slightly raised to reduce the ceiling effort, given that the Murphy method performed poorly in the simulated financial market and therefore, resulted in more automation failures. The performance of the Murphy method is being discussed in section 4.4.2 with a title named “financial market data”. The misconnection between the stages of automation would be eliminated in the improved-high DOA configuration with the use of the Murphy method, as the SMA information presented on the market panel matched the logic of the Murphy method. Additional training was provided to ensure the participants understand the new algorithm logic being used, as opposed to in Experiment 1 the participants were not thoroughly trained with the logic of the trading algorithm. In Experiment 2, the participants were explicitly told that the trading algorithm used the Murphy method when they experienced the improved-high DOA configuration.

4.3 Adaptive Configuration

To develop the adaptive configuration for this experiment, possible use cases of adaptable automation and adaptive automation in the financial trading context are reviewed as follows.

Adaptable automation may be common in automated trading but requires the participant to have adequate expertise in financial trading. In both the original and the improved-high DOA configurations, the participant was not allowed to stop the trading algorithm from buying or selling shares as these configurations were intended to simulate automation that has a high and typically fixed DOA. In a real-world setting, however, most automated trading systems are inherently adaptable due to the existence of the stop feature, also known as the stop order. Since the trader is ultimately responsible for the profit and loss of the trading system, they may submit a stop order to the market, and the stop order will be executed when a specified price level is breached. The stop order can be a trailing stop in which the stop is made relative to the financial product’s market price, or a fixed stop that can be triggered by rigorous algorithm logic or by the trader. In the language of human factors, strong expertise in financial trading is instrumental in determining the best timing for choosing the DOA of this adaptable system and therefore, it is less meaningful to use adaptable automation with a participant who was a novice in this field.

Since the participant was familiar with and ultimately responsible for the performance of the trading system that was essentially characterized by *RPL*, it is more feasible to implement a performance-based adaptive system. With the adaptive system reasonably designed for this experiment, the authority in making decisions and performing executions - two decisive functions in financial trading that are associated with later stages of automation - would be allocated to the automation when the trading system generally performed well (i.e., triggering a higher DOA) and to the participant when the trading system performed poorly (i.e., triggering human intervention). It has been known from the first experiment that the performance of AUTRASS could be characterized using system state ($RPL > 0$: profiting; $RPL = 0$: neutral; $RPL < 0$: losing). In this second experiment, suppose the simulation started in a higher DOA configuration in which the trading algorithm selected the decisions and implemented actions, the participant should be responsible for monitoring the temporal performance of the automation regarding the profit or loss gained through each buy-sell pair. If the automation consistently performed poorly, a significant realized loss (i.e., a negative *RPL*) would be achieved, and the participants were told to interpret this situation as a system failure. The adaptive system automatically switched the automation to a lower DOA configuration similar to the moderate DOA configuration of the previous experiment, in which the participant was requested to make decisions and performed executions manually. One benefit of using the moderate DOA configuration is that the participant might be able to regain SA with a lower DOA. On the other hand, if the trading system gained enough realized profit (i.e., $RPL > 0$) during this intervention, the trading system should be assumed to enter a market that favours the trading algorithm, and the DOA should be able to return to a higher degree to reduce the participant's workload.

The DOA layering approach introduced earlier in this dissertation was useful in determining the design requirements with different DOAs which together build an adaptive system. With this approach, DOA shifts in trend following trading can be modeled on a DL as demonstrated in section 2.5.3. Although the DOA layering approach was preliminary in modeling adaptive automation and no direct design support was provided on DOA shifts at the moment, with the improved automation design provided for this experiment, the ecological displays were expected to provide support in the profiting state of the adaptive configuration at the least.

4.4 Apparatus

This experiment continued to use AUTRASS as the simulator. The Gazepoint GP3 eye tracker continued to provide support in collecting eye-tracking data. Most apparatus of Experiment 1 and 2 were identical, including the conventional and ecological displays being used. The differences were generally in the automation design and will be highlighted in the following sections.

4.4.1 Automation Design

Table 27 shows that the improved-high DOA configuration and the profiting state of the adaptive configuration inherited all task phases of the original high DOA configuration in Experiment 1 with the stage of decision selection modified to use the Murphy method. The author has discussed that the Murphy method had poorer performance than the inverted two moving average method previously used in Experiment 1. Not surprisingly, the poorer performance created more needs for monitoring the algorithm trading in the market and more opportunities for the participants to intervene during the losing state of the adaptive configuration.

Table 27. Distinct Operation Logic of the Two Moving Averages Methods.

Stages of Automation	Task Phases	Improved-High		Adaptive	
		Trader	Automation	Trader	Automation
Information acquisition	Collect information from the market quotes and display it on a display (in favor of the traders); Observe market and portfolio data		×		×
Information analysis	Calculate a short-period SMA and a long-period SMA and plot the curves onto a display (in favor of the traders)		×		×
Decision selection	Interpret the current and predict situations to decide signals to buy and sell		×	× (DOA shifted to moderate, *RPL* < 0)	× (DOA shifted to high, *RPL* > 0)
Action implementation	Determine and perform a buying or a selling task		×	× (DOA shifted to moderate, *RPL* < 0)	× (DOA shifted to high, *RPL* > 0)

4.4.2 Financial Market Data

The improved-high DOA configuration replaced the original high DOA configuration of the first experiment to use data sets TD3, TD4, D3 and D4 (Table 28). The Murphy method used in improved-high DOA has been known to perform worse with those simulated markets in in comparison to the inverted "two moving average" method. However, since the primary task for the participants in the high DOA configurations was to monitor the automation trading, it is more important to ensure the total numbers of losing buy-sell pairs to which the participants should respond for both D3 and D4 were approximately the same. The author has previously reported that for the original high DOA configuration using the two-moving average method, there were in total four losing buy-sell pairs with the D3 and five losing buy-sell pairs with the D4. For the improved-high DOA configuration, there were in total eight losing buy-sell pairs with the D3 and seven losing buy-sell pairs with the D4, suggesting the task performance could be measured in a consistent way. The slightly increased incidence rate that could be defined by the number of losing buy-sell pairs may reduce the ceiling effect which might have occurred in Experiment 1. However, participants experienced a more familiar environment in the improved-high DOA configuration than in the original high DOA configuration, as the algorithm logic was known and consistent with the perceptual support provided on the conventional displays. This topic will be reviewed in the data analysis.

The adaptive configuration replaced the moderate DOA configuration to use data sets TD1, TD2, D1, and D2. Since the participants were expected to experience both the profiting state (in which automation performed the executions, and the participants had less task involvement, using a higher DOA) and the losing state (in which the participants performed a flexibility trading task, using a lower DOA), it is important to ensure that the Murphy method could generate both profiting states and losing states in these simulated markets. Using the Murphy method, the trading system would enter the first losing state at approximately the same time with both the D1 and the D2 data sets, providing the participants a consistent starting point to recover the trading system from further failure. With D1, the *RPL* would fall to -\$5.00 at the 12th timestamp. With D2, *RPL* would fall to -\$26.00 at the 11th timestamp. The D1 and D2 markets were consistent at the starting point. However, the performance of the trading system after this point was beyond the control of the experimental design.

Table 28. SPY Historical Market Data Sets (Experiment 1 and 2 Combined).

Data Set	Starting Price	Original Trading Day	Total Timestamps	Total Duration (min)	DOA		
					Experiment 1		Experiment 2
TD1	111.76	August 2, 2010	60	5	Moderate	→	Adaptive
TD2	110.63	July 26, 2010	60	5	Moderate	→	Adaptive
TD3	106.76	July 8, 2010	60	5	High	→	Improved-high
TD4	117.67	April 8, 2010	60	5	High	→	Improved-high
D1	115.75	November 30, 2010	120	10	Moderate	→	Adaptive
D2	114.11	October 18, 2010	180	15	Moderate	→	Adaptive
D3	118.37	October 7, 2010	120	10	High	→	Improved-high
D4	118.98	September 22, 2010	180	15	High	→	Improved-high

4.4.3 Conventional Displays

The conventional displays used in the improved-high DOA configuration and the profiting state of the adaptive configuration (i.e., DOA shifted to high) were identical to that used in the original high DOA configuration in Experiment 1. The losing state of the adaptive configuration (i.e., DOA shifted to moderate) used the moderate DOA conventional display in Experiment 1, as demonstrated in Table 29.

Table 29. Display Elements of the Conventional Displays for the Improved-High DOA Configuration and the Adaptive Configuration.

High DOA (Improved)	Adaptive	
	DOA shifted to moderate, *RPL* < 0	DOA shifted to high, *RPL* > 0
Market panel		
Fundamental history panel		
Portfolio panel		
Trading history panel		
Execution panel (high DOA)	Execution panel (moderate DOA)	Execution panel (high DOA)

4.4.4 Ecological Displays

The improved-high DOA configuration and the profiting state of the adaptive configuration each adopted the same ecological display used in the original high DOA configuration of the first experiment. The losing state of the adaptive configuration used the moderate DOA configuration ecological display, as shown in Table 30.

Table 30. Display Elements of the Ecological Displays for the Improved-High DOA Configuration and the Adaptive Configuration.

High DOA (Improved)	Adaptive	
	DOA shifted to moderate, *RPL* < 0	DOA shifted to high, *RPL* > 0
Market panel		
Fundamental history panel		
Portfolio panel		
Execution history panel		
Market-portfolio-execution visualization		
Execution panel for high DOA and the states-task visualization	Execution panel for moderate DOA	Execution panel for high DOA and the states-task visualization

4.5 Method

4.5.1 Experimental Design

Similar to Experiment 1, this experiment was generally a one-way (scenario type: improved-high-conventional, improved-high-ecological, adaptive-conventional or adaptive-ecological), within-subject design.

4.5.2 Procedure

A request for ethics clearance of a modification was approved by a University of Waterloo research ethics committee on July 12, 2016 (ORE #: 21061). The recruitment process and the experimental procedure were generally identical to those of Experiment 1, except for the training slides and training scenarios.

The training slides were modified to include an introduction to the Murphy method. The participants were explicitly told that the Murphy method was the operation logic of the trading algorithm used in the improved-high DOA configuration and the profiting state of the adaptive configuration. The participant experienced the first training scenario, which was improved-high-conventional. The participant was then introduced to the ecological display and completed the second training scenario – improved-high-ecological. After that, the participant was trained with the adaptive automation configuration and completed the other training scenarios.

4.5.3 Participants

None of the participant attended Experiment 1 prior to attending this experiment. All participants were registered undergraduate and graduate students at the University of Waterloo. Eight females and sixteen males voluntarily participated in this study, and each was remunerated 30 Canadian dollars for their participation. The average age of the participants was 22.7 (SD = 3.172), which was similar to that of Experiment 1 participants (*M* = 25.1 years, *SD* = 3.256). All participants self-reported they have a normal or corrected normal visual acuity and normal color vision, and they would be comfortable interacting with numeric and colour visualizations displays (rated at least 3 on the 5-point scale in both cases). All participants have successfully completed at least one computer programming course to facilitate automated trading. Five participants stated they had had experience in personal investment. One participant had worked in the forex industry. All other participants were novice in financial trading.

4.5.4 Task Descriptions

4.5.4.1 Improved-High DOA Configuration: Fault detection task

In the improved-high DOA scenarios, the participants were requested to monitor the algorithm trading and completed a fault detection task used in original high DOA.

4.5.4.2 Adaptive Configuration: Flexible trading task and fault detection task

The participants performed the flexible trading task when the DOA shifted to moderate and performed the fault detection task when the DOA shifted back to high.

4.5.5 Independent Variables

The independent variables were generally the same as those of Experiment 1, as shown in Table 31.

Table 31. Summary of the Independent Variables (Experiment 2).

Independent Variable	Type	Level	Used with
Scenario type	Within-subject	Improved-high-conventional, improved-high-ecological, adaptive-conventional or adaptive-ecological	All dependent variables
Financial market data*	Within-subject	D1, 2, 3 or 4	Manipulated in the experimental design (D3 and 4 for improved-high DOA and D1 and 2 for DOA adaptive) but was not included in data analysis
System state	Within-subject	Profiting or losing	Used in the evaluation of mean position size portfolio and decision preference in a guaranteed profiting situation
Area Of Interest (AOI)	Within-subject	Market, portfolio, trading history, market-portfolio-execution or states-task	Used in the evaluation of all eye-tracking measures: 1) the market, portfolio and trading history AOIS were evaluated with all scenario types; 2) the market-portfolio-execution AOI was involved in the evaluation of conventional display scenarios; 3) the states-task AOI was evaluated within the improved-high DOA scenarios only
SA Level	Within-subject	1, 2, or 3.	Used in the evaluation of SA rating
NASA TLX subscale	Within-subject	Mental, physical, temporal demand, performance, effort or frustration	Used in the evaluation of perceived workload

For DOA, the moderate DOA configuration and the original high DOA configuration were replaced with the adaptive configuration and the improved-high DOA configuration respectively. For system state, the neutral state was used as a buffer for the transition between the profiting state and the losing state of the adaptive configuration in which the participants used distinct control mechanisms; therefore, the neutral state was excluded from the analysis.

4.5.6 Dependent variables

Most dependent variables examined in Experiment 1 were kept, as presented in Table 32. A new dependent variable was included to the losing state of adaptive for understanding how well the ecological displays would help the participants to bring *RPL* back to positive. Minor modifications to the dependent variables are being introduced as follows.

Table 32. Summary of the Dependent Variables (Experiment 2).

Research Question to Answer	Category	Dependent Variables	Type	Independent Variables (Number of Levels)	Data Collection Method
Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?	Task performance	End of scenario *RPL*	Ratio	Adaptive DOA only: scenario type (2)	Simulation data
		Mean accumulating *RPL* (only when DOA shifted to moderate) *	Ratio	Adaptive DOA only: scenario type (2)	Simulation data
		Duration of losing state*	Ratio	Adaptive DOA only: scenario type (2)	Simulation data
		Fault detection accuracy	Ratio	Scenario type (2)	Simulation data
	Situation awareness	Mean SA score	Ratio	Scenario type (4), SA Level (3)	Computer-administrated questionnaire
	Eye tracking	Total dwell time	Ratio (log-transformed)	Scenario type (4), AOI (6)	Eye tracking data
	Workload	NASA TLX rating	Ratio	Scenario type (4), NASA TLX subscale (6)	Paper-based questionnaire
Research question 3: Can ecological displays	Risk preference	Choice of options (as part of the fourfold pattern of preferences)	Binary (nominal)	Scenario type (4), fourfold situations (4)	Paper-based questionnaire

Research Question to Answer	Category	Dependent Variables	Type	Independent Variables (Number of Levels)	Data Collection Method
influence trader's risk preference? If they can, in which DOA does this influence exist?		Mean portfolio's size	Ratio	Scenario type (2), system state (2)	Simulation data
		Decision preference in a guaranteed profiting situation	Ratio	Scenario type (2), system state (2)	Simulation data
		Decision preference in a guaranteed losing situation	Ratio	Scenario type (2), system state (2)	Simulation data

* New measure in Experiment 2.

4.5.6.1 Task performance measures

Similar to Experiment 1, *end of scenario RPL* was used as the measure of task performance for the adaptive-conventional and adaptive-ecological scenarios. Note that this measure reflected the joint performance of trader performance (when DOA shifted to moderate) and trading algorithm performance (when DOA shifted to high) and therefore, may be subject to known or unknown confounds related to DOA shifts. However, this measure was kept for consistency with Experiment 1.

Mean accumulating RPL was a newly introduced performance measure that only described trader performance in the flexible trading task (when DOA shifted to moderate). It has been previously introduced that at each timestamp, the AUTRASS back-end calculated the accumulating *RPL* and immediately plotted its value on the simulator screen. Consequently, the trading system experienced either a profiting, a neutral or a losing state at any timestamp, as demonstrated in section 3.4.5.3. Since the participants only performed the flexible trading task when the system was in a losing state, mean accumulating RPL was collected by adding up the ending *RPL* when each losing state ended then dividing the result by the number of losing states.

Fault detection accuracy was the same measure that has been used in Experiment 1, high DOA scenarios. Fault detection accuracy continued to be used as the measure to understand how well the participants detected the temporal failure of the trading algorithm in both the improved-high DOA scenarios and the profiting state of the adaptive DOA scenarios (when DOA shifted to high).

4.5.6.2 Situation awareness measure

The measure of SA was based on SAGAT and was essentially the same as in Experiment 1. A modified SA query pool as shown in Table 33 was used. Three modifications were made to the SA query pool:

1. Several new SA queries about the function allocation were added to the pool. A new level 1 SA query asked about who (automation or the participant) was recently taking the control in trading on AUTRASS. A new level 2 SA query examined whether the participant could successfully identify DOA shifts in the stages of decision selection and action implementation;

2. Existing SA queries representing the same situation of the simulation through opposite angles were merged to eliminate the confusions on these queries as identified from Experiment 1 results. These questions included 2 level 2 SA queries (i.e., "is there a buying opportunity? why?" and "is there a selling opportunity? why?") and 4 level 3 SA queries (i.e., "do you think there will be a buying opportunity in the next 5 seconds? why?", "do you think there will be a selling opportunity in the next 5 seconds? why?", "do you think there will be a buying execution in the next 5 seconds? why?", "do you think there will be a selling execution in the next 5 seconds? why?");

3. One of the level 1 SA queries ("in the last 5 seconds, what was the most recent trade?") pertaining to the trading execution was previously removed from the analysis of Experiment 1 data and was revised with an additional option ("nothing") added for the participant to accurately describe all situations.

Table 33. SA Query Pool (Experiment 2).

SA Level	SA Query	Answer Choice
1	"In the last 5 seconds, the market price has gone:"	Up, down or flat
	"In the last 5 seconds, the slower moving average curve (yellow) has gone:"	Up, down or flat
	"In the last 5 seconds, the faster moving average curve (purple) has gone:"	Up, down or flat
	"In the last 5 seconds, was there a crossover of the two moving average curves?"	Yes or no
	"In the last 5 seconds, ___ was taking the control in trading."	Automation or I
	"In the last 5 seconds, what was the most recent trade?"	Buy, sell or nothing
2	"Is there a shift in control (e.g., you and automation)? Why?"	(Open-ended)
	"What is happening with the unrealized profit and loss in your portfolio? Why?"	
	"Is there a trading opportunity? What kind of opportunity, and why?"	
	"What is happening with the quantity in your portfolio? Why?"	
	"Has the most recent trade made any realized profit, or loss? Why?"	
3	"What will happen to the market price in the next 5 seconds? Why?"	(Open-ended)
	"Do you think there will be a trading opportunity in the next 5 seconds? What kind of opportunity, and why?"	
	"Do you think there will be a trade in the next 5 seconds? What kind of trade, and why?"	
	"What will happen to the quantity in your portfolio in the next 5 seconds? Why?"	

4.5.6.3 Eye tracking measure

The assignment of AOIs and the eye-tracking measure remained unchanged given the same conventional and ecological displays were used in Experiment 1.

4.5.6.4 Workload measure

Unweighted NASA TLX has previously been used as the workload measure and was kept in this experiment.

4.5.6.5 Risk preference measures

The fourfold pattern of preferences continued to serve as the qualitative measure of the participants' risk preference. The quantitative measures, including the mean position size of the portfolio and the decision preference in a guaranteed profiting situation and a guaranteed losing situation, were only examined for the losing state of the adaptive configuration where the flexible trading task was taken place.

4.6 Research Hypotheses

The author has three research questions:

Research question 1: How to model automated trading systems with a variable DOA using CWA?

Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

This second experiment examined research question 2 and 3 with new automation design. Several hypotheses were modified to better describe the difference in apparatus of Experiment 2.

4.6.1 Research Hypotheses for Examining Performance

H1: The participants should achieve better performance with ecological displays in comparison to conventional displays.

Hypothesis 1 generally remained consistent and was examined by dependent variable:

H1a: Within improved-high overall and the profiting state of the adaptive configuration (DOA shifted to high), the participants could perform better in the fault detection task with the ecological displays compared to the conventional displays. Within the losing state of the adaptive configuration (DOA shifted to moderate), no difference in flexible trading task performance would be found between the conventional and the ecological displays. Hypothesis 1a was slightly modified on the side of the flexible trading task. It has been known from Experiment 1 results that the flexible trading task involved both detection and mitigation and might not be supported by ecological displays. The losing state of the adaptive configuration was similar to the moderate DOA configuration and might follow the same pattern. On the other hand, with design changes in the improved-high DOA configuration, the ecological display should effectively support monitoring for unanticipated situations in the improved-high DOA configuration, and this hypothesis should be examined to verify the design change.

H1b: The participants' SA would be higher with improved-high-ecological than with improved-high-conventional. The participants' SA would be higher with adaptive-ecological than with adaptive-conventional. There might also be some evidence in the eye-tracking measure data that can support the SA results

H1c: The participants would neither perceive higher workload with improved-high-ecological than with improved-high-conventional. The participants would neither perceive higher workload with adaptive-ecological than with adaptive-conventional.

Hypothesis 1b and 1c remained unchanged from the previous experiment. However, since the high DOA was improved and the new adaptive configuration was used, the author expects the results of this experiment reveal different effects of DOA may be different in the results.

4.6.2 Research Hypotheses for Examining Risk Preference

H2: The participants could have different risk preferences with the ecological displays in comparison to the conventional displays.

A breakdown of hypothesis 2 would be:

H2a: The participants' fourfold experience-based choice in an automated trading environment would be different from that as identified in McAndrew and Gore's observations (2013), under the influence of DOA (moderate DOA and high DOA) and display (conventional or ecological).

H2b: In the losing state of the adaptive configuration, the mean position size of the participants' portfolio (S_p) with the ecological displays would be different with that with the conventional displays.

H2c: In the losing state of the adaptive configuration, the participants' decision preference in a guaranteed profiting situation with the ecological displays would be different with that with the conventional displays.

H2d: In the losing state of the adaptive configuration, the participants' decision preference in a guaranteed losing situation with the ecological displays would be different with that with the conventional displays.

Hypothesis 2b, 2c and 2d can only be meaningfully examined within the losing state of the adaptive configuration.

4.7 Results

4.7.1 Conventions

The data analysis process followed all conventions that have been previously used in Experiment 1. The conventions included significance level, effect size, and the measure of central tendency.

4.7.2 Data Analysis Scripts

The scripts used in Experiment 1 were modified to handle the Experiment 2 data. The modification took approximately four person-months to complete.

4.7.3 Summary of Results

Training data showed no evidence of confusions while completing the scenarios.

4.7.3.1 Task performance

End of scenario RPL (adaptive DOA)

End of scenario *RPL* was analyzed for adaptive DOA scenarios. End of scenario *RPL* for each participant was obtained from AUTRASS and was compared between two scenario types (adaptive-conventional or adaptive-ecological), $N = 24$. The assumption of normality was violated, $ps < .05$, therefore, a non-parametric test was used instead. Results of the Wilcoxon Signed test showed that the scenario type effect was not significant, $p > .05$. Empirical results shown in Table 34 suggested that the participants gained more end of scenario *RPLs* within the adaptive-ecological scenario.

Table 34. Summary of End of Scenario *RPL* for Adaptive DOA (Experiment 2).

Scenario Type	End of Scenario *RPL*			
	Mdn (Middle Most)	*M* (Arithmetic Mean)	*Mo* (Most Frequent)	*SD* (Deviation)
Adaptive-conventional	\$60.6	\$48.9	\$100	49.2
Adaptive-ecological	\$80.5	\$71.0	\$83	48.7

A follow-up analysis was performed to look at whether financial market data has become a confounding variable. The end of scenario *RPL* data were divided into two groups by the financial market data set they used (i.e., D1 or D2). Since the participants were randomly assigned with a financial market data set when they experienced the adaptive DOA scenarios, similar to Experiment 1, again the sample sizes of the two groups were unequal. Empirical analysis results showed that with D1, end of scenario *RPL* was slightly less with adaptive-ecological (*Mdn* = \$91.25) than with adaptive-conventional (*Mdn* = \$91.5). With D2, a reverse pattern was observed whereby end of scenario *RPL* was more with adaptive-ecological (*Mdn* = \$37) than with adaptive-conventional (*Mdn* = \$22). Since the end of scenario performance is subject to both trader's detection and mitigation performance and the trading algorithm performance, end of scenario *RPL* as a performance measure is arguably too robust.

Mean accumulating RPL (adaptive DOA)

An in-depth analysis was performed to understand the traders' performance when DOA shifted to moderate. Mean accumulating *RPL* of all losing states for each participant was calculated and subsequently compared between two scenario types (adaptive-conventional or adaptive-ecological), N = 24. The assumption of normality was violated, $ps < .05$. Results of the Wilcoxon Signed test showed that the scenario type effect was not significant, $p > .05$. Table 35 summarizes the descriptive statistics.

Table 35. Summary of Mean Accumulating *RPL* for Adaptive DOA (Experiment 2).

Scenario Type	Mean Accumulating *RPL*			
	Mdn (Middle Most)	*M* (Arithmetic Mean)	*Mo* (Most Frequent)	*SD* (Deviation)
Adaptive-conventional	-$12.5	-$58.125	-$5	111.5
Adaptive-ecological	-$15	-$36.25	-$5	46.9

The pattern was generally consistent with that of the Experiment 1 performance measure. Mean accumulating *RPL* data were divided into two groups by financial market data. For D1, the participants gained more shares in the adaptive-conventional scenario (*Mdn* = -$5) than with the adaptive-ecological scenario (*Mdn* = -$7.5). For D2, mean accumulating RPL was not empirically different (adaptive-conventional or adaptive-ecological: *Mdn* = -$47.5).

Mean duration of losing state (adaptive DOA)

Mean duration of losing state was caculated by counting how many timestamps the participants experienced DOA shifted to moderate in each adaptive DOA scenario, $N = 24$. The assumption of normality was violated, $ps < .05$. Results of the Wilcoxon Signed test showed that the scenario type effect was also not significant, $p > .05$, as shown in Table 36. The participants generally did not experience many DOA shifts during the scenarios. Only 66.7% of the participants have successfully made the trading system to be profitable (*RPL* > 0) after the initial mandatary loss by design, and therefore experienced a DOA shift from moderate to high. All participants have only experienced one DOA shift from high to moderate, which was the mandatory shift at the beginning of the scenario. Therefore, no detailed patterns of dynamic function allocations have been further studied.

Table 36. Summary of Mean Durations of Losing State for Adaptive DOA (Experiment 2).

Scenario Type	Mean Duration of Losing State (*Mdn*)
Adaptive-conventional	94 timestamps
Adaptive-ecological	86 timestamps

Fault detection accuracy (improved-high DOA and adaptive DOA)

Significant individual differences have been found in the distribution pattern of the profiting state in the adaptive-conventional scenarios and the adaptive-ecological scenarios. Four out of the twenty-four participants have experienced no trial throughout their scenarios, because they performed the flexible trading task poorly and the system remained at the losing stage after the initial shift from the higher DOA (also demonstrated in the long duration of losing state). These data were not analyzed statistically since not a single losing buy-sell pair occurred.

Data for improved-high-conventional and improved-high-ecological scenarios were separated from the raw data set and were analyzed statistically. Two participants did not report any losing buy-sell pairs in certain scenarios. To keep a balanced within-subject design, data for these two participants were excluded. Failure detection accuracy and response time to correctly detect a fault were analyzed similarly as the Experiment 1 (automated trading with implicit logic) analysis. The simulation data were divided over 2 (scenario type: improved-high-conventional or improved-high-ecological) × 3 (system state: profiting, neutral or losing) groups, $N = 22$. The assumption of normality was heavily violated, *ps* $< .05$. Non-parametric tests showed the scenario type effect was significant, $z = 2.004$, $p = .045$, $r = .427$, a medium effect. It can be seen from Figure 39 that the participants reported temporal automation failures in the form of losing buy-sell pairs significantly more accurately in the improved-high-ecological scenarios.

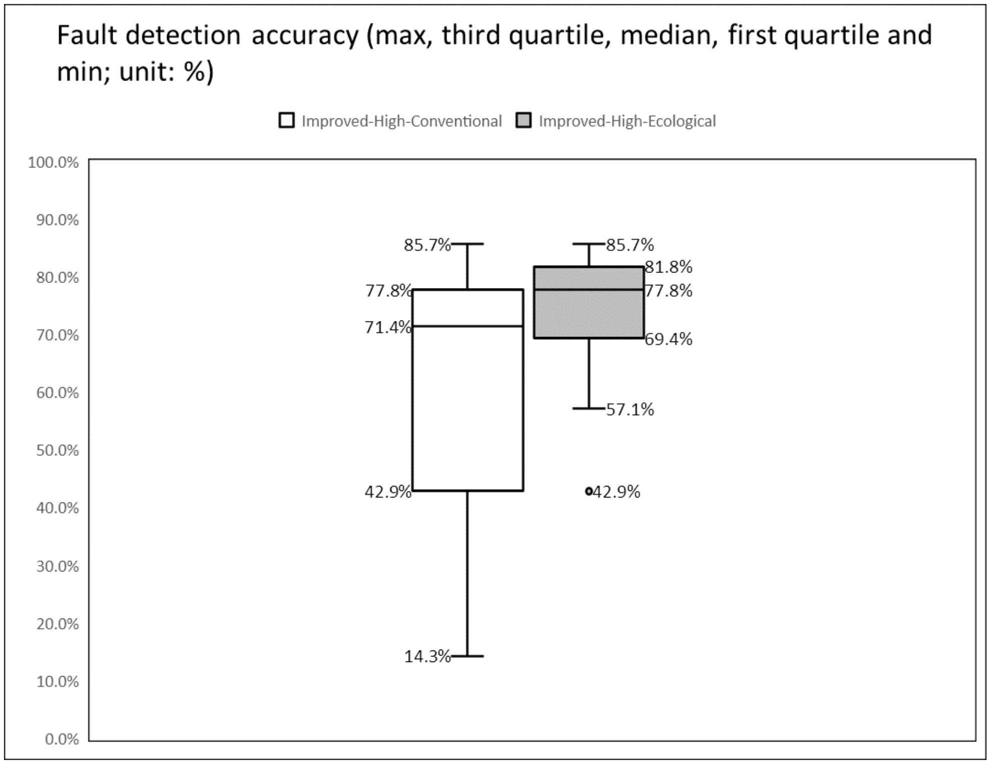


Figure 39. Fault detection accuracy for improved-high DOA: Scenario type effect (Experiment 2).

All participants have successfully completed the adaptive-conventional scenarios and the adaptive-ecological scenarios. Descriptive statistics for the profiting state of these scenarios suggested that the added visual elements of the ecological display may be an important attribute that distinguishes the performance of the fault detection task between the two scenario types, *N* = 24 (adaptive-conventional: *Mdn* = 60.0%, *M* = 56.7%, *SD* = .294; adaptive-ecological: *Mdn* = 80.0%, *M* = 61.3%, *SD* = .374). A summary of fault detection accuracies described in median and other central tendency measures for all configurations in Experiment 1 and 2 is presented in Table 37.

Table 37. Summary of Fault Detection Accuracies (Experiment 1 and 2 Combined).

Experiment	Scenario Type	Fault Detection Accuracy			
		Mdn (Middle Most)	*M* (Arithmetic Mean)	*Mo* (Most Frequent)	*SD* (Deviation)
Experiment 1	High-conventional	75.0%	73.2%	100%	.289
	High-ecological	100.0%	87.2%	100%	.213
Experiment 2mproved-high)E	Improved-high-conventional	71.4%	62.8%	77.8%	.227
	Improved-high-ecological	77.8%	73.0%	85.7%	.125
	Adaptive-conventional	56.7%	56.7%	60.0%	.294
	Adaptive-ecological	80.0%	61.3%	100%	.374

4.7.3.2 Situation awareness

The participants' responses to the SA queries were scored based on the simulation data in a way similar to Experiment 1. Abnormal responses were handled similarly. Table 38 shows the SA queries and the scoring rubrics.

Table 38. Summary of SA Queries and Scoring Rubrics (Experiment 2).

SA Level	SA Query	Answer Choice	Scoring Rubrics (Score 1)
1	"In the last 5 seconds, the market price has gone:"	Up, down or flat	Correctly compared P_m at t - 1 to t.
	"In the last 5 seconds, the slower moving average curve (yellow) has gone:"	Up, down or flat	Correctly compared SMA_{slow} at $t - 1$ to t.
	"In the last 5 seconds, the faster moving average curve (purple) has gone:"	Up, down or flat	Correctly compared SMA_{fast} at $t - 1$ to t.
	"In the last 5 seconds, was there a crossover of the two moving average curves?"	Yes or no	Correctly identified whether there was a SMA crossover during t - 1 to t.
	"In the last 5 seconds, ___ was taking the control in trading."	Automation or I	Correctly identified the operator (automation or the subject) during $t - 1$ to t.

SA Level	SA Query	Answer Choice	Scoring Rubrics (Score 1)
	"In the last 5 seconds, what was the most recent trade?"	Buy, sell or nothing	Correctly identified the trade (buying, selling or neither) during t - 1 to t.
2	"Is there a shift in control (e.g., you and automation)? Why?"	(Open-ended)	Correctly identified a shift or no shift between manual trading and automated trading during $t – 1$ to t.
	"What is happening with the unrealized profit and loss in your portfolio? Why?"		Correctly identify the positive or negative and direction *UPL* during timestamps $t – 1$ to t.
	"Is there a trading opportunity? What kind of opportunity, and why?"		Identified a buying opportunity if $P_p > P_m$ at the timestamp $_t$, or Identified a selling opportunity if $P_p < P_m$ and $Q_p > 500$ at the timestamp t, or Identified no trading opportunity in other cases.
	"What is happening with the quantity in your portfolio? Why?"		Correctly compared Q_p at $t – 1$ with Q_p at t.
	"Has the most recent trade made any realized profit, or loss? Why?"		Correctly compared P_m with P_p, if there was a selling trade between $t – 1$ and t, or Addressed if there was a buying trade or no trade.
3	"What will happen to the market price in the next 5 seconds? Why?"	(Open-ended)	Correctly compared P_m at t and Pm at $t + 1$ (predicted).
	"Do you think there will be a trading opportunity in the next 5 seconds? What kind of opportunity, and why?"		Predicted a buying opportunity if $P_m < P_p$ at $t + 1$,

SA Level	SA Query	Answer Choice	Scoring Rubrics (Score 1)
			or
			Predicted a selling opportunity if P_m > Pp and Q_p > 500 at $t + 1$, or
			Predicted no trading opportunity in other cases.
	"Do you think there will be a trade in the next 5 seconds? What kind of trade, and why?"		Correctly predicted the trade (buying, selling or neither) during t to $t + 1$.
	"What will happen to the quantity in your portfolio in the next 5 seconds? Why?"		Correctly compared Q_p at t and Q_p at $t + 1$ (predicted).

All participants successfully completed the SA queries. Mean SA score data were divided into 4 (scenario type: improved-high conventional, improved-high ecological, adaptive conventional or adaptive ecological) × 3 (SA level: 1, 2 or 3) within-subject groups, $N = 24$. The assumption of normality was violated in all groups, $ps < .05$. Non-parametric tests were performed. Results of the Friedman's test showed that there was a SA level significant effect, $\chi^2 = 23.239$, $p < .001$ (Level 1: $Mdn = .625$; Level 2: $Mdn = .438$; Level 3: $Mdn = .375$). Post hoc tests showed there was a significant difference between Level 1 and 2 SAs ($z = 4.022$, $r = .821$, a large effect) and a significant difference between Level 1 and 3 SAs ($z = 3.634$, $r = .742$, a large effect), $ps < .001$. The difference between Level 2 and 3 SAs was not significant, $p > .05$. The scenario type effect was not significant, $p > .05$. Descriptive statistics for the mean SA scores are summarized in Table 39.

Table 39. Summary of Mean SA Scores (Experiment 2).

Scenario Type	Per Scenario Type (*Mdn*)	Per SA Level (*Mdn*)		
		1	2	3
Improved-high-conventional	.458	.750	.375	.333
Improved-high-ecological	.500	.500	.500	.446
Adaptive-conventional	.500	.500	.500	.385
Adaptive-ecological	.486	.500	.500	.316

4.7.3.3 Eye tracking

Total dwell time (market price AOI, portfolio AOI and trading history DOA)

Eye tracking data for three participants were excluded from the analysis due to data losses. Total dwell time for the 5-second window prior to such SA pause in each scenario were calculated and submitted to the statistical model. The assumption of normality was violated in most of the 4 (scenario type: improved-high-conventional, improved-high-ecological, adaptive-conventional or adaptive-ecological) × 3 (AOI: market, portfolio, trading history) groups, $ps < .05$, $N = 21$. A repeated measures ANOVA was performed on the log 10 transformation eye-tracking data. The assumption of Sphericity was met, $p > .05$. There was a significant AOI main effect, $F(2, 40) = 12.989$, $p < .001$, $\eta^2 = .135$ (a medium effect). No other effects were significant, $ps > .05$. The AOI effect plot is presented in Figure 40.

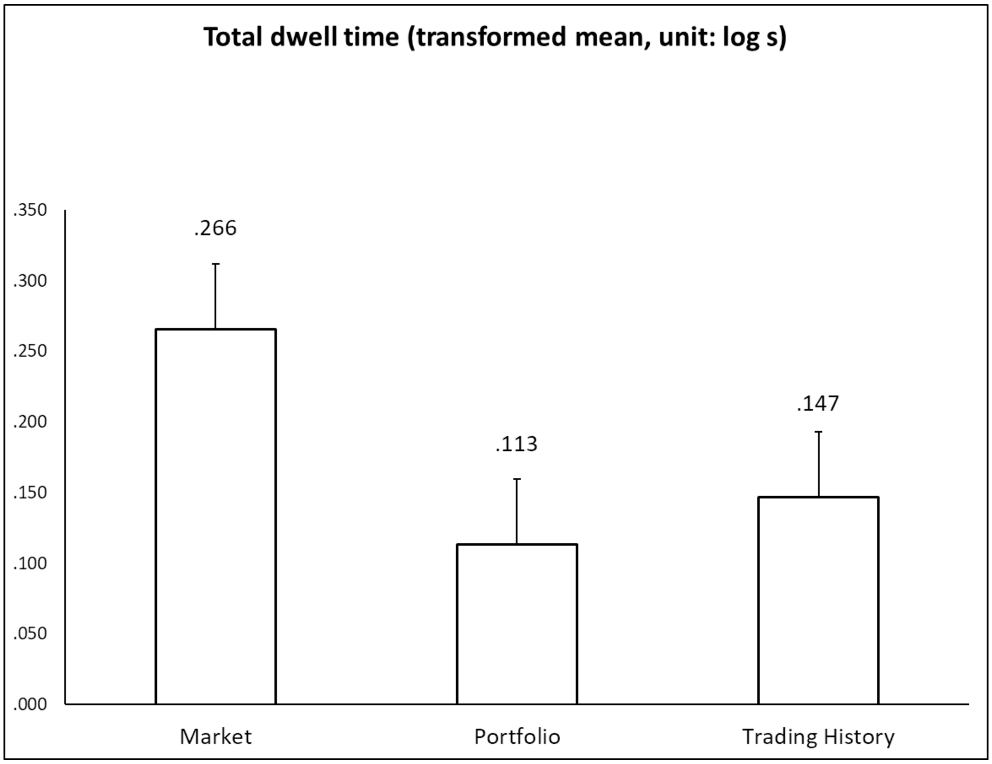


Figure 40. Total dwell time for the common AOIs: AOI interaction (Experiment 2).

Post hoc tests were performed using Bonferroni corrections. Total dwell time in each AOI was compared between the four scenario types. Total dwell time was significantly longer on the market AOI than on the portfolio AOI, in the case of improved-high-ecological (p = .046), adaptive-conventional (p = .001) and adaptive-ecological (p = .045). Total dwell time was significantly longer on the market AOI than on the trading history AOI only with improved-high-conventional (p = .034). There was no significant difference between the portfolio AOI and the trading history AOI, ps > .05.

A summary of total dwell time for all the common AOIs is presented in Table 40.

Table 40. Summary of Total Dwell Time for the Common AOIs (Experiment 2).

Scenario Type	Area of Interest (*M*, *SD*, unit: log s)		
	Market	Portfolio	Trading History
Improved-high-conventional	.287 (.214)	.156 (.180)	.117 (.158)
Improved-high-ecological	.260 (.182)	.124 (.162)	.124 (.162)
Adaptive-conventional	.299 (.199)	.099 (.110)	.157 (.150)
Adaptive-ecological	.218 (.188)	.075 (.133)	.188 (.176)

Total dwell time (market-portfolio-execution AOI)

Total dwell time for the market-portfolio-execution AOI with the ecological displays were log transformed and statistically compared between two scenario types, improved-high-ecological and adaptive-ecological, using a paired *t* test, $N = 21$. The scenario type effect was not significant, $p > .05$, as shown in Table 41.

Table 41. Summary of Total Dwell Time for the Market-Portfolio-Execution AOI (Experiment 2).

Scenario Type	Area of Interest (*M*, *SD*, unit: log s)
	Market-Portfolio-Execution
Improved-high-ecological	.135 (.175)
Adaptive-ecological	.143 (.135)

Total dwell time (states-task AOI)

Total dwell time for the states-task AOI was log transformed and submitted to a paired *t* test to examine the scenario type effect (adaptive-conventional or adaptive-ecological), $N = 21$. The scenario type effect was not significant, $p > .05$, as shown in Table 42.

Table 42. Summary of Total Dwell Time for the States-Task AOI (Experiment 2).

Scenario Type	Area of Interest (*M*, *SD*, unit: log s)
	States-Task
Adaptive-conventional	.172 (.177)
Adaptive-ecological	.168 (.175)

4.7.3.4 Workload

All participants successfully completed the paper-based workload questionnaires. Unweighted NASA TLX ratings were submitted to a 4 (scenario type: improved-high or adaptive) × 6 (NASA TLX subscales: mental workload, physical workload, temporal workload, performance, frustration or effort) robust repeated measures ANOVA, $N = 24$. The assumption of sphericity was violated in the scenario type × subscale interaction, $p < .05$. A Greenhouse-Geisser correction was applied to the statistical model.

There was a significant subscale effect, $F(3.441, 115) = 12.647$, $p < .001$, $\eta^2 = .193$ (Figure 43). Post hoc test results are presented in Table 43. The pattern in the data is generally consistent with that in the Experiment 1 data. The mental demand was moderate, and the physical demand was relatively low. All remaining effects in the statistical model were not significant, *ps* > .05.

A summary of NASA TLX ratings is presented in Table 43.

Table 43. Summary of NASA TLX Ratings (Experiment 2).

Scenario Type	Subscale (*M*, *SD*)					
	Mental Demand	Physical Demand	Temporal Demand	Performance	Frustration	Effort
Improved-high-conventional	12.75 (5.007)	5.417 (4.442)	11.000 (4.978)	9.625 (4.707)	8.208 (4.191)	11.208 (4.961)
Improved-High-ecological	11.959 (4.196)	5.125 (4.047)	8.75 (5.152)	9.250 (5.227)	7.792 (5.099)	10.542 (4.433)
Adaptive-conventional	12.833 (3.784)	6.167 (4.697)	10.750 (4.513)	9.458 (5.013)	9.375 (5.747)	12.958 (4.486)
Adaptive-ecological	14.042 (4.704)	6.000 (4.314)	11.167 (5.427)	10.375 (3.954)	8.208 (4.191)	11.208 (4.961)

4.7.4 Risk Preference

4.7.4.1 Fourfold pattern of preferences

Fourfold pattern of preferences, the qualitative measure of risk preference, was analyzed. All participants completed this measure, *N* = 24. The author summarized the percentage of the participants who made risk-seeking or risk-aversion preference after they experienced each measurement scenario. Table 44 shows the participants' choices in each scenario type. It can be concluded that most participants were risk seeking with medium- and large-probability gains (61 out of 96) and small-probability losses (69 out of 96), and risk aversion with small-probability gains (67 out of 96). These patterns were generally consistent across the four scenario types. No consistent pattern of choices across of the four scenario types has been found with in the case of medium- and large-probability losses (risk seeking: 45 out of 96; risk aversion: 51 out of 48).

Table 44. Fourfold Pattern of Preferences: Scenario Type Effect (Experiment 2).

Medium- and large-probability gains

	Improved-High-Conventional	Improved-High-Ecological	Adaptive-Conventional	Adaptive-Ecological
Risk seeking	<u>17</u>	<u>16</u>	<u>17</u>	<u>12</u>
Risk aversion	7	8	7	12

Small-probability losses

	Improved-High-Conventional	Improved-High-Ecological	Adaptive-Conventional	Adaptive-Ecological
Risk seeking	<u>21</u>	<u>16</u>	<u>15</u>	<u>17</u>
Risk aversion	3	8	9	7

Small-probability gains

	Improved-High-Conventional	Improved-High-Ecological	Adaptive-Conventional	Adaptive-Ecological
Risk seeking	10	11	8	11
Risk aversion	<u>14</u>	<u>13</u>	<u>16</u>	<u>13</u>

Medium- and large-probability losses

	Improved-High-Conventional	Improved-High-Ecological	Adaptive-Conventional	Adaptive-Ecological
Risk seeking	12	<u>13</u>	<u>13</u>	11
Risk aversion	12	11	11	<u>13</u>

4.7.4.2 Mean portfolio's size of portfolio (adaptive DOA)

Mean position size was calculated the losing state of each adaptive automation scenario. The data were divided over two groups by scenario type (adaptive-conventional or adaptive-ecological). The assumption of normality was violated in either group, *ps* < .05. Wilcoxon signed test results showed that the scenario type effect was not significant, $p > .05$ (adaptive-conventional: *Mdn* = 468.5 shares; adaptive-ecological: *Mdn* = 660.9 shares). Since the automation performance was a mediating factor in deciding the portfolio's size, no further analysis similar to Experiment 1 was performed.

4.7.4.3 Decision preference in a guaranteed profiting situation (adaptive DOA)

To help the trading system recover from the losing state, the participants must profit from guaranteed profiting opportunities when the system was in the losing state of the adaptive configuration. The participants may perform an immediate sell-off in the shares of SPY (risk-aversion) or to wait for future profiting opportunities by holding the portfolio (moderate risk-seeking) or buying more shares of SPY (high risk-seeking). The percentage of executions was calculated in each of the 2 (scenario type: adaptive-conventional or adaptive-ecological) × 3 (risk level: risk-aversion, moderate risk-seeking or high risk-seeking) groups, $N = 24$. The system state effect (profiting, neutral or losing) was not analyzed because the participants only perform this task during the losing state. The assumption of normality was violated in 2 × 3 all groups, $ps < .05$. The scenario type effect was not significant, $p > .05$ (adaptive-conventional: *Mdn* = 33.3%; adaptive-ecological: *Mdn* = 33.3%). There was a significant risk level effect, $\chi^2 = 42.000$, $p < .001$ (risk-aversion: *Mdn* = 10.2%; moderate risk-seeking: *Mdn* = 78.6%; high risk-seeking: *Mdn* = 5.8%). All pairwise differences between the three risk levels were significant, $ps < .05$. The pairwise differences can be interpreted as the participants were most likely to perform moderate risk-seeking decisions in a guaranteed profiting situation, followed by risk-aversion decisions than high risk-seeking decisions, as presented in Figure 41.

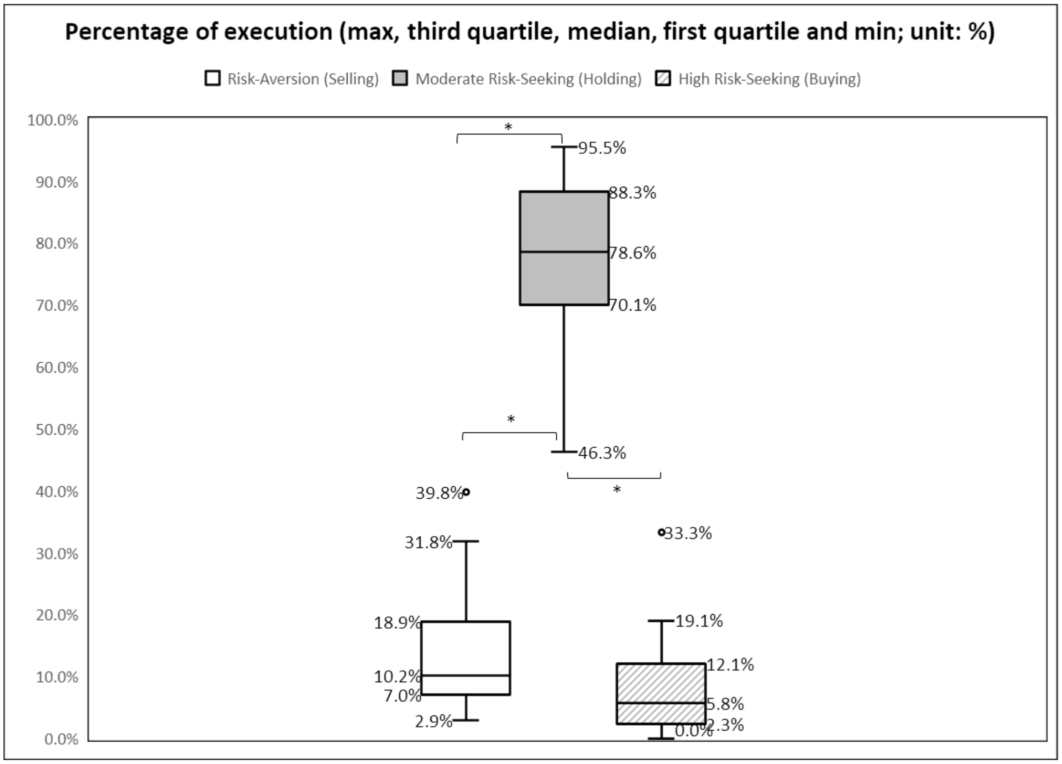


Figure 41. Percentage of execution: Risk level effect (Experiment 2).

To examine whether the type of display can influence the percentage of executions at each risk level, the 2 × 3 statistical model was collapsed to three models representing these risk levels. Non-parametric tests were performed for each level:

1. For the risk-averse risk level, the percentage of selling executions was analyzed. The scenario type effect was not significant, $p > .05$ (adaptive-conventional: *Mdn* = 9.5%; adaptive-ecological: *Mdn* = 12.6%).

2. For the moderate-risk seeking risk level, the percentage of holding executions made was analyzed. The scenario type effect was significant, $p = .022$, $r = .467$, a large effect (adaptive-conventional: *Mdn* = 89.4%; adaptive-ecological: *Mdn* = 78.3%). Figure 42 shows that the participants performed significantly less holding executions with the adaptive-ecological scenario.

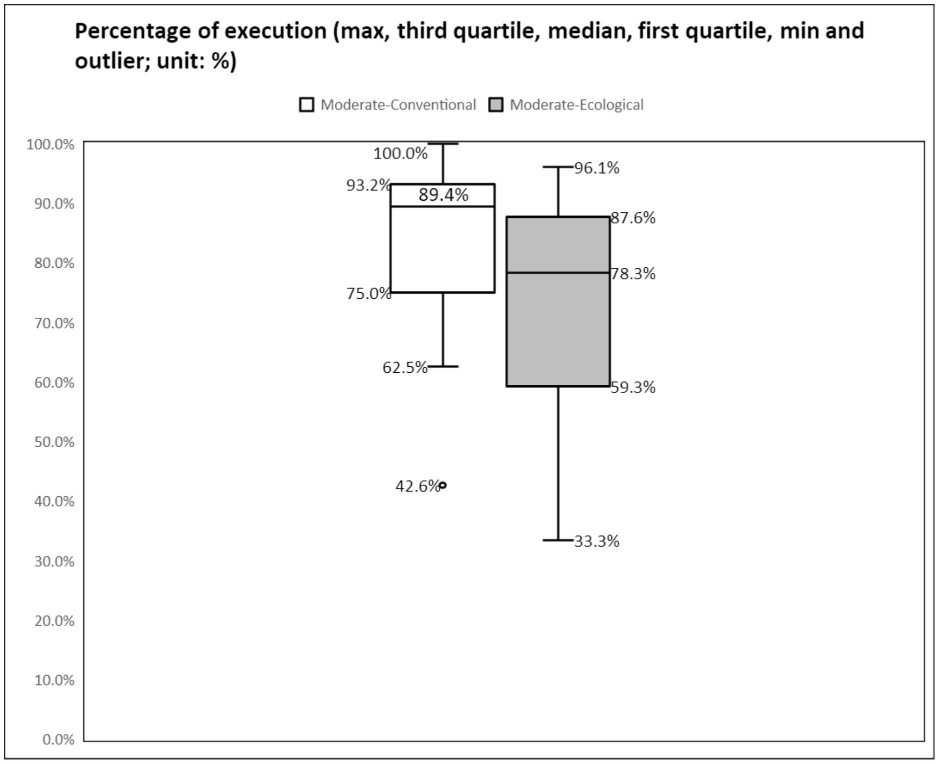


Figure 42. Percentage of holding executions: Scenario type effect (Experiment 2).

3. For the high-risk seeking risk level, the percentage of buying executions was analyzed. The scenario type effect was not significant, $p > .05$ (adaptive-conventional: *Mdn* = 1.7%; adaptive-ecological: *Mdn* = 5.7%). Table 45 summarizes the median percentages of execution divided by scenario type and risk level.

Table 45. Percentage of Executions (Adaptive-Conventional or Adaptive-Ecological, Experiment 2)

Scenario Type	Percentage of Execution (*Mdn*)		
	Risk-Aversion	Moderate Risk-Seeking	High Risk-Seeking
Adaptive-conventional	9.5%	89.4%	1.7%
Adaptive-ecological	12.6%	78.3%	5.7%

4.7.4.4 Decision preference in a guaranteed losing situation (adaptive DOA)

The percentage of executions in each guaranteed losing situation was calculated in each of the 2 (scenario type: adaptive-conventional or adaptive-ecological) × 3 (risk level: risk-aversion, moderate risk-seeking or high risk-seeking) groups, $N = 24$. The scenario type effect was not significant, $p > .05$ (adaptive-conventional: *Mdn* = 33.3%; adaptive-ecological: *Mdn* = 33.3%). The risk level effect, however, was significant, $p < .001$. All pairwise differences between the three risk levels were significant, *ps* < .05. As Figure 44 shows, the participants were generally moderate risk-seeking while facing a guaranteed losing situation (*Mdn* = 85.5%), followed by high risk-seeking (*Mdn* = 10.0%) then risk averse (*Mdn* = 4.5%). Unlike the previous analysis performed on guaranteed profiting situations, the current result did not elaborate the qualitative results, Figure 43.

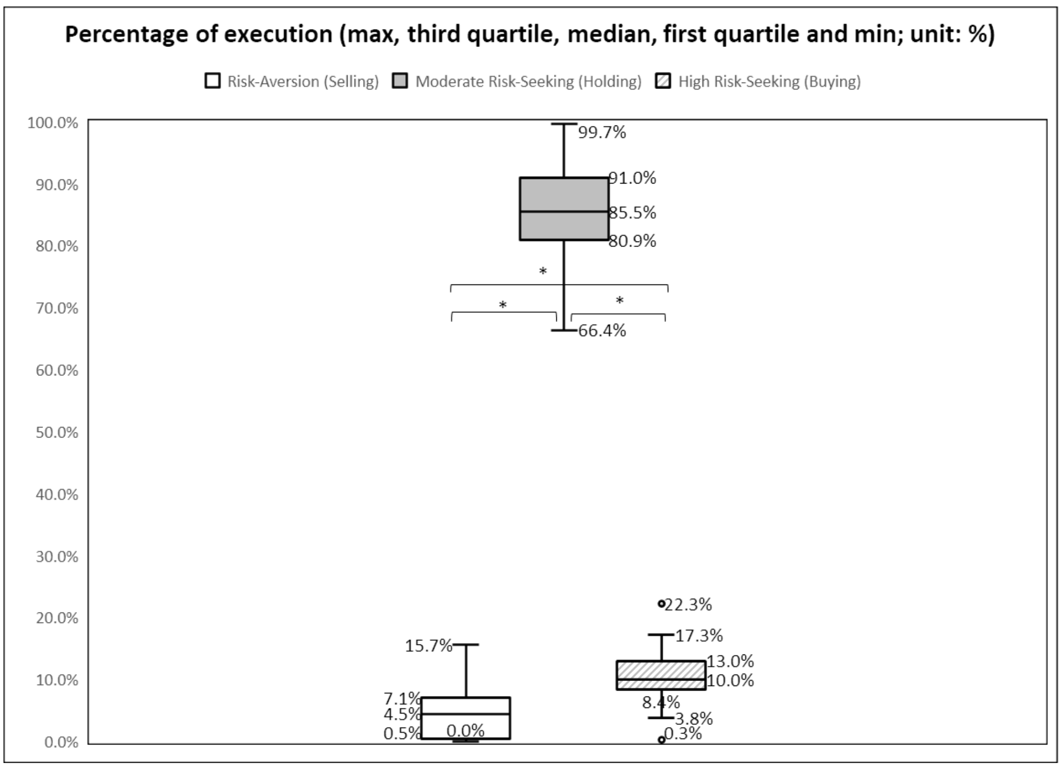


Figure 43. Percentage of execution: Risk level effect (Experiment 2).

1. For risk-aversion, the percentage of selling executions was analyzed. The scenario type effect was not significant, $p > .05$ (adaptive-conventional: *Mdn* = 3.3%; adaptive-ecological: *Mdn* = 4.3%).

2. For moderate risk-seeking, the percentage of holding executions was analyzed. The scenario type effect was not significant, $p > .05$ (adaptive-conventional: *Mdn* = 84.6%; adaptive-ecological: *Mdn* = 87.4%).

3. For high risk-seeking, the percentage of buying executions was analyzed. The scenario type effect was not significant, $p > .05$ (adaptive-conventional: *Mdn* = 11.8%; adaptive-ecological: *Mdn* = 7.7%).

4.8 Discussion

The discussion focused on interpreting Experiment 2 results but also made several comparisons to the results of Experiment 1.

4.8.1 Performance

4.8.1.1 Task performance

Performance degradation with ecological displays due to risky actions?

During the losing state of the adaptive configuration, mean accumulating *RPL* replaced end of scenario *RPL* as the measure of task performance. However, it was not significantly different with the ecological displays compared to the conventional displays, which is consistent with the end of scenario *RPL* results of Experiment 1. Therefore, hypothesis 1a was rejected. The detection-mitigation confound in the task as identified in Experiment 1 is a possible explanation. However, there might be another explanation.

A comparison of Experiment 1 and 2 results is presented in Table 46. This comparison empirically suggested that the ecological displays may have degraded the participants' performance in the flexible trading task, which can also be characterized as overall trading performance. Almost all participants have not been professionally trained (with one exception in Experiment 1 and another one in

Experiment 2). Therefore, there were likely to be huge individual differences in the choices of trading strategies and their signal detection and mitigation abilities, all of which are crucial in financial trading. The consistent pattern in the mean *RPL* results suggested a possible performance degradation with the ecological displays, which may be due to the risk-seeking behaviours fostered by the ecological displays, as demonstrated in the results of the quantitative risk preference measures. This explanation is merely hypothetical and cannot be statistically analyzed with the current experimental setting. Future research is warranted on this topic to recruit professionally-trained traders and examine how they performed the flexible trading task in an experimental setting similar to AUTRASS.

Table 46. Summary of Moderate Task Performance (Experiment 1 and 2 Combined).

Experiment	Display	End of Scenario *RPL (Mdn)*
Experiment 1	Moderate-conventional	$35
	Moderate-ecological	$30.5
Experiment 2	Adaptive-conventional	-$12.5
	Adaptive-ecological	-$15.0

Fault detection accuracy: effective ecological display support with improved automation design

Results of the fault detection accuracy in the improved-high DOA configuration supported hypothesis 1a. The ecological display effectively supported the fault detection task in unanticipated situations. Fault detection accuracy was compared amongst the original and the improved-high DOA configurations and the adaptive configuration using a wide range of central tendency measures. Results of the comparison are presented in Table 47.

Table 47. Summary of Fault Detection Accuracies (Experiment 1 and 2 Combined).

Experiment	Display	Fault Detection Accuracy			
		Mdn (Middle Most)	*M* (Arithmetic Mean)	*Mo* (Most Frequent)	*SD* (Deviation)
Experiment 1Expe	High-conventional	75.0%	73.2%	100%	.289
	High-ecological	100.0%	87.2%	100%	.213
Experiment 2Experi)	Improved-high-conventional	71.4%	62.8%	77.8%	.227
	Improved-high-ecological	77.8%	73.0%	85.7%	.125
	Adaptive-conventional	56.7%	56.7%	60.0%	.294
	Adaptive-ecological	80.0%	61.3%	100%	.374

* Significant difference.

No evidence of a ceiling effect was observed in the improved-high DOA configuration data, though the participants performed the fault detection task generally well as demonstrated in the mode data. The ecological display effect was not examined statistically in the profiting state of the adaptive configuration due to the limited sample size. However, a consistent pattern can be found in the adaptive configuration.

Since the same ecological display was used in both the original and the improved-high DOA configurations, the significance in the type of display is most likely to be a direct result of the improvement of automation design. With appropriate training on the algorithm logic and features provided by the ecological display, a mental model can be established to utilize the problem-solving support provided by the states-task visualization. On the other hand, the procedural support provided in the states-task visualization ensured a consistent rule-based mapping between the cues and the automation performance. Since no knowledge-based reasoning was required to use the procedural support, the rule-based mapping should not be influenced by the unfamiliar algorithm logic. However, with the effect of the ecological display significant in the improved-high DOA configuration in this experiment but not significant in the original high DOA configuration in the previous experiment, it seems that the misconnection between different stages of automation brings more complexity to the problem space and may detract from the participants' ability to utilize the rule-based mapping. With a higher DOA, the participants generally had less task involvement and may be more vulnerable to the loss of SA. Therefore, it is possible that in situations where the automation logic is unfamiliar, operators may not be able to effectively use the rule-based support without a well-established knowledge-based understanding.

Fault detection accuracy is a customized task performance measure for monitoring in a high DOA automated trading setting that involves sophisticated algorithm operations. It is noticeably different from other task performance measures that have been examined in the literature. For example, H. Li (2013) simulated a space mission to investigate a number of performance consequences measures that were suitable for a wide range of domains. Her measures included task completion time, accuracy and hazard occurrence. While accuracy has been adopted as a performance measure in Experiment 1 and 2, task completion time and occurrence of real hazards were not used in the data analysis of this dissertation due to two distinctions in the automated trading setting. First, AUTRASS simulated a financial trading system which receives and presents market data in discrete time series and the trading executions issued by the trader, or the automation must proceed in a time-by-time manner. The response time to correctly reported buy-sell pairs was not logged by AUTRASS due to the disruption of the latency. Task completion time is also subject to the latency in order processing and does not always reflect the operator performance. Second, although this experiment tried to simulate automation failure which was rare with the well-performing trading algorithm, more hazardous events (known as the "black swans", e.g., market crash) in a real market are much rarer. Considering how rare these hazardous events may occur, to practically simulated these hazards, a much longer scenario or a longitudinal study should be considered in the design of future studies. The occurrence of hazards is a realistic measure comparable to many fault detection measures that have been documented in the literature (e.g., perceived urgency of hazards: Arrabito et al., in review; number of control actions taken: Reising & Sanderson, 2000b). Further development of this measure for the automated trading setting is warranted.

4.8.1.2 Situation awareness

Participants attending Experiment 2 had moderate SA (*Mdn* = .500) similar to those participating in Experiment 1 (*Mdn* = .500). The participants had a highest level 1 SA, followed by the level 2 SA then the level 3 SA, which generally reflected the definition of SA. No display type or DOA effect was observed. As a result, hypothesis 1b was rejected.

It can be concluded that no solid SA improvement has been found throughout Experiment 2 according to the results of SAGAT. The eye-tracking data represented a similar pattern in comparison to the moderate DOA configuration with the conventional displays in Experiment 1 data. With no effect of the ecological displays observed, it seems that adaptive automation is a new concept in the context of financial trading, and future studies are necessary to identify specific design requirements for this type of automation.

4.8.2 Workload

Neither DOA nor the type of display significantly influenced workload as predicted in hypothesis 1c. The ecological display did not add additional workload to the trading tasks.

4.8.3 Risk Preference

Results of the qualitative risk preference measures matched those of Experiment 1. It can be seen in Table 48 that the fourfold patterns of choice presented in both experimental studies were consistent with those identified in Hertwig and Erev's study in the case of small-probability gains and losses. In the case of medium- and large-probably gains, the fourfold patterns of choice were consistent with those identified in McAndrew and Gore's study. Similar to the Experiment 1 results, no consensus was found in the risk preference of medium- and large-probability losses. The consistent pattern between Experiment 1 and 2 demonstrated that the fourfold pattern of choice was probably determined by the overall simulated trading environment, not the specific tasks or the experimental conditions, Table 48.

Table 48. Summary of Fourfold Patterns of Choice (Experiment 1 and 2 Combined).

	Description (Prospect Theory)		Experience Through Learning (Hertwig & Erev, 2009)		Experience Through Professional Training (McAndrew & Gore, 2013)		Experiment 1		Experiment 2	
	Gains	Losses	Gains	Losses	Gains	Losses	Gains	Losses	Gains	Losses
Small probability	Risk seeking	Risk aversion	Risk aversion	Risk seeking	Risk aversion	Risk aversion	Risk aversion	Risk seeking	Risk aversion	Risk seeking
Medium and large probability	Risk aversion	Risk seeking	Risk aversion?	Risk seeking?	Risk seeking	Risk aversion	Risk seeking	(No consensus)	Risk seeking	(No consensus)

Ecological displays and risk-seeking actions

The mean portfolio size examined in the losing state of the adaptive configuration has a similar pattern in comparison to the results of the moderate DOA configuration (Table 49), which again, supported Borst et al. (2015) that ecological displays may be prone to risky actions if the intentional constraints of the work domain were not made visible. Since the market-portfolio-execution visualization merely represents the physical constraints, with a goal of achieving a maximum rate of revenue, no intentional constraints were represented in this visualization.

Table 49. Summary of Mean Portfolio's Size (Experiment 1 and 2 Combined).

Scenario Type	Mean Portfolio's Size (*Mdn*)
Moderate-Conventional	618.0
Moderate-Ecological	904.3
Adaptive-Conventional	468.5
Adaptive-Ecological	660.9

A comparison of the decisions made in guaranteed profiting situations between the moderate automation scenarios and the losing state of the adaptive automation scenarios showed that the participants were generally moderate risk-seeking with AUTRASS, as being summarized in Table 50.

Table 50. Summary of Percentage of Execution (Experiment 1 and 2 Combined).

Scenario Type	Risk Level	Percentage of Execution (*Mdn*)
Moderate	Risk aversion	22.4%
	Moderate risk-seeking	55.3%
	High risk-seeking	21.9%
Adaptive	Risk aversion	10.2%
	Moderate risk-seeking	78.6%
	High risk-seeking	5.8%

In this experiment, with the adaptive configuration, the participants performed significantly less holding executions that were associated with moderate risk-seeking with the ecological display than with the conventional display (adaptive-conventional: *Mdn* = 89.4%; adaptive-ecological: *Mdn* = 78.3%). This pattern was not empirically identical with that of Experiment 1 where the participants experienced the moderate DOA configuration (adaptive-conventional: *Mdn* = 51.2%; adaptive-ecological: *Mdn* = 51.5%). My speculation is that with adaptive automation being used, the influence of the ecological displays on risky actions was limited to a certain extent, and the participants preferred taking extreme actions (i.e., either high risk-seeking or risk-aversion actions). This speculation, along with the impact of the missing support for the intentional constraints, require more in-depth explorations in future.

Here the author provides a final remark on the development of risk preference measures. The quantitative measures make a distinct contribution toward the description- and experience-based choice research. AUTRASS provided the technological foundation for the development of quantitative measures for evaluating risk preference. These measures not only supplemented the qualitative measures in identifying the patterns of choice for medium- and large-probabilities which were largely underexplored (McAndrew & Gore, 2013), but also suggested that risk seeking for medium- and large-probability gains may be attributable to the heuristic cues provided by the ecological displays, and thus connected the CWA/EID approach to the prospect theory/description- and experience-based choice research.

4.8.4 Other Findings

4.8.4.1 A DOA layering approach to display design

The distinct statistical results of fault detection performance in the original and the improved-high DOA configurations demonstrated the applicability of the DOA layering approach to designing ecological displays. Based on several aspects of the DOA-layered models, the ecological displays provided different forms of support according to the DOA. The market-portfolio-execution visualization has been developed in a way similar to how ecological displays have been developed in the literature. Built on the base AH, this visualization did not address the DOA and, similar to most ecological displays, may not support a specific DOA situation.

The success of the ecological displays in supporting fault detection performance in the improved-high DOA configuration has suggested important implications for using information on a DL to derive ecological displays that were specific to a DOA, which was an underexplored area. Rasmussen pointed out that that in information-processing systems the operator is connected to the environment in various ways, and the connections can be organized into three categories according to the SRK taxonomy. These categories are skill-based, rule-based and knowledge-based. Although most EID applications in the literature only described constraints that were inherent in the work domain, there may be other constraints imposed by certain control tasks as suggested by Vicente and Rasmussen in their early development of EID (1992). Bennett and Flach's work has provided some practical guidance for mapping DL functions to skill-, rule-, and knowledge-based information-processing. They described that skill-based processing reflects the direct links at the bottommost of the DL, between activation and execution. Rule-based processing utilizes the middle region of the DL, including observe, system state, goal state, formulate procedure and the various states of knowledge developed by these functions. Knowledge-based processing, however, is not directly supported by most perceptual cues and the operators must analyze the situation and develop a solution. The design of the states-task visualization is an initial attempt to design based on Bennett and Flach's mapping and requires further development in the future.

4.8.4.2 A DOA layering approach to automation design

The two experimental studies added a practical caveat to automation design. The presented design exercise included determining initial function allocation (i.e., stages and levels of automation), evaluating automation design according to evaluative criteria (e.g., dependent variables in the two experimental studies) and modifying automation design, and demonstrated the potentials of the DOA layering approach. With this approach fitted into the existing automation design framework (Parasuraman et al., 2000), CWA can be performed at an early phase of automation design. Further, ecological displays and automation can be designed concurrently, based on the rich information provided by the DOA layered models. Future research should extend the DOA layering approach as an automation design approach.

4.9 Chapter Summary and Connections to Research Questions

4.9.1 Key Findings

This chapter elaborates on the findings in chapter 6 and suggests new findings.

Automation design and EID: This chapter further explores the effectiveness of the ecological displays in supporting a specific DOA. With the improved automation design, task performance at a higher DOA was supported. It can be concluded that in cases where the knowledge-based support did not provide necessary help, the rule-based support alone may not be able to support the detection performance.

Adaptive automation, moderate risk-seeking, and EID: With the adaptive automation, although the participants were moderate risk-seeking, the participants were significantly less moderate risk-seeking with the ecological displays, suggesting a limit on the influence of ecological displays on risky actions.

4.9.2 Connections to Research Questions

The author has three research questions for this dissertation:

Research Question 1: How can we model automated trading systems with a variable DOA using CWA?

Research Question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research Question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

This chapter is a further exploration of research question 2 and 3. Both research questions are partially supported. Connections to the first research question are also addressed.

Part D

Conclusion

This last part of the dissertation is a conclusion. In chapter 5, the author summarizes key findings and contributions and suggest areas for future work.

Chapter 5

Conclusion

Automated trading is an underexplored domain in the human factors research, and it involves a wide range of research topics. A thorough literature review was performed on automated trading, human-automation interaction and CWA and was documented in Appendix A. This dissertation narrowed these topics down to 3 research questions which were proposed in chapter 1. The rest of this dissertation adopted a two-phase approach to explore automated trading. First, a DOA layering approach was proposed in chapter 2 as the theoretical foundation to investigate the complexity relevant to the variable DOA of automated trading. Second, design concepts implied by the DOA layering approach were used to develop the automation and the ecological displays that were expected to support the variable DOA situation in the automated trading domain. Some DOA configurations, including adaptive automation, were evaluated through two experiments with novice participants recruited from a university population (chapter 3 and 4).

To conclude this dissertation, this chapter reviews key findings and contributions, and suggest future work.

5.1 Summary of Key Findings

This summary is structured according to the three research questions with findings included under each question.

5.1.1 Model Automated Trading Systems with a Variable DOA Using CWA

Automation in financial trading is versatile regarding the various stages and levels of automation involved. The DOA layering approach was proposed during an early investigation of this domain to characterize the complexity related to the variable DOA. With the stages and levels of automation model adopted in CWA, the DOA layering approach suggested new opportunities for designing automation and ecological displays.

5.1.2 Ecological Interface Design for Supporting Performance in Financial Trading

AUTRASS is the first to create a simulation of trend following trading that is typical in automated trading systems. Automation and ecological displays were derived from CWA models and layers following the DOA layering approach. The two automation configurations and the conventional and the ecological displays together formed four scenario types – moderate-conventional, moderate-ecological, high-conventional and high-ecological. Experiment 1 examined these four scenarios types, two of which had a higher DOA than the other two due to less participant task involvement. Experiment 2 examined the same conventional and ecological displays with two different DOA configurations in four new scenario types – improved-high-conventional, improved-high-ecological, adaptive-conventional and adaptive-ecological. Among them, improved-high-conventional and improved-high-ecological were similar to the high-conventional and the high-ecological scenario types evaluated in Experiment 1 with the automation design improved to be better supported by the ecological displays. Adaptive-conventional and adaptive-ecological were two adaptive automation conditions, the first of this kind in an automated trading setting.

There was no strong statistical evidence on how ecological displays could influence trading performance in a moderate automation configuration in which the traders were largely involved in making decisions and performing actions. However, descriptive statistics showed a possible pattern of worse trading performance with the ecological displays in the moderate-ecological scenario, as the traders were prone to taking risky actions with such displays. In a higher DOA situation, similar to supervisory control where the traders performed monitoring on a trading algorithm, the ecological display significantly improved the fault detection performance only if both rule- and skill-based processing was supported by the displays. The fault detection performance was not fully supported by the ecological display with only the skill-based processing was supported. This finding should inspire new directions for developing future ecological displays that can better support ecological displays based on the proposed DOA layering approach.

The empirical data showed that an automation trade-off that was typical with automated systems has been found in the financial trading setting: with a higher DOA, traders' perceived workload was likely to be lower and their SA was likely to be degraded. With the adaptive automation, the workload was not degraded, but neither the SA was improved. The current ecological displays that have been evaluated in the two experimental studies did not seem to influence this trade-off.

5.1.3 Ecological Interface Design to Influence Trader's Risk Preference

Results of the two experimental studies demonstrated a fourfold pattern of preferences which was partially consistent with the findings of the experience-based choice research. Quantitative measures developed in this dissertation provided new ways to evaluate risk preference in a simulation of a real-world environment. Ecological displays imposed risky actions with a moderate DOA configuration closest to manual control, suggesting that risk preference should be taken into account by ecological display designers. In the adaptive automation setting, although the participants were moderate risk-seeking, they were less moderate risk-seeking with the ecological displays. Therefore, there may be a limit on the influence of ecological displays on risky actions.

The findings on the effectiveness of the ecological displays support Borst et al.'s observation that operators sometimes perform riskier actions with ecological displays if the displays have not made the intentional constraints (e.g., laws and regulations) adequately visible to the operators (2015). The market-portfolio-execution visualization, the major part of the ecological displays used by the two experiments, was solely based on a physical constraint to maximizing the profitability of trading, as demonstrated by a Functional Purpose of the AH. Minimizing the risk has been identified as the other Functional Purpose as an intentional constraint but has not been implemented in the interface design. In a realistic trading environment, with more intentional constraints regarding risk control represented on the ecological displays (e.g., market fundamentals, laws and regulations), real traders facing realistic scenarios may be less prone to risk-seeking decisions.

The EID approach has originally been developed for life-critical domains such as aviation in which Borst et al.'s observation took place. It should be noted that these domains are in general professional and the ultimate goal of display research in these domains is to support expert operation, given that a minimum level of expertise is required for even the novice operators to work in these domains. Extending the EID approach to financial trading, a non-life critical domain but has a substantial social impact, requires more aspects of the effectiveness of ecological displays being examined, given that both expert and novice traders participate in the same market. According to the game theory, financial trading is a zero-sum game in which expert traders (or better automation) profit from competitors who are less skillful. The expert and the novice traders participate in the same market and, as shown in the qualitative results of risk preference, may have different types of risk tolerance. Future research should be conducted to evaluate how EID could influence the risk preference of traders who have different extents of professional knowledge.

5.2 Summary of Contributions

A summary of contributions to the fields of human factors and finance is provided as follows.

5.2.1 Contributions to Human Factors

First, this dissertation is the first to demonstrate CWA as a useful modeling tool for understanding the complexity in the automated trading domain. The DOA layering approach was developed by adopting Parasuraman et al.'s stages and levels of automation model (2000) to CWA.

Kaber's (2017) recent comments on the stages and level of automation model summarized many challenges in developing a human factors approach to studying automation. Stages and levels of automation should continue to be used as an important modeling tool but further development is needed. As Burns suggested (2017), with the emergence of intelligent automation, human-automation interactions would be more complex than "basic psychological performance" that has been well addressed in legacy models. New models are warrantied and can be established by studying new domains for specific design challenges. The DOA layering approach was demonstrated as a versatile approach for modeling systems that have a variable DOA. This dissertation described examples of a financial system, but arguably, the DOA layering approach should be applicable to other automated systems where automation is equally or more pervasive.

Secondly, the design implications of the DOA layering approach were not limited to display design, but also included made to automation design. With most CWA works only have implications for display design, this current work made a unique contribution. Examples of automation design based on this approach were provided in Experiment 1 and 2. The ecological display design addressed different task requirements that were captured in the CWA models.

Third, AUTRASS, an automated trading microworld was developed to encourage more human factors experimental studies in this domain. Finance is an important domain that requires more human factors research, but previously there was only limited tool available to researchers. Being able to simulate the dynamics of financial trading in a lab-controlled environment has been a considerably important achievement.

Lastly, the experiment results demonstrated the possible influence with automation and display design on task performance and risk preference in a variable DOA setting. Monitoring performance improvement has been achieved with automation and display appropriately designed, shedding some light on improving software design in the financial trading domain. The evaluation of risk preference introduced an important new line of research to human factors, but further work is needed.

5.2.2 Contributions to Finance

This dissertation has implications for improving the design of trading algorithms and the broader automation design. Links were created between the social and the technological aspects of financial trading. The introduction of the CWA approach to this domain was not intended to eliminate the complexity relevant to automation. Automated trading was evaluated through a human factors angle, with its complexity explicitly revealed in the interrelated processes of modeling, design, and evaluation.

Introducing ecological displays to the finance domain makes a unique contribution. Rather than replacing the traders with highly automated technologies, which is a common practice in the finance industry, this dissertation introduced a different view to improving the resilience of financial trading systems by exploring ways to utilize the flexibility of human beings to cope with unexpected situations. This dissertation explored ways in which technology can facilitate human adaptivity and flexibility to cope with unforeseen events (i.e., to enhance resilience in trading with more effectiveness automation and display design).

The influence of ecological displays on traders' risk preference makes an important contribution. With different risk preference patterns among novice and professionally-trained traders, it might be possible to develop new trading systems and strategies that have a better understanding of the different risk preferences of the market participants. Indeed, a subject-domain expert has started to implement a trading strategy inspired by this dissertation.

5.3 Limitations

The DOA layering approach is preliminary and has only been developed at the WDA and ConTA levels. This dissertation has discussed the future possibility of extending this approach to other

CWA levels, including StrA. StrA could make important contributions in guiding the design of adaptive automation, as well as display support that would be appropriate.

The two experiments demonstrated examples of automated systems in the financial trading domain. The distinct automation and display designs in a variety of scenario types showed influence on trader behaviour as well as risk preference. It is possible that the magnitudes of the relative differences are not fully generalizable – to different design concepts or to other domains. These concerns are valid, however, the facts that the consistent patterns in participants' responses have been found across the two experiments have demonstrated that reasons deeper than simple visual design difference could have contributed to the observed differences. Nevertheless, financial trading is an emerging domain that requires more human factors research, and more work is required.

More on the experimental study side, is the possibility that factors such as subject-domain knowledge and market dynamics could have mediated the participants' responses. This limitation, however, was subject to the high fidelity of AUTRASS and experimental design. Being able to simulate financial trading in a lab-controlled environment and performing one of the first human factors in this domain has been challenging. The design of the experimental study has been impacted by a number of constraints that made the trade-off between fidelity and scientific validity more complex. Arguably, in this initial attempt to understand a brave new world, it is common for researchers to take a bottom-up approach, trying to understand the fundamentals of the specific domain, then moving towards the next iteration of refined design.

5.4 Future Work

5.4.1 Model Development

The DOA layering approach has great potentials to be used as a modeling approach and a design approach, as demonstrated in this dissertation. As a modeling approach, the DOA layering approach can be extended to include more phases of the CWA. For example, adaptive automation has only been preliminary described in this dissertation, with strategies that dominate the DOA shifts not modeled. This might suggest opportunities to explore the StrA.

As a design approach, the DOA layering approach can be improved to better support the automation design as well as the display design. With the results of this dissertation showing that the ecological displays have only supported monitoring in a specific DOA, more research is warranted on this topic.

5.4.2 Improving Simulation Fidelity

This first-attempted exploration of automated trading used a practical simplification of real financial trading to conduct human factors research but may not represent the full work environment of the professional traders. The professional traders worked with the higher workload, monitoring multiple algorithms and financial products in a much longer timeframe which were not simulated in this experiment.

The simulation required adherence to strict experimental protocols and therefore limited the participants' authority in financial trading to some extent. For example, several restrictions were applied to the participants' task in moderate DOA scenarios to reduce individual differences in inputting trading parameters. In original high DOA, the fault detection task provided intrinsic motivation for monitoring the automation trading. Obviously, this task was hypothetical and probably not typical in real-world automated trading. Further, AUTRASS simulated trend following trading using historical data in a way similar to back-testing. The performance of a back-tested trading system may only be achieved in a certain market; therefore, back-testing is limited by this potential over-fitting problem (Hu & Watt, 2014). This experiment may not be able to provide immediate help in improving the profitability of automated trading due to the negative impact of the over-fitting problem. A higher fidelity, agent-based simulation of automated trading, using multiple computerized agents to simulate market participants was considered in the planning stage of this experiment in collaboration with the University of Western Ontario, London, Canada. This plan was unfortunately abandoned due to concerns over the difficulty in controlling for the potentially confounding influences of the participants' lack of trading experience.

Bibliography

Achonu, J., & Jamieson, G. A. (2003). Work domain analysis of a financial system: an abstraction hierarchy for portfolio management. In *Proceedings of the 22nd European Annual Conference on Human Decision Making and Control* (pp. 102–109). Linköping, Sweden: Citeseer. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.92.8202&rep=rep1&type=pdf

Ahn, H.-J., & Cheung, Y.-L. (1999). The intraday patterns of the spread and depth in a market without market makers: The Stock Exchange of Hong Kong. *Pacific-Basin Finance Journal*, *7*(5), 539–556.

Alaieri, A., & Vellino, A. (Eds.). (2016). Ethical Decision Making in Robots: Autonomy, Trust and Responsibility Autonomy Trust and Responsibility (pp. 159–168). Presented at the International Conference on Social Robotics 2016: Social Robotics, Cham: Springer International Publishing.

Arrabito, G. R., Ho, G., Li, Y., Giang, W., Burns, C. M., & Hou, M. (in review). Supervisory Control of an Unmanned Aircraft System: Evaluation of Sonification and Tactification for Encoding Urgency of System Health Status.

Ashoori, M., & Burns, C. (2010). Reinventing the Wheel: Control Task Analysis for Collaboration. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*(4), 274–278. https://doi.org/10.1177/154193121005400402

Ashoori, M., & Burns, C. M. (2013). Team Cognitive Work Analysis Structure and Control Tasks. *Journal of Cognitive Engineering and Decision Making*, *7*(2), 123–140.

Avellaneda, M. (2011). Algorithmic and High-frequency trading: an overview. *Retrieved May*, *11*, 2013.

Bailey, N. R., Scerbo, M. W., Freeman, F. G., Mikulka, P. J., & Scott, L. A. (2006). Comparison of a Brain-Based Adaptive System and a Manual Adaptable System for Invoking Automation. *Human Factors*, *48*(4), 693–709. https://doi.org/10.1518/001872006779166280

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775–779.

Bartlett, M. S. (1947). The use of transformations. *Biometrics*, *3*(1), 39–52.

Bisantz, A. M., & Burns, C. M. (Eds.). (2009). *Applications of cognitive work analysis*. Boca Raton: CRC Press.

Bloomberg. (2017, January 24). Aleynikov on the Hook Again for Taking HFT Code From Goldman. *Bloomberg.Com*. Retrieved from https://www.bloomberg.com/news/articles/2017-01-24/aleynikov-s-conviction-is-reinstated-by-state-appeals-court

Bollinger, J. (2001). *Bollinger on Bollinger bands*. New York, NY: McGraw-Hill.

Borst, C., Flach, J. M., & Ellerbroek, J. (2015). Beyond Ecological Interface Design: Lessons From Concerns and Misconceptions. *IEEE Transactions on Human-Machine Systems*, *45*(2), 164–175. https://doi.org/10.1109/THMS.2014.2364984

Burns, C. M. (2017). Automation and the Human Factors Race to Catch Up. *Journal of Cognitive Engineering and Decision Making*, 1555343417724975. https://doi.org/10.1177/1555343417724975

Burns, C. M., Bisantz, A. M., & Roth, E. M. (2004). Lessons From a Comparison of Work Domain Models: Representational Choices and Their Implications. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(4), 711–727. https://doi.org/10.1518/hfes.46.4.711.56810

Burns, C. M., Enomoto, Y., & Momtahan, K. (2009). A cognitive work analysis of cardiac care nurses performing teletriage. In A. M. Bisantz & C. M. Burns (Eds.), *Applications ofcognitive work analysis* (pp. 149–174). Boca Raton, FL: CRC Press.

Burns, C. M., & Hajdukiewicz, J. R. (2004). *Ecological interface design*. Boca Raton, FL: CRC Press.

Burns, C. M., Skraaning Jr., G., Jamieson, G. A., Lau, N., Kwok, J., Welch, R., & Andresen, G. (2008). Evaluation of ecological interface design for nuclear process control: situation awareness effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(4), 663–679.

Calhoun, G. L., Ward, V. B. R., & Ruff, H. A. (2011). Performance-based Adaptive Automation for Supervisory Control. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*(1), 2059–2063. https://doi.org/10.1177/1071181311551429

Canadian Securities Institute. (2004). *Canadian Securities Course Volume I*. CSI Global Education.

Cartea, Á., Jaimungal, S., & Penalva, J. (2015). *Algorithmic and high-frequency trading*. Cambridge University Press. Retrieved from https://books.google.ca/books?hl=en&lr=&id=5dMmCgAAQBAJ&oi=fnd&pg=PR13&dq=%5B14%5D%09%C3%81.+Cartea,+S.+Jaimungal,+and+J.+Penalva,+Algorithmic+and+High-Frequency+Trading,+Cambridge:+Cambridge+University+Press,+2015.&ots=49HjSQDGb_&sig=jyzttvGDUjRVWkYxEm3qjZ9OhS0

Chan, E. P. (2009). *Quantitative trading: how to build your own algorithmic trading business*. Hoboken, N.J: John Wiley & Sons.

Chen, Y., Elenee Argentinis, J., & Weber, G. (2016). IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. *Clinical Therapeutics*, *38*(4), 688–701. https://doi.org/10.1016/j.clinthera.2015.12.001

Clarke, R. G., Krase, S., & Statman, M. (1994). Tracking errors, regret, and tactical asset allocation. *The Journal of Portfolio Management*, *20*(3), 16–24.

CodePlex. (2011). Dynamic Data Display. Retrieved March 19, 2017, from http://dynamicdatadisplay.codeplex.com/wikipage?title=D3v1.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cornelissen, M., Salmon, P. M., McClure, R., & Stanton, N. A. (2012). What are they doing: testing a structured cognitive work analysis-based approach for identifying different road user strategies. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 363–367. https://doi.org/10.1177/1071181312561083

Cullen, R. H., Rogers, W. A., & Fisk, A. D. (2012). The Effects of Automation Reliability and Experience on Attention in a Computer Environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 511–515. https://doi.org/10.1177/1071181312561106

Dainoff, M. J., & Dainoff, C. A. (2003, May 1). Financial information display system.

Dainoff, M. J., Dainoff, C. A., & McFeeters, L. (2004). On the Application of Cognitive Work Analysis to the Development of a Commercial Investment Software Tool. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*(3), 595–599. https://doi.org/10.1177/154193120404800368

Davis, M., Kumiega, A., & Van Vliet, B. (2013). Ethics, Finance, and Automation: A Preliminary Survey of Problems in High Frequency Trading. *Science and Engineering Ethics*, *19*(3), 851–874. https://doi.org/10.1007/s11948-012-9412-5

Dikmen, M., & Burns, C. M. (2016). Autonomous Driving in the Real World: Experiences with Tesla Autopilot and Summon. In *Proceedings of the 8th International Conference on Automotive User*

Interfaces and Interactive Vehicular Applications (pp. 225–228). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=3005465

Dredge, S. (2015, March 18). Elon Musk: self-driving cars could lead to ban on human drivers. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2015/mar/18/elon-musk-self-driving-cars-ban-human-drivers

Easley, D., & O'hara, M. (2010). Microstructure and ambiguity. *The Journal of Finance*, *65*(5), 1817–1846.

Ellerbroek, J., Brantegem, K. C. R., van Paassen, M. M., de Gelder, N., & Mulder, M. (2013). Experimental Evaluation of a Coplanar Airborne Separation Display. *IEEE Transactions on Human-Machine Systems*, *43*(3), 290–301. https://doi.org/10.1109/TSMC.2013.2238925

Ellis, C. A., & Parbery, S. A. (2005). Is smarter better? A comparison of adaptive, and simple moving average trading strategies. *Research in International Business and Finance*, *19*(3), 399–411. https://doi.org/10.1016/j.ribaf.2004.12.009

Endsley, M. R. (1988). Situation awareness global assessment technique (SAGAT). In *Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National* (pp. 789–795). IEEE. Retrieved from http://ieeexplore.ieee.org/abstract/document/195097/

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 65–84.

Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 4–32.

Endsley, M. R. (2017). Autonomous Driving Systems: A Preliminary Naturalistic Study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making*, 1555343417695197.

Endsley, M. R., Hoffman, R., Kaber, D., & Roth, E. (2007). Cognitive engineering and decision making: An overview and future course. *Journal of Cognitive Engineering and Decision Making*, *1*(1), 1–21.

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462–492. https://doi.org/10.1080/001401399185595

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, *37*(2), 381–394.

Fabozzi, F. J., Focardi, S. M., & Jonas, C. (2011). High-frequency trading: methodologies and market impact. *Review of Futures Markets*, *9*, 7–38.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Field, A. (2005). *Discovering Statistics Using SPSS. Second Edition*. London: SAGE Publications.

Fortune. (2016, April 5). From Cancer to Consumer Tech: A Look Inside IBM's Watson Health Strategy. Retrieved April 26, 2017, from http://fortune.com/ibm-watson-health-business-strategy/

Furukawa, H., & Parasuraman, R. (2003). Supporting System-Centered View of Operators Through Ecological Interface Design: Two Experiments on Human-Centered Automation. In *Proceedings of the HumProceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (Vol. 47, pp. 567–571). Denver, CO: SAGE Publications. https://doi.org/10.1177/154193120304700366

Gary, A. K. I., Schluetter, J.-U., & Brumfield, H. (2004, July). Click based trading with intuitive grid display of market depth. Retrieved from http://www.google.com/patents/US6766304

Gugerty, L. (2011). Situation awareness in driving. In D. L. FIsher, M. Rizzo, J. K. Caird, & J. D. Lee, *Handbook for driving simulation in engineering, medicine and psychology* (Vol. 1, pp. 265–272). Boca Raton, FL: CRC Press.

Hajdukiewicz, J., & Burns, C. (2004). Strategies for bridging the gap between analysis and design for ecological interface design. In *Proceedings of the Human Factors and Ergonomics Society*

Annual Meeting (Vol. 48, pp. 479–483). SAGE Publications. Retrieved from http://pro.sagepub.com/content/48/3/479.short

Hajdukiewicz, J. R., Vicente, K. J., Doyle, D. J., Milgram, P., & Burns, C. M. (2001). Modeling a medical environment: an ontology for integrated medical informatics design. *International Journal of Medical Informatics*, *62*(1), 79–99.

Hassall, M. E., & Sanderson, P. M. (2014). A formative approach to the strategies analysis phase of cognitive work analysis. *Theoretical Issues in Ergonomics Science*, *15*(3), 215–261. https://doi.org/10.1080/1463922X.2012.725781

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523. https://doi.org/10.1016/j.tics.2009.09.004

Hey, J. D., & Morone, A. (2004). Do markets drive out lemmings—or vice versa? *Economica*, *71*(284), 637–659.

Hilliard, A., & Jamieson, G. A. (2014). A strategy-based ecological (?) display for time-series structural change diagnosis. In *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on* (pp. 347–352). IEEE.

Hilliard, A., & Jamieson, G. A. (2015). Representing energy efficiency diagnosis strategies in cognitive work analysis. *Applied Ergonomics*. https://doi.org/10.1016/j.apergo.2015.10.009

Hoffmann, A. O. I., Post, T., & Pennings, J. M. E. (2013). Individual investor perceptions and behavior during the financial crisis. *Journal of Banking & Finance*, *37*(1), 60–74. https://doi.org/10.1016/j.jbankfin.2012.08.007

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Horiguchi, Y., Burns, C. M., Nakanishi, H., & Sawaragi, T. (2013). Visualization of Control Structure in Human-Automation System Based on Cognitive Work Analysis. In *Analysis, Design, and Evaluation of Human-Machine Systems* (Vol. 12, pp. 423–430). Retrieved from http://www.researchgate.net/profile/Catherine_Burns3/publication/236890155_Visualization_of_control_structure_in_human-automation_system_based_on_Cognitive_Work_Analysis/links/0c960519f64441bc93000000.pdf

Hu, R., & Watt, S. M. (2014). An Agent-Based Financial Market Simulator for Evaluation of Algorithmic Trading Strategies. In *6th International Conference on Advances in System Simulation* (pp. 221–227). Nice, France: International Academy, Research, and Industry Association. Retrieved from http://www.csd.uwo.ca/~watt/pub/reprints/2014-simul-agent-based.pdf

Iati, R. (2009). The real story of trading software espionage. Retrieved from http://www.advancedtrading.com.

Interactive Brokers. (n.d.). Interactive Brokers. Retrieved May 5, 2017, from https://www.interactivebrokers.ca/en/home.php

Investopedia. (2003a, November 18). 52-Week High/Low. Retrieved June 7, 2017, from http://www.investopedia.com/terms/1/52weekhighlow.asp

Investopedia. (2003b, November 26). Short Selling. Retrieved May 15, 2017, from http://www.investopedia.com/terms/s/shortselling.asp

Investopedia. (n.d.-a). Slippage. Retrieved May 26, 2017, from http://www.investopedia.com/terms/s/slippage.asp

Investopedia. (n.d.-b). What is the difference between investing and trading? Retrieved May 14, 2017, from http://www.investopedia.com/ask/answers/12/difference-investing-trading.asp

Jenkins, D. P., Stanton, N. A., Salmon, P. M., Walker, G. H., & Young, M. S. (2008). Using cognitive work analysis to explore activity allocation within military domains. *Ergonomics*, *51*(6), 798–815.

Jenkins, D. P., Stanton, N. A., & Walker, G. H. (2009). *Cognitive work analysis: coping with complexity*. Ashgate Publishing, Ltd. Retrieved from https://books.google.ca/books?hl=en&lr=&id=KxJnw4RYiPMC&oi=fnd&pg=PR9&dq=Jenkins

,+D.P.,+Stanton,+N.+A.,+Salmon,+P.+M.,+%26+Walker,+G.+H.+(2009).+Cognitive+work+analysis:+Coping+with+complexity.+Aldershot:+Ashgate&ots=_w4LexWqfA&sig=QibcXwCPF7x3q7l4WsHhYzyMxeA

Jiancaro, T., Jamieson, G. A., & Mihailidis, A. (2014). Twenty years of cognitive work analysis in health care: a scoping review. *Journal of Cognitive Engineering and Decision Making*, *8*(1), 3–22.

Kaber, D. B. (2017). Issues in Human-Automation Interaction Modeling: Presumptive Aspects of Frameworks of Types and Levels of Automation. *Journal of Cognitive Engineering and Decision Making*, 1555343417737203. https://doi.org/10.1177/1555343417737203

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, *5*(2), 113–153. https://doi.org/10.1080/1463922021000054335

Kaber, D. B., Riley, J. M., Tan, K.-W., & Endsley, M. R. (2001). On the design of adaptive automation for complex systems. *International Journal of Cognitive Ergonomics*, *5*(1), 37–57.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Kumiega, A., & Van Vliet, B. E. (2012). Automated Finance: The Assumptions and Behavioral Aspects of Algorithmic Trading. *Journal of Behavioral Finance*, *13*(1), 51–55. https://doi.org/10.1080/15427560.2012.654924

Lau, N., Jamieson, G. A., Skraaning Jr., G., & Burns, C. M. (2008). Ecological Interface Design in the Nuclear Domain: An Empirical Evaluation of Ecological Displays for the Secondary Subsystems of a Boiling Water Reactor Plant Simulator. *IEEE Transactions on Nuclear Science*, *55*(6), 3597–3610. https://doi.org/10.1109/TNS.2008.2005725

Lau, N., Skraaning Jr, G., Jamieson, G. A., & Burns, C. M. (2008). Enhancing operator task performance during monitoring for unanticipated events through ecological interface design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, pp. 448–452). SAGE Publications. Retrieved from http://pro.sagepub.com/content/52/4/448.short

Lau, N., Skraaning Jr., G., Jamieson, G. A., & Burns, C. M. (2008). *The Impact of Ecological Displays on Operator Task Performance and Workload*. OECD Halden Reactor Project, Halden, Norway HWR-888. Retrieved from http://cel.mie.utoronto.ca/wp-content/uploads/CEL08-02.pdf

Leaprate. (2014, August 1). Exclusive: Quantica Trader prepares for launch. Retrieved May 8, 2017, from https://www.leaprate.com/news/exclusive-quantica-trader-prepares-for-launch/

Leaver, M., & Reader, T. W. (2015). Non-technical skills for managing risk and performance in financial trading. *Journal of Risk Research*, 1–35. https://doi.org/10.1080/13669877.2014.1003319

Leaver, M., & Reader, T. W. (2016). Human Factors in Financial Trading: An Analysis of Trading Incidents. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *58*(6), 814–832.

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lees, M. N., & Lee, J. D. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems. *Ergonomics*, *50*(8), 1264–1286.

Li, H. (2013). *Supporting Human - Automation Collaboration through Dynamic Function Allocation The Case of Space Teleoperation*.

Li, H., Sarter, N., Wickens, C., & Sebok, A. (2013). Supporting Human-Automation Collaboration Through Dynamic Function Allocation: The Case of Space Teleoperation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 359–363. https://doi.org/10.1177/1541931213571078

Li, Y., Burns, C., & Hu, R. (2016). Representing Stages and Levels of Automation on a Decision Ladder: The Case of Automated Financial Trading. In *Proceedings of the Human Factors and Ergonomics Society 60th Annual Meeting* (Vol. 60, pp. 328–332). Washington, D. C.: SAGE Publications.

Li, Y., & Burns, C. M. (2017). Modeling Automation with Cognitive Work Analysis to Support Human-Automation Coordination. *Journal of Cognitive Engineering and Decision Making*.

Li, Y., Burns, C. M., & Hu, R. (2015). Understanding automated financial trading using work domain analysis (Vol. 59, pp. 165–169). Presented at the Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting, Los Angeles, US: SAGE Publications.

Luger, E., & Sellen, A. (2016). "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents (pp. 5286–5297). ACM Press. https://doi.org/10.1145/2858036.2858288

Manzey, D., Reichenbach, J., & Onnasch, L. (2008). Performance consequences of automated aids in supervisory control: The impact of function allocation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, pp. 297–301). SAGE Publications. Retrieved from http://pro.sagepub.com/content/52/4/297.short

Mazaeva, N., & Bisantz, A. M. (2003). Modeling automation within an abstraction hierarchy. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 47, pp. 649–652). SAGE Publications. Retrieved from http://pro.sagepub.com/content/47/3/649.short

Mazaeva, N., & Bisantz, A. M. (2007). On the representation of automation using a work domain analysis. *Theoretical Issues in Ergonomics Science*, *8*(6), 509–530. https://doi.org/10.1080/14639220600647816

McAndrew, C., & Gore, J. (2013). Understanding preferences in experience-based choice: A study of cognition in the "wild." *Journal of Cognitive Engineering and Decision Making*, *7*(2), 179–197. https://doi.org/10.1177/1555343412463922

McIlroy, R. C., & Stanton, N. A. (2015). Ecological interface design two decades on: Whatever happened to the SRK taxonomy. *IEEE Transactions on Human-Machine Systems*, *45*(2), 145–163. https://doi.org/10.1109/THMS.2014.2369372

Mendelson, M. (1972). *From automated quotes to automated trading: restructuring the stock market in the US*. New York University, Graduate School of Business Administrtion, Institute of Finance.

Metford, P. (2010). Securities trading method.

Metz, C. (n.d.). The Rise of the Artificially Intelligent Hedge Fund. Retrieved April 24, 2017, from https://www.wired.com/2016/01/the-rise-of-the-artificially-intelligent-hedge-fund/

Militello, L. G., & Hutton, R. J. (1998). Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, *41*(11), 1618–1641.

Miller, C. A., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*(1), 57–75.

Minotra, D., & Burns, C. M. (2016). Understanding safe performance in rapidly evolving systems: A risk management analysis of the 2010 US financial market Flash Crash with Rasmussen's risk management framework. *Theoretical Issues in Ergonomics Science. Advance Online Publication. DOI: 10.1080/1463922X.2016.1254837*.

Modha, D. S., Ananthanarayanan, R., Esser, S. K., Ndirango, A., Sherbondy, A. J., & Singh, R. (2011). Cognitive computing. *Communications of the ACM*, *54*(8), 62. https://doi.org/10.1145/1978542.1978559

Modulus. (n.d.). TradeScript - The Easiest Trading Strategy Programming Language. Retrieved April 24, 2017, from http://www.modulusfe.com/products/trading-system-developer-components/tradescript/

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology: Applied*, *6*(1), 44.

Mosier, K., & Fischer, U. (2012). Impact of Automation, Task and Context Features on Pilots' Perception of Human-Automation Interaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, pp. 70–74). Sage Publications Sage CA: Los Angeles, CA. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1071181312561035

Mouloua, M., & Parasuraman, R. (1994). *Human performance in automated systems: Current research and trends*. Erlbaum Hillsdale, NJ. Retrieved from http://userpages.umbc.edu/~norcio/papers/1994/Chen-RecUBAM-HPAS.pdf

Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin.

Naikar, N., Pearce, B., Drumm, D., & Sanderson, P. M. (2003). Designing teams for first-of-a-kind, complex systems using the initial phases of cognitive work analysis: Case study. *Human Factors*, *45*(2), 202–217.

National Highway Traffic Safety Administration. (2013). Preliminary statement of policy concerning automated vehicles. *Washington, DC*, 1–14.

N.D. Ill. v. Sarao, United States District Court Northern District of Illinois Eastern Division (2015). Retrieved from http://www.justice.gov/sites/default/files/opa/press-releases/attachments/2015/04/21/sarao_criminal_complaint.pdf.

Ninja Trader. (n.d.). Trading Software & Brokerage. Retrieved May 5, 2017, from http://ninjatrader.com/

Nygren, T. E. (1991). Psychometric Properties of Subjective Workload Measurement Techniques: Implications for Their Use in the Assessment of Perceived Mental Workload. *Human Factors*, *33*(1), 17–33. https://doi.org/10.1177/001872089103300102

Odean, T. (1998). Do investors trade too much? Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=94143

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(3), 476–488. https://doi.org/10.1177/0018720813501549

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors*, *29*(6), 695–706.

Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G., & Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems*. CATHOLIC UNIV OF AMERICA WASHINGTON DC COGNITIVE SCIENCE LAB. Retrieved from http://www.dtic.mil/docs/citations/ADA254595

Parasuraman, R., Cosenzo, K. A., & De Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology*, *21*(2), 270.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced'complacency'. *The International Journal of Aviation Psychology*, *3*(1), 1–23.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions On*, *30*(3), 286–297.

Parasuraman, R., & Wickens, C. D. (2008). Humans: Still Vital After All These Years of Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 511–520. https://doi.org/10.1518/001872008X312198

Pogue, D. (2012). Siri, why aren't you smarter? *Scientific American*, *307*(2), 33–33.

Pritchett, A. R., Kim, S. Y., & Feigh, K. M. (2014). Modeling human–automation function allocation. *Journal of Cognitive Engineering and Decision Making*, *8*(1), 33–51.

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*(1), 1–14.

Rasmussen, J. (1974). *The Human Data Processor as a System Component: Bits and Pieces of a Model (Report No. Riso-M1722)*. Retrieved from http://orbit.dtu.dk/fedora/objects/orbit:89096/datastreams/file_e2e379ec-324b-4497-b8eb-e571d574ddf9/content

Rasmussen, J. (1976). Outlines of a hybrid model of the process plant operator. In *Monitoring behavior and supervisory control* (pp. 371–383). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4684-2523-9_31

Rasmussen, J. (1979). *On the structure of knowledge-a morphology of metal models in a man-machine system context*. DTIC Document. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA145684

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. New York: North-Holland.

Rasmussen, J., & Goodstein, L. P. (1987). Decision support in supervisory control of high-risk industrial systems. *Automatica*, *23*(5), 663–671.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). Cognitive systems engineering. Retrieved from http://www.forskningsdatabasen.dk/en/catalog/2185760125

Reising, D. V. C., & Sanderson, P. M. (2000a). Testing the impact of instrumentation location and reliability on ecological interface design: Control performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, pp. 124–127). SAGE Publications Sage CA: Los Angeles, CA. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/154193120004400133

Reising, D. V. C., & Sanderson, P. M. (2000b). Testing the impact of instrumentation location and reliability on ecological interface design: fault diagnosis performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, pp. 591–594). SAGE Publications Sage CA: Los Angeles, CA. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/154193120004402225

Reuters. (2017, April 25). ABB, IBM team up on industrial artificial intelligence. Retrieved April 26, 2017, from http://www.reuters.com/article/us-abb-ibm-idUSKBN17R0DE

Riley, V. (1989). A General Model of Mixed-Initiative Human-Machine Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *33*(2), 124–128. https://doi.org/10.1177/154193128903300227

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 5–19.

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. *Handbook of Human Factors and Ergonomics*, *2*, 1926–1943.

Schlossberg, B. (2005, February 17). Using Bollinger Band® "Bands" to Gauge Trends. Retrieved June 12, 2017, from http://www.investopedia.com/articles/trading/05/022205.asp

Sharpe, W. F. (1994). The sharpe ratio. *The Journal of Portfolio Management*, *21*(1), 49–58.

Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. MIT press. Retrieved from https://books.google.ca/books?hl=en&lr=&id=eu41_M2Do9oC&oi=fnd&pg=PA1&dq=Sheridan,+1992&ots=Z6u07VtF9d&sig=HlbnDY3I2KImmZQ4BsYkF9sjRf4

Sheridan, T. B. (2011). Adaptive automation, level of automation, allocation authority, supervisory control, and adaptive control: Distinctions and modes of adaptation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *41*(4), 662–667.

Sheridan, T. B. (2017). Musings on Models and the Genius of Jens Rasmussen. *Applied Ergonomics*, *59*, 598–601.

Sheridan, T. B., & Parasuraman, R. (2006). Human-Automation Interaction.

Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. DTIC Document. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA057655

Shiller, R. J. (2015). *Irrational exuberance, 3rd edition*. Princeton, NJ: Princeton University Press.

Smith, A. G., & Jamieson, G. A. (2012). Level of Automation Effects on Situation Awareness and Functional Specificity in Automation Reliance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 2113–2117.

Solaimani, S., Keijzer-Broers, W., & Bouwman, H. (2015). What we do – and don't – know about the Smart Home: An analysis of the Smart Home literature. *Indoor and Built Environment*, *24*(3), 370–383. https://doi.org/10.1177/1420326X13516350

Sundström, G., & Hollnagel, E. (2011). *Governance and Control of Financial Systems: A Resilience Engineering Perspective*. Farnham, UK: Ashgate.

TD thinkorswim. (n.d.). Online Trading Platform. Retrieved May 5, 2017, from https://www.thinkorswim.com/t/innovation.html

Terrell, E. (2006). History of the NASDAQ and American Stock Exchanges (Business Reference Services, Library of Congress). Retrieved May 14, 2017, from http://www.loc.gov/rr/business/amex/amex.html

Trading Physics. (n.d.). Download. Retrieved from http://www.tradingphysics.org/Feeds/DownloadHistoricalOHLCV.aspx.

TradingSim. (n.d.). 6 Reasons Not to Trade During the First 30 Minutes. Retrieved from https://tradingsim.com/blog/6-reasons-not-to-trade-during-the-first-30-minutes.

Treleaven, P., Galas, M., & Lalchand, V. (2013). Algorithmic trading review. *Communications of the ACM*, *56*(11), 76–85.

Treynor, J. L. (1981). What does it take to win the trading game? *Financial Analysts Journal*, *37*(1), 55–60.

U.S. Commodity Futures Trading Commission, & U. S. Securities & Exchange Commission. (2010). *Findings Regarding the Market Events of May 6, 2010*. Retrieved from https://www.sec.gov/news/studies/2010/marketevents-report.pdf.

van de Merwe, K., van Dijk, H., & Zon, R. (2012). Eye movements as an indicator of situation awareness in a flight simulator experiment. *The International Journal of Aviation Psychology*, *22*(1), 78–95.

Vicente, K. J. (1991). *Supporting knowledge-based behavior through ecological interface design*.

Vicente, K. J. (1999). *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. Boca Raton, FL: CRC Press.

Vicente, K. J. (2002). Ecological interface design: Progress and challenges. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *44*(1), 62–78.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *Systems, Man and Cybernetics, IEEE Transactions On*, *22*(4), 589–606.

Wickens, C. D., Li, H., Santamaria, A., Sebok, A., & Sarter, N. B. (2010). Stages and Levels of Automation: An Integrated Meta-analysis. In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (Vol. 54, pp. 389–393). Los Angeles, CA: SAGE Publications.

Wiener, E. L., & Curry, R. E. (1980). Flight-deck automation: promises and problems. *Ergonomics*, *23*(10), 995–1011. https://doi.org/10.1080/00140138008924809

Woods, D. D., & Sarter, N. B. (2000). Learning from automation surprises and “going sour” accidents. *Cognitive Engineering in the Aviation Domain*, 327–353.

Appendix A

Literature Review

The following literature review focuses on three topics: automated trading, human-automation interaction, and CWA. After each topic is reviewed, the author provides a summary of the review and discuss the connections of the review to the three research questions that were proposed in Part A of this dissertation.

Automated Trading

This dissertation assumes the readers have no prior knowledge of finance; therefore, this literature review serves as a knowledge base of concepts, methods, and propositions of automated trading. On this topic, the author reviews the basic concepts of finance, automated trading, and behavioural finance. These three aspects have been briefly covered in Part A of this dissertation, and here the author reviews their origins and summarize key concepts. The review of automated trading and behavioural finance later in this section is not intended to be comprehensive; the coverage is limited to providing some references for developing a modeling approach and measures for design evaluation in this domain.

Basic Concepts of Finance

According to Canadian Securities Institute's definition (2004), *financial trading* is the buying or selling activity completed in and between three elements that formalize the flow of capital (e.g., money) and the flow of information (e.g., quotation): financial products, financial markets, and market intermediaries.

Financial products, also known as financial assets, or securities, are instruments with economic value that is owned by an individual or company (e.g., stocks, debts, and derivatives).

Financial markets are trustable vehicles that allow transactions of financial products between buyers and sellers. Transactions are achieved by auctions for a specific financial product using a shared centralized order book. The order book lists all buy and sell orders ranked by price and order arrival time, generally following a first-in-first-out rule. This auction process has been described many times in the literature (e.g., Hu & Watt, 2014; Treleaven et al., 2013). Exchanges, such as NYSE and NASDAQ, are institutes that maintain financial markets. In most financial markets, financial products are traded in

---

[1] Although this dissertation takes a systematic view on automated trading, this view is centralized on individual traders in automated trading and does not focus on market intermediaries (e.g., broker).

a secure and publicly transparent way. There are also exceptions; for example, in over-the-counter markets buyers and sellers traded directly with each other in a private manner.

Market intermediaries are individuals or institutions that provide services to facilitate the transactions (e.g., banks, and credit unions).

Traders maintain a *portfolio* of diversified financial products and try to make profits. In a bull (winning) market, profits can be made through buying at a lower price and selling at a higher price. For examples, a trader may buy certain *shares* of a stock when the *market price* is relatively low. When the market price jumps, the trader's portfolio is in a profitable *position* that has yet to be sold for cash. In this case, there is an *unrealized profit* in the portfolio. If the traders decide to sell these shares, the unrealized profit will become *a realized profit*. Profiting in a bear (falling) market is also possible, where profits can be generated by short selling at a higher price and buying to cover at a lower price (Investopedia, 2003b).

Automated Trading

This dissertation uses a historical and broad-sense definition of automated trading, in which case, "automated trading" does not equal to "algorithmic trading". This section first reviews the history of automated trading, then provides this definition and other details of automated trading that have not been covered in Part A.

History and Definition

At an earlier time, the concept of automated trading only included electronic quoting. NASDAQ, founded in 1971 (for a review, see Terrell, 2006), is the acronym of National Association of Securities Dealers Automated Quotations, and it is the first electronic quotation system. Later, NASDAQ became a stock market exchange with the electronic trading feature amended. Automated trading, in its broad sense, is an assembly of software and hardware platforms that utilizes high performance electronic processing units and telecommunication technologies for buyers and sellers to submit trading quotes to market exchanges (Mendelson, 1972). In a narrow and more common sense, automated trading is interchangeable with *algorithmic trading* which is “any form of trading using sophisticated algorithms (programmed systems) to automate all or some part of the trade cycle” (Treleaven et al., 2013).

While this dissertation certainly focuses on modern financial trading, automated trading is broadly defined to include algorithmic trading as well as powerful computers and communications infrastructure. This broad definition of automated trading aligns with how automation is defined in other domains.

Market Analysis

Analyzing which financial product to trade, when, and in which financial market to trade are instrumental to trading. There are two distinct market analyses in financial trading that differ in their beliefs about the mechanism of financial markets and, typically, lead to different styles in trading (Investopedia, n.d.-b).

Fundamental analysis looks to understand the value of financial products by considering a variety of political (e.g., government policies, laws, and regulations), economic (e.g., labor costs, taxes, and interest rates), and financial factors (e.g., company financial statements). Fundamental analysis, in theory, can be used to identify undervalued and overvalued financial products to advocate investment opportunities. Market participants who are more concerned with the fundamentals are investors and typically, investments are held for a long period.

Technical analysis, on the other hand, is the “study of market action, primarily through the use of charts, for the purpose of forecasting future price trends” (Murphy, 1999, p. 1). Charts are graphical presentations of data (e.g., price data plotted on a bar chart or a candlestick chart). Technical analysis studies the directions and magnitudes of financial markets and identifies reversal and continuation patterns. Technical analysis is normally employed in more frequent trading activities in comparison to fundamental analysis, therefore, it is the more commonly used market analysis in automated trading.

Trading Algorithm

Trading algorithms contain signals (when to do) and market executions (what to do) to complete transactions. A signal is an abstraction of a certain market condition decided by changes in values and charts. A pair of entry and exit signals can decide a trading window in which the trading automation can perform market executions (e.g. buying or selling). Trading algorithm designers have much flexibility to implement almost any trading strategy; however, expertise in specific programming languages is required to implement the desired trading strategy on top of a proprietary Application Program Interface (API). Some trading software, such as TradeScript (Modulus, n.d.), allow users to develop algorithms using scripting languages. However, these scripting languages are also proprietary and thus require a large amount of knowledge of the specific trading software.

Many trading algorithms utilize indicators of technical analysis to capture entry and exit signals in a dynamic financial market, and automate market executions based on triggering events defined by trading algorithm designers. Here the author briefly introduces *moving average*, the most commonly used technical indicator for *trend following trading* that will be revisited in the case study of this dissertation (Ellis & Parbery, 2005; Murphy, 1999, p. 195). The purpose of moving average is to identify when an old trend has reversed and a new trend has begun. A moving average averages a certain period of price data. As an hypothetical example, a 50-day moving average is calculated by summing up prices for the last 50 days then dividing the total by 50. According to Murphy, moving averages can be either unweighted or weighted. For examples, a simple moving average (SMA) is an unweighted moving average, which is calculated by assigning the equal weight to each day’s price. An exponentially smoothed moving average assigned greater weight to prices from more recent days, and therefore, it is an weighted moving average.

Human Factors in Automated Trading

Most efforts in the finance industry have focused on improving the profitability of trading algorithms, improving their ability of identifying patterns in the market data, limiting the amount and type of risk, and minimising transaction costs (Treleaven et al., 2013). In the finance literature, academic research has just begun to address the topic of automated trading and generally stays at a broad level (Avellaneda, 2011; Davis et al., 2013; Kumiega & Van Vliet, 2012; Treleaven et al., 2013). If, as Kumiega and Van Vliet stated, automated trading is a marriage between two standards - the first one being a relatively loose standard of subjective judgements of empirical financial data by the trader, and the second one being a more rigorous engineering standard of justification by automation - there is an new area for studying this relationship from an engineering point of view.

The reality is traders and automation usually buy and sell financial products electronically through a front-end system (e.g., Gary, Schluetter, & Brumfield, 2004; Interactive Brokers, n.d.; Ninja Trader, n.d.; TD thinkorswim, n.d.), and therefore, any research into this domain from an engineering point of view must not ignore the graphical displays as part of the front-end system. Graphical displays are ubiquitously used in financial trading for displaying information of financial markets (e.g., price and volume charts) and trading algorithms, and they are important in facilitating human and automation coordination in this domain. There should be a new line of research on improving the experience of using these graphical displays. Most trading front-end systems provide scripting languages that allow algorithm designers to program trading algorithms, and novel ways of improving the programming experience are being explored in the industry. For example, Metford (2010), on behalf of Quantica Trading, a Canadian-based automated trading software company, patented a method to allow for designing sophisticated strategy templates through an intuitive user interface. The templates determine the logic for the trading algorithms to enter or exit a financial market, and are independent of parameters such as financial products, prices, and quantities. Algorithm designers may program popular trading strategies into templates and reuse these templates in developing future trading algorithms. Relating to this topic, the author was embedded at Quantica Trading as a user experience designer in 2014 to design a drag and drop design tool that allows algorithm designers to develop and reuse trading strategy templates as described by Metford (Leaprate, 2014).

While intuitive user experience design could improve the experience of formalizing trading algorithms, there is another facet of automated trading that involves monitoring the trading system, including the trading algorithms, through a graphical display. Trading algorithms must be "back-tested" with historical market data on the front-end system to evaluate their profitability and stability, before being sent to live-trading (Treleaven et al., 2013). Back-testing and live-trading are similar processes in terms of the graphical displays being used and, of course, live-trading is certainly associated with real financial risks. In both processes, and particularly in live-trading, traders must effectively monitor the trading algorithms and intervene if necessary. So far, to the author's knowledge, there is no attention in the industry or academic research on using improved displays to support monitoring performance, after automation is developed.

Treleaven and colleagues pointed out that with increasing DOA in financial trading, there is a practical need for human behaviour to change from the old observe-and-execute model to a three-phase sequence (2013). This three-phase sequence, in line with the author's observations on the design of automated trading software at Quantica Trading, includes formalizing trading algorithms with estimating trends in the financial market, back-testing trading algorithms with historical data, and high performance live-trading to utilize the strong processing power. In practice, this three-phase sequence is an iterative process; however, humans may only improve the trading algorithms after each unsuccessful back test before sending the algorithms in for live-trading. The author argues though, a gulf lies between these three phases, as a rule of thumb is the financial market may deviate from human estimation at any time, and humans have limited abilities to intervene after the algorithm design is completed. While most efforts in the industry are to improve the experience of algorithm design, as the author stated in Part A of this dissertation, no research has been done to understand how humans interact with automation with which this gulf can be narrowed.

This argument is also supported by Kumiega and Van Vliet's statement (2012) on the relationship between humans and automation in automated trading. Humans, typically trading algorithm designers, set up estimations of the probability always with biases, while automation provided purely objective estimations. Automation has a stronger ability to perform calculations so that computers can obtain information that cannot be captured by humans. Kumiega and Van Vliet proposed three assumptions to take a systematic research approach on automated trading. First, the inputs into automated trading systems are driven by human behaviours (e.g., market predictions, programmed into the control law of the automation) and therefore they are not stable. Second, trading windows are short but during these windows the inputs are regulated by automation and considered as steady and exploitable trading opportunities; whereas, it is difficult for humans to identify these trading windows in automated trading, without adequate information support. Lastly, automation should generate stable outputs with an engineering repeatability. In this regard, automated trading is a complex information system with human-automation interactions heavily involved in its monitoring process, thereby more research is required in this area.

Behavioural Finance

The efficient market hypothesis (EMH, Fama, 1970) is the cornerstone of the neoclassical economic research. EMH describes an efficient market of rational behaviour. EMH assumes market participants are in general rational, thereby they value each financial product rationally for its fundamental value. In the weak assumptions of EMH, also known as the random walk hypothesis (Malkiel, 1973), even a relatively large group of market participants become irrational, the efficient market contains a self-correcting mechanism to defend market rationality. EMH suggests that prices of financial products are completely random, as the prices contain all available information and therefore, cannot be used to predict the future of the prices.

On the other hand, behavioural finance responds to the difficulties faced by EMH in explaining irrational phenomena (e.g., the U. S. stock market boom started in 1982 and the world financial crisis in 2007, described in Shiller, 2015). Behavioural finance argues that humans do not always make decisions rationally as assumed. Social psychological theories were introduced to explain biases in behaviors. For example, prospect theory (Kahneman & Tversky, 1979) described a loss aversion bias whereby people feel more pains at the prospect losses than an equal amount of gains. After that, there have been many studies reporting decision biases leading to non-rational market decisions, most of which had psychological evidence. There are many other examples, such as overconfidence, (e.g., Odean, 1998), herding behavior (e.g., Hey & Morone, 2004), ambiguity aversion (e.g., Easley & O'hara, 2010) and regret effect (e.g., Clarke, Krase, & Statman, 1994). There are critiques on the behavioural finance theory, as it only provides descriptive, but no formative explanations of these imperfections in decision-making and how to overcome these biases.

Summary and Connections to Research Questions

The author has three research questions:

Research question 1: How can we model automated trading systems with a variable DOA using CWA?

Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

This first literature review on automated trading provided a detailed view of the domain that has been explored with all three research questions. The basic concepts of finance described at the beginning of this literature review served as a glossary of terms that was revisited several times in this dissertation. This dissertation took a broad view of automated trading, suggesting that automated trading should be studied as a complex system, using a systematic approach. Later in this literature review, the author briefly introduced the history and the current state of behavioural finance. This psychological facet of automated trading was evaluated in dissertation, using a simulation approach with human participants.

Human-Automation Interaction

As the author introduced in Part A of this dissertation, the modeling section of this dissertation (Part B) aims to model a complex system with varying DOAs. DOA models, as defined in the human factors literature, identify the extent of functional distribution of automation in a specific work environment. This literature review presents an overview of human-automation interaction. The author review several popular topics in this area, including the existing DOA models as bases for developing new modeling approaches, performance consequences of using automation, the trade-off between the benefits and the risks of using automation, and adaptive automation.

Definitions of Automation

Parasuraman and Riley (1997) defined automation as the re-allocation of functions from humans to machine while these functions were previously assigned to humans. Sheridan and Parasuraman (2006), through a different angle, defined automation as the use of automatic control to replace human labor (both physical and mental labor) in any industry or science. Both definitions correspond to the spread of automation from the manufacturing industry to other domains, such as aviation (e.g., Wiener & Curry, 1980; Woods & Sarter, 2000), road transportation (e.g., Lees & Lee, 2007), teleoperation and robots (Sheridan & Verplank, 1978), and more recently, as the author mentioned in Part A of this dissertation, home and work automation domains (e.g., conversational agents and IBM Watson computers) that have not been well explored in the literature.

More importantly, the definitions of automation suggest what is more important is the interaction between humans and automation who take different roles in an automated system. Humans and automation must work well in a coordinated way. It is unlikely for humans to be fully passive users who only receive benefits of the powerful processing capabilities provided by the automation, in which case, the automation completely displaces human operators in the execution of all system functions

(Parasuraman & Riley, 1997); whereas, in most cases, humans still need to take actions on tasks that have not yet been allocated to the automation or cannot be performed by the automation. Nevertheless, humans are at least responsible for specifying the control laws for the automation, commanding the automation to start or stop a specific task, or stopping the automation entirely.

Degree of Automation Models

DOA addresses the extent of functional distribution of automation, which can be used to guide the representational design of the automation (i.e., designing a display that can support the use of automation) and predict the automation impacts on human and system performance (Wickens et al., 2010).

Several models have been developed to provide guidance to understand the partnership between humans and automation. Early works in the field identified automation as not a yes or no concept, and characterized automation as varying degrees of support in sensing information and taking control actions. Sheridan and Verplank proposed the taxonomy of levels of automation (1978) to represent the automatic control of remote surveillance devices. Their taxonomy was presented in a 10-point scale, from the lowest level of automation, with “no assistance provided by computer” to the highest level, with “full automation, completely ignoring human intervention”.

A well-known extension to the Sheridan and Verplank taxonomy is the Parasuraman, Sheridan and Wickens stages and levels of automation model (2000). In the development of this model, Parasuraman et al. incorporated four consecutive information-processing stages (i.e., four stages of automation) to Sheridan and Verplank’s taxonomy. These four stages are: 1) information acquisition (looking at sensory processing), 2) information analysis (including perception and working memory), 3) decision selection (effective decision-making), and 4) action implementation (including responses and control actions). The four stages of automation are orthogonal to the Sheridan and Verplank taxonomy, and together, the stage and levels of automation generalize a variety of components in a linear information-processing sequence. This two-dimension paradigm contains two continuums of degrees that may be involved in any automated system with automation levels across information-processing stages.

An alternative extension to the Sheridan and Verplank taxonomy was developed by Endsley and Kaber (1999). In the Endsley and Kaber model, automation was categorized as the amount of human mental or physical activities. This model extended the use of the original 10-point scale in the Sheridan and Verplank taxonomy - specifically made for describing teleoperation - to a broader range of domains such as aviation. Endsley and Kaber described that an agent, either human or computer, may carry out four functions that are generalized by stages in human information-processing. The four functions are: 1) monitoring (perceiving system status through an information display), 2) generating (creating strategies to achieve certain goals), 3) selecting (performing decision-making to choose a desired strategy), and 4) implementing (taking control actions). Endsley and Kaber also identified different roles required for taking different operations, and these roles can be distributed or shared between humans and computer (the term agent was used in Endsley and Kaber's work to include both human and automation as system operators).

The Parasuraman et al. model and the Endsley and Kaber model have the same theoretical common ground: automation has many facets and, since automation could span all information-processing stages, humans and automation must work together effectively in each stage. The DOA models, in general, have served as the foundation of the human-automation interaction research, for their ability to guide the selection of appropriate DOAs for human use, and to inform design concepts that focus on facilitating human-automation coordination. However, the DOA models have been criticized as "broad-brush descriptions of function allocation" (Pritchett et al., 2014), because they are not fully capable to describe the discrepancy between the significant authority gained by the automation (e.g., autopilot in Pritchett et al.'s example) and the responsibility ultimately kept by the human operators (e.g., safety). As shown in the case study provided by Pritchett et al., although functions related to safety have been allocated to the automation (as shown in the DOA models), human operators must still perform a number of critical check-ups to ensure the automation behaves as expected – a function that cannot be represented with the DOA models. Therefore, this discrepancy indicates a complex situation where there can be different human-automation relationships at the same DOA. Pritchett el al.'s criticism suggests that the DOA models are far from mutual to describe complex automated systems, and the author would argue, an opportunity lies in improving the DOA models with other modeling tools to cope with more complex scenarios of function allocation. Although the DOA models are not ideal, these models retain a richness of knowledge that can imply for later design and evaluation works in this area. Nevertheless, it is important for any new modeling works in this area to fully utilize the knowledge adopted from the DOA models.

Performance Consequences and Automation Trade-Off

In the early literature, incident and accident reports of the misuse of automation have been empirically studied (Parasuraman & Riley, 1997). Automation has been found to provide both benefits and risks to performance (e.g., Bainbridge, 1983; Parasuraman, 1987; Parasuraman & Riley, 1997; Sarter, Woods, & Billings, 1997; Wiener & Curry, 1980).

The performance consequences of using automation have been studied with different metrics, using both modeling and experimental approaches. For example, workload is one of the first studied metrics in this field. Riley (1989) proposed a mixed-initiative model to describe the structure of automated systems with automation on one side, the human operators on the other side and the environment in the middle. The Riley model contained workload as an important parameter of the world along with perceived workload as a parameter of human operations. The model suggested that the effects of automation to workload may be complicated, as many factors such as system reliability, task complexity, and time constraint could influence the use of automation. On the experimental side, Wiener and Curry conducted a series of studies with human operators who worked with automation in the aviation domain and identified automation impacts on human operators' workload (Wiener & Curry, 1980). Results of the Wiener and Curry studies show that a positive feature of automation is to reduce physical workload because massive control actions have been allocated to the automation. However, mental workload may increase with the use of automation, as human operators must remain monitoring for status of the system and the automation. It is therefore difficult to conclude how automation would affect human operators' workload, and a dilemma clearly exists in choosing the right automation to use.

Similar dilemmas have been found with studying other performance metrics, including loss of awareness (Endsley & Kiris, 1995) and complacency (Hoffmann, Post, & Pennings, 2013; Lee & See, 2004; Parasuraman, Molloy, & Singh, 1993), multi-task attention allocation (Cullen, Rogers, & Fisk, 2012) and perception change (Mosier & Fischer, 2012). This first attempt to systematically address the trade-off of using automation was Parasuraman and Riley's paper (1997). In this paper, Parasuraman and Riley stated that many performance issues come from humans misusing the technologies, and the underlying reason may be the misunderstanding of human roles in an automated system. Humans are unlikely to be removed from automated systems because they are more flexible and adaptable than automation in dynamic conditions. On the contrary, there are also risks because humans react to changing situations in various ways, especially when automation is no longer reliable as it was considered and therefore, humans need to retake the control independently.

Among all attempts to reason the causes of degraded human performance in automated systems, an important *automation trade-off* between benefits and costs has been concluded. In the literature, this automation trade-off has been described in different ways, for one example, as “the more support the automation provides, the higher the risk of degraded human performance” (Onnasch et al., 2014), and for another, as “the more reliable the automation in non-failure conditions is, the greater the level of complacency when automation fails” (Bainbridge, 1983).

It has been found that the trade-off becomes more sophisticated once the DOA continuum and more performance metrics are taken into account. The most recently work in this field is to extend the automation trade-off to include four variables: routine performance, failure performance, situation awareness (SA), and workload (Wickens et al., 2010). SA is characterized as a requirement for humans to maintain an up-to-date assessment of changing contexts in the forms of events, information and incidents. A high DOA reduces cognitive workload and improves task performance, but degrades situation awareness (SA) of system status and behaviour (Endsley & Kiris, 1995). Endsley and Kiris explained that when automation transfers the human workload of system executions to machines, it also increases the human workload of system monitoring. Humans now need to spend more time monitoring the running of automation. Therefore, as Endsley and Kiris suggested, to keep a proper level of SA, the DOA should not pass an optimal point.

The automation trade-off suggests that to balance the benefits and risks brought by the automation, functions and tasks must be carefully allocated between humans and automation, leading to a function allocation question. Function allocation unifies the DOA models and the human performance consequences. Research about this topic typically aims at hunting for a fixed DOA that is optimal in automation trade-off – trying to add most benefits and keeping risks at the minimal level that are suitable for a specific context (e.g., domain, task). Example domains include cabin life support system (Smith & Jamieson, 2012) and process control (Manzey, Reichenbach, & Onnasch, 2008). It has been found that an intermediate DOA – more specifically, closest to manual control (Endsley & Kaber, 1999) - may have more human performance benefits over higher DOAs.

Adaptive Automation

An emerging, alternative approach to function allocation is *adaptive automation*, a context-sensitive approach. Adaptive automation sets up a dynamic allocation scheme, with dynamically manipulated DOA to changes in a physical environment (Mouloua & Parasuraman, 1994; Parasuraman & Riley, 1997; Sheridan, 2011). An allocation authority agent, either automation or a human supervisor, is the core of this allocation scheme. The allocation authority agent has been added to the control loop to distribute the tasks between the human operator and automation.

Adaptive automation, in its narrow sense, requires the allocation authority agent to be automation to trigger changes in function allocation. These automated systems typically use automation at a higher DOA to achieve better routine performance. When a disturbance occurs in the environment, the automated systems use a lower DOA for the human operator to gain better awareness which is instrumental to save a system failure.

On the other hand, if the allocation authority agent is a human supervisor, this allocation scheme would be more precisely called *adaptable automation*, in which case, the human supervisor must require the automation to return the control responsibilities whenever necessary. Over the past decade, adaptive (and adaptable) automation has become an interest in automation literature. The results of using adaptive automation may be improved situation awareness and a higher and acceptable DOA (e.g., Calhoun, Ward, & Ruff, 2011; Kaber & Endsley, 2004). There are also arguments, for example, adaptable automation may bring additional workload to humans because humans are responsible to allocate the system functions, and clearly, the human supervisor that is critical in adaptable automation must have expertise to facilitate choosing an appropriate function allocation (Miller & Parasuraman, 2007).

Summary and Connections to Research Questions

Here the author revisits the research questions of this dissertation:

Research question 1: How can we model automated trading systems with a variable DOA using CWA?

Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

This second part of the literature review looked at the main themes of the human-automation interaction research. The DOA models provide useful guides to designing automated systems; however, as Pritchett et al. (2014) pointed out, these models may not have the sufficient resolution to represent complex situations where functions are allocated differently within the same DOA and therefore, these models can be misleading to designers in some applications. The author argues though, the DOA models can be improved by abstracting knowledge from these models to extend CWA, a modeling tool that is more suitable for handling complex situations (for example, the variable DOA situation in the automated trading case). In the next part of the literature review, the author will review the CWA, an analysis approach. Together, the CWA and the DOA models provide a better approach to research question 1. This theoretical work involves extending the CWA approach and has been reported in Part B of this dissertation.

This human-automation interaction literature review further reviewed performance consequences, automation trade-off, and adaptive automation. These contents are useful for designing an experimental approach to research question 2 and 3, by examining whether the ecological displays, which derived from the extended CWA models with the DOA information incorporated, would have a performance advantage over conventional displays. If this performance advantage exists, as Wickens and colleagues recently suggested (2010), it would be interesting to further examine whether these ecological displays can reduce, or even reverse the automation trade-off.

Research question 3 adds new perspective to the experimental approach for studying human-automation interaction. Risk preference in an automated trading setting has not yet been evaluated in the literature, and apparently, studying this topic in an experimental setting would make unique contributions to this topic.

Cognitive Work Analysis

CWA is an analytic approach to cognitive system engineering (Rasmussen et al., 1994), a field emerged in the response to the accidents in complex socio-technical systems (e.g., the Three-Mile Island nuclear power plant accident in 1979). The complexity of these systems results in situations that system designers had not anticipated; whereas, in these situations, human operators must be provided with additional support for taking appropriate actions while they are monitoring the system. CWA has been useful in dealing with such complexity at an earlier phase of the system design. In Part A of this dissertation, the author has explained that automated trading is a socio-technical system that is underexplored in the literature; therefore, naturally, one goal of writing this dissertation is to explore the use of CWA in this domain and brings similar benefits of the CWA to traders' monitoring performance.

CWA is a framework featuring five interrelated analysis phases, looking at different facets of the complexity of a system (Vicente, 1999). The five analysis phases start from a first phase providing a broad description of the domain fundamentals (work domain analysis, WDA), with subsequent phases detailing tasks being conducted by the operator (control task analysis, ConTA), various strategies adopted in performing the tasks (strategy analysis, StrA), the functional distribution amongst all operators or roles (social organization and cooperation analysis, SOCA), and cognitive requirements for the operator to behave (worker competency analysis, WCA). Most of the five analysis phases have one or more dominant modeling tools that have been widely used in the literature. The five analysis phases of the CWA retain an enormous richness of the information provided to system analysts; whereas in practical use, there is a trade-off between the time and resources that allocated to the work analysis and the values brought by the work analysis to the later system design. Therefore, it is not uncommon to perform some, but not all phases of the CWA.

This literature review presents an overview of the current state of the CWA. In this chapter, all five analysis phases are being introduced with the goal of providing the readers with a full picture of the CWA. The focus of this overview is in the WDA, the ConTA, and the SOCA for the objective of each analysis phase, the most commonly used modeling tools, and the extensions being done to these tools. To be sure, the author reviews the literature on the very few examples of applying the CWA to automated systems and finance, two topics related to this dissertation research. The author remarks on the opportunities in the current modeling tools that can be further developed to facilitate the development of new modeling approaches. Being introduced in Part B, the modeling stage of this dissertation research takes the view of functional distribution that the SOCA suggests, and in practical use, focuses on extending the first two phases, the WDA and the ConTA, which have the most maturity and may be most suitable for this first attempt to model an automated system with a variable DOA.

Further, to facilitate the design and evaluation stage of this dissertation research (being reported in Part C), the author brieflys introduce EID, the accompanying design approach for graphically representing the rich information identified with the CWA. In particular, the author reviews previous works using EID for designing displays for automated systems and financial systems, and suggest new opportunities for further research in this field.

Five Analysis Phases

Work Domain Analysis

WDA is the first phase of the CWA, and it plays a fundamental role in identifying the system boundary and specifying the relationships between functions and constraints in an abstraction-decomposition space (Rasmussen, 1979). Rasmussen's abstraction-decomposition space represents the work domain of a complex socio-technical system in two dimensions. The first dimension is a part-whole, decomposition hierarchy, ranging from the broadest level of "system" to the most detailed level of “component”. WDA is most frequently performed with the second dimension only, as a modeling tool: a means-end abstraction hierarchy (AH), detailing functional distribution of system functions and constraints (Vicente, 1999). In a typical AH, each work domain function or constraint is illustrated in a *box* and is placed at a designated *level of abstraction*. There are five levels of abstraction in an AH (**Error! Reference source not found.**): the level of *functional purpose* stays at the top of the abstraction, howing major objectives of the domain; the *abstract function* level explains the priorities, values and principles that can be used to achieve the functional purposes; the intermediate level, *generalized function*, describes processes to achieve the abstract functions; the lowest two levels, *physical function* and *physical form*, describe functions related to components of the domain and their attributes. As shown in Figure 42, the five AH levels are interrelated, with any two adjacent levels representing a means-end causal relationship, graphically represented as a line connecting two boxes. Typically, this relationship can be interpreted as: a higher-level box shows *why* an interrelated, lower-level box exists; whereas, a lower-level box shows *how* features of an linked higher-level box can be achieved.

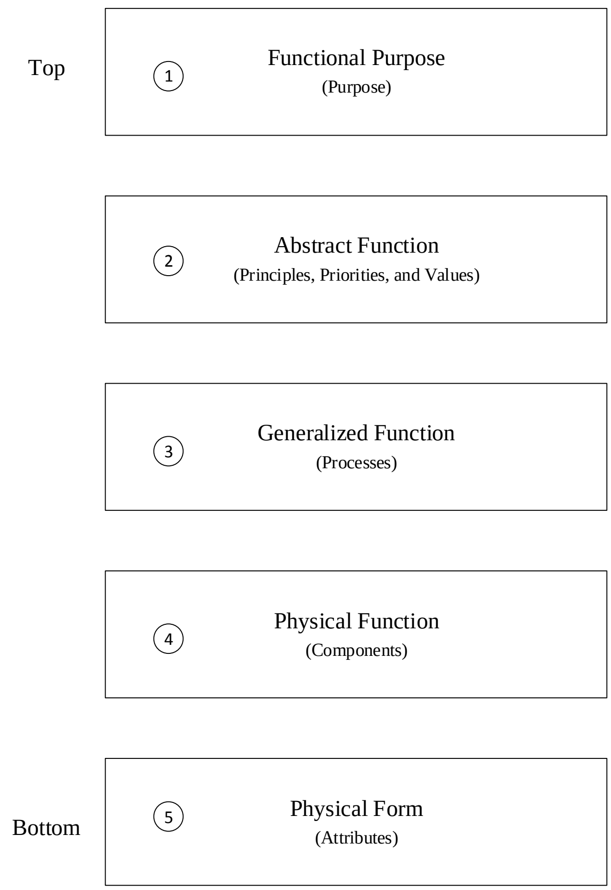


Figure 44. The five-level abstraction hierarchy.

Control Task Analysis

The second phase of the CWA is the ConTA, an analysis for helping analysts understand known, recurring classes of situations in a work domain (Vicente, 1999). The decision ladder (DL, Figure 43) is a tool typically used in a ConTA for modeling information-processing that constructs a *control task* (in CWA terminology, a task that covers all information-processing activities, despite that its name might seem to only include the action stage, see Vicente, 1999, p. 181). A DL extends the traditional, linear-formed representations of information-processing to portray richer information pertaining to expert behaviour. The expert behaviour in complex socio-technical systems, as Rasmussen observed in a nuclear power domain, has a unique feature that is to take efficient shortcuts from one information-processing step to another (1974). Rasmussen structured the DL in a ladder form that is “folded” in the middle stage of the information-processing activity, allowing for shortcuts being illustrated between the left (showing *activation*, in Rasmussen’s words) and the right (showing *execution*) legs of the ladder (Figure 43). Expert behaviour can be portrayed on a DL; on the other hand, the basic stages of the information-processing activity represented on a DL (typically, with annotations referring to a specific context) are independent of how (e.g., strategy) or by whom (e.g., a human operator or automation). These two unique features add more value to the CWA as a multi-faceted analytic approach, making DL the de facto standard tool for performing a ConTA. In the realm of the CWA, a ConTA details the work domain preliminarily studied in its predecessor, a WDA, through a task angle. A control task must act on a work domain, with both physical and functional information provided by the work domain (which can be identified through an AH). This relationship between an AH and a DL was graphically presented in Vicente’s work (1999, p. 193).

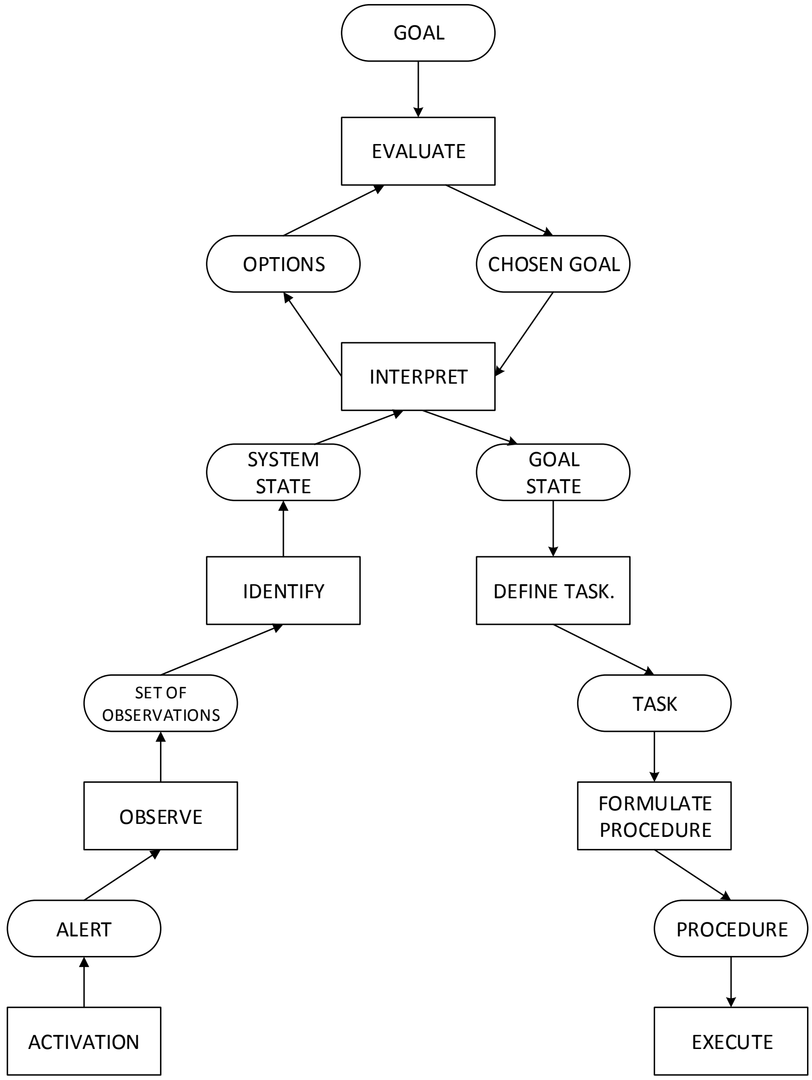


Figure 45. Rasmussen's decision ladder (adopted from Rasmussen, 1974).

Strategy Analysis

The third CWA phase, StrA, studies different *strategies* to perform a control task that has been identified in a prior ConTA. *Strategies* were defined as categories, rather than as instances of cognitive procedures (Rasmussen et al., 1994; Vicente, 1999), to represent how the same control task can be performed in different ways in different situations. These many strategies are operator- and context-dependent, and the StrA, apparently, takes a deeper look at the CWA at a finer level. In the literature, there are remarkably fewers studies using the StrA in comparison to those using the two prior CWA phases, and in general, the StrA is a less developed analysis phase (for a review, see Hassall & Sanderson, 2014). In the literature, the information flow map has been recommended as a modeling tool to perform this analysis (Vicente, 1999), and in practical use, an interview approach can be taken to prompt strategies that are commonly used by operators (Burns, Enomoto, & Momtahan, 2009). Further, although the StrA is theoretically grounded as part of the CWA, formative steps to connect the StrA to the other analysis phases of the CWA have just recently been studied in the literature (e.g., Cornelissen, Salmon, McClure, & Stanton, 2012; Hassall & Sanderson, 2014; Hilliard & Jamieson, 2015). As an example of these recent attempts, Hassall and Sanderson studied the grouping of the categories of strategies by the types of domain functions and constraints, and cognitive procedures involved.

Social Organization and Cooperation Analysis

The fourth phase of the CWA, SOCA, inherits layers of functions and constraints identified in the previous three analysis phases (Vicente, 1999) and addresses function allocation[1]. The SOCA suggests extension opportunities for the modeling tools used in the first three analysis phases of the CWA to represent the allocation of team responsibilities. Most of the extensions focus on modeling a human teamwork environment. Automation must become team players while working with humans (Borst et al., 2015), therefore, a similar team approach should be developed for modeling human-automation coordination.

In the next three subsections, the author provides a brief summary of the existing works in this field. The focus of this summary is the modeling tools that have been developed.

---

[1] As I have discussed about function allocation in automated systems in the introduction chapter of this dissertation (Part A), readers should be aware that Vicente used function allocation in a much broader sense, including both functional distribution of automation and the cooperation of multiple human roles in a teamwork environment.

Extensions to Work Domain Analysis

Responsibility map. A responsibility map can be layered on a work domain model (e.g., abstraction-decomposition space or AH) to show different information needs for various operator roles (Hajdukiewicz, Vicente, Doyle, Milgram, & Burns, 2001).

Collaboration table. Collaboration tables (Ashoori & Burns, 2013) supplement to an AH layered with responsibility map, summarizing shared and individual functions and constraints of a work domain in a tabular form.

Dual abstraction hierarchies. Functional distributions to a human operator and automation can be represented on two separate AHs (Mazaeva & Bisantz, 2003, 2007).

Extensions to Control Task Analysis

Chained (dual) decision ladders. Different roles (e.g., humans or automation) can be represented on multiple DLs, with each DL showing a single control task and interactions between these roles portrayed on the "chains" that connect these DLs (Mazaeva & Bisantz, 2007; Rasmussen et al., 1994).

Decision wheel. The chained decision ladders can be extended to a decision wheel to reduce the complexity of modeling large teams. A decision wheel contains multiple slices, with each slice showing a control task for a specific role (Ashoori & Burns, 2010, 2013). Different decision wheels represent different teams, with communications different in synchronicity (synchronous or asynchronous) and scope (inter- or intra-team) portrayed.

Contextual activity template. Function changes over different situations and operator roles can be illustrated on a tabular, contextual activity template (Jenkins, Stanton, Salmon, Walker, & Young, 2008; Naikar, Pearce, Drumm, & Sanderson, 2003).

Extensions to Strategy Analysis

Colour-coded information flow map. An information flow map can be colour coded and mapped to a contextual activity template to represent different strategies (Jenkins, Stanton, & Walker, 2009). Extensions to the StrA in this area are in general less developed in comparison to those to the WDA and ConTA.

Worker Competencies Analysis

The last analysis phase of the CWA, WCA, identifies the competencies that human operators must exhibit by examining whether the requirements identified in the previous CWA phases are consistent with human limitations and capabilities (Vicente, 1999). The skills, rules, and knowledge taxonomy (SRK, Rasmussen, 1976) is the most-used modeling tool to help decide appropriate requirements for the functions extracted from earlier CWA models. Skills of a human operator represent sensory-motor performance, typically in a simple feedback control task; rule-based behaviour takes place in a situation where there is a known, one-to-one correspondence between the situations and actions (e.g., a look-up table); knowledge is related to sophisticated decision-making, and in many cases, the human operator requires support from the information system to deal with unfamiliar problems.

Cognitive Work Analysis for Automated Systems

Automation is typically modeled with the SOCA to facilitate the representation of function allocation. CWA has been well-established to model complex socio-technical systems. Although Vicente has considered automation as an important aspect of complex socio-technical systems and identified that computer algorithms may direct the tasks in an automated work domain (Vicente, 1999), there have not been many studies discussing how to treat automation within the CWA.

The author reviews these studies in the following subsections. As part of this review, the author revisits some of the approaches discussed in the previous SOCA review which are specifically used to model automated system. The following review overlaps with the previous SOCA but has a different focus on the occasions where each approach fits, and opportunities to develop new approaches.

Modeling Automated Systems with Work Domain Analysis

There have been two distinct approaches for modeling automation with the WDA. The first approach considers automation as a work domain constraint, not as a physical component. This approach is essentially consistent with the traditional way of using the SOCA, after the WDA and the ConTA have been performed (Vicente, 1999). the author has previously reviewed the responsibility map, an extension to the WDA that follows this first approach. However, studies in this field generally looked at human-human interactions to develop a representation of shared, human responsibilities (Hajdukiewicz et al., 2001). An extension to this team approach for describing how work is distributed over humans and automation is yet to be developed.

The second approach treats automation as a system component and explicitly models the automation in an AH that is being shared with other non-automated components. For example, during his candidacy, the author took this second approach to analyze a physiotherapy work domain where an automated motion-tracking device was incorporated to provide physiotherapists and patients with additional, quantitative information regarding the patients' motion (Li, Burns & Kulic, 2014). In the AH the author developed, the automation was described as components at the lower WDA levels (e.g. physical forms) but not at the higher WDA levels. In other words, the automation shared the same values and principles with the humans. A better-documented example using this approach is the modeling of a U.S. warship with the WDA, reported by Burns, Bisantz and Roth (2004). In this U.S. warship model, the sensor system of the warship was explicitly modeled as a physical function in the AH. Burns et al. compared the U.S. warship model to another WDA model, a Canadian frigate model. In the Canadian frigate model, the sensor system was left out of the AH. Burns et al. remarked that modeling sensors as components, as in the U.S. warship model, inherits the original structure of the WDA approach and keeps the model as broad as possible. This "modeling automation as component" approach clearly implies for understanding the current sensors and designing displays that can help with the management of the current sensors; whereas, leaving out the physical structure of the automation, the Canadian frigate model is robust enough to allow users considering un-sensed threats that are not currently picked up by the automation. A variation of this second approach, a *dual-model* approach, is the newest and probably the most developed approach representing the view of modeling automation as component. The dual-model approach was first introduced in Burns et al.'s work, including building a broad AH independent of all control methods and another AH that specifically represents the automation. The dual-model approach was best described in Mazaeva and Bisantz's study on a camera system (2007). A full review of the pros and cons of this dual-model approach was provided in section 2.5.1 in the Part B of this dissertation.

In summary, the first approach, modeling automation as a domain constraint is distinct from the second approach which treats automation as a component. As Burns et al. commented, these two approaches should be used in different occasions. The first approach does not include automation as a component and therefore, may be suitable at an early stage of the automation design to provide directions of future design. The second approach is better in modeling how humans and automated components interact, however, as Burns et al. remarked, it requires a well-defined DOA which is typically unavailable at an early stage of automation design.

Modeling Automated Systems with Control Task Analysis

The ConTA has generally been used to model human information-processing sequence, with differences in expert and novice behavior portrayed as shortcuts on a DL. For modeling automation, existing works typically involve developing an automation-specific DL, showing a control task that has been conducted by the automation. The trail blazer work in this area is Rasmussen and Goodstein's work (1987), using three DLs to represent three different controls tasks for a human designer, a human operator, and a computer. Horiguchi, Burns, Nakanishi, and Sawaragi (2013) presented DLs related to automated finishing mill production and did not include a human operator DL. Despite that the DOA models also describe automation spans all information-processing stages, no knowledge elicited from the DOA models has been transferred to CWA.

Modeling Automated Systems with Strategy Analysis

So far, no extension to the StrA has been made in the literature for modeling automation. However, in the original introduction of the StrA by Vicente (1999), there is an example of allocating resource-demanding aspects of a system troubleshooting strategy from humans to automation. Since StrA identifies that the cognitive procedures to complete a task are different with different operators and under different circumstances, there might be an opportunity in using this analysis phase to understand how a variable DOA might affect the choice of strategies.

Modeling Automated Systems with Worker Competencies Analysis

Acknowledging the recent advances in intelligent automation, Sheridan (2017) proposes that SRK may also be used as a model of how modern automation behaves, and this proposal may extend the scope of the WCA from analyzing the competencies of humans to those of automation. According to Sheridan, skills of automation are subjected to the direct mapping of data sensing to actions, which is typical in classical feedback control; rules are related to the analysis of situations and the selection of control parameters; knowledge extends the use of automation to the decision-making stage and is influenced by human supervision, suggesting the capabilities of the intelligent automation, and the possible roles of humans in such automated systems. In practice, Sheridan's SRK representation for automation, apparently, can only be achieved when the complexity in the modes of automation behaviour becomes transparent to designers, through the use of the preceding CWA phases to analyze automation.

Cognitive Work Analysis for Finance

It should be noted that in Vicente's introduction to CWA (1999), a systematic introduction of CWA toward better computer-based information systems, automated trading was used as an example to characterize the complexity of sociotechnical systems. However, there are only a few examples of using CWA in the finance domain. Achonu and Jamieson (2003) modeled portfolio management mutual fund with a modified WDA. Their model did not include any physical components, as "a portfolio is not primarily a physical system", which is probably not case in automated trading. Achonu and Jamieson removed the physical function level from the AH, the second lowest level from the original AH. Dainoff, Dainoff and McFeeters (2004) used a modified five-level AH model inspired by Reising and Sanderson (2000b)to guide the interface design of a company value judgment supporting tool for investors.

Both examples look at aggregating the vast amount of information to facilitate fundamental analysis. So far, no application of the CWA has been found in supporting automated trading in which technical analysis and sophisticated physical structure of technologies are being used.

Ecological Interface Design

CWA delivers unique products that not only make contributions to understanding and modeling a complex system. Based on the literature, deliverables of the WDA (e.g., functions, constraints, and means-end relationships) and the WCA (e.g., the identifications of SRK behaviours) were typically used to derive appropriate graphical forms for designing ecological displays that can improve monitoring performance (EID, see Hajdukiewicz & Burns, 2004; Vicente, 2002). However, knowledge of the other CWA analysis phases was rarely used in EID.

In practical use, EID first extracts functions and constraints from an AH to generate variables and limits of graphical objects (e.g. indicators and screens). After that, EID organizes these graphical objects in an integrated manner that take into account the capabilities and limitations of the human operator according to the AH levels. EID has been experimentally shown to improve human operator performance (Lau, Skraaning Jr., Jamieson, & Burns, 2008; Lau, Skraaning Jr, et al., 2008). Ecological displays could also enhance human operators' SA without adding workload, in situations where procedural support is not well provided by the system (Burns et al., 2008).

Ecological Interface Design for Automated Systems

Computer displays in automated systems have a conventional role of presenting information. This conventional role can be characterized as the displays carry outputs from machine to human. An automated system is multi-faceted and typically has various *system modes* (e.g., system start-up, normal operation, or failure situations in a process control system). As the complexity of the system grows, there are vast amounts of information to be presented, corresponding to the status of the system and the behaviour of the automation. The increasing amount of information has resulted in the effort of distributing the indications over a great number of displays, with each display representing a specific system mode (Sarter et al., 1997). It has been suggested that an integrated display that replaces the distributed displays may help human operators detect failures quickly (Sheridan, 1992). Such displays are designed in a way to support direct perception that requires little or no additional effort in probing information of system status through a glance.

The call for investing on integrated displays for automated systems has suggested EID as a suitable approach to guide the development of integrated displays (Borst et al., 2015; Sheridan & Parasuraman, 2006). Relatively high DOA is one characteristic of complex socio-technical systems, and EID is inherently a design approach to these systems. The foundational work in this field was the two examples given by Furukawa and Parasuraman (2003) to examine the efficacy of EID displays in an automated system. The first example compared pilot performance difference of an integrated display and a distributed display in a failure detection task in a cockpit simulation. The benefit of EID lies in displaying the deep functional structure of human-machine systems (see also Borst et al., 2015). The second example was conducted on a simulated heated water supply plant, modified from DURESS, the earliest microworld for testing ecological displays. The results of this experiment showed that ecological displays may be particularly helpful to improve human performance under system failure conditions.

An opportunity lies in evaluating ecological displays with new forms of automation to align with emerging topics in the human-automation interaction research. Adaptive automation is emerging but there are only a few empirical design guidelines based on meta-analysis results (Kaber et al., 2001). Kaber et al. suggested that the display should provide humans with information relevant to the transition of system modes (e.g. from one degree of automation to another). The effect of ecological displays to support this type of automation has not been addressed in the literature.

A new but probably fairly important suggestion in the literature is to explore whether ecological displays could influence the automation trade-offs, involving routine performance, automation failure performance, workload and situation awareness (Wickens et al., 2010). Wickens et al. suggested that it might be possible to design automation that can minimizing the system failure costs while keeping the performance benefits, using a properly develop display. Such display may mitigate or even eliminate the automation trade-offs (for example, the system remains at high DOA to keep good routine performance but results in little or no degraded failure performance). EID was, again, suggested as a potentially possible design approach to develop such displays which may help the automated system to “buffer high automation degree from human performance costs when things fail” (Kaber, Hancock, Jagacinski, Parasuraman, Wickens, Wilson, Bass, Feigh & Ockerman, 2011). However, to date, no studies have been conducted to explore this line of research.

Ecological Interface Design for Finance

A recent review showed that EID has been mostly applied to time- and safety-critical domains, mostly in aviation, medicine, power generation and road transport domains (McIlroy & Stanton, 2015). There is no mention of automated trading, financial trading, or finance in general as a domain of EID application in McIlroy and Stanton's review.

Dainoff, Dainoff and McFeeters's work (2004), looking to design an ecological display for an investment tool to support fundamental analysis which was on a par with technical analysis used by most trading algorithms, has been captured in McIlroy and Stanton's review, but has not been categorized as a unique application domain. As the author reviewed previously, Dainoff et al.'s work (2004) involves an AH of fundamental analysis and based on the AH, designed several mock-up designs. On the other hand, Dainoff et al. documented several challenges in working toward an ecological display for a financial system, with which the author agreed according to the author's experience designing automate trading software for Quantica Trading. They stated that EID principles (e.g., AH and SRK) facilitate an "logical analysis" rather than "empirical research" to guide the interface design. In fact, no empirical research has been done, as "there is no competing product available". Further, it was not feasible for them to perform an experimental verification. Dainoff et al.'s displays were part of a commercial product and was submitted as a patent application at the time of writing (Dainoff & Dainoff, 2003).

Summary and Connections to Research Questions

Here is a revisit to the three research questions:

Research question 1: How can we model automated trading systems with a variable DOA using CWA?

Research question 2: Do ecological displays have an advantage in supporting financial trading performance? If so, in which DOA does this advantage exist?

Research question 3: Can ecological displays influence trader's risk preference? If they can, in which DOA does this influence exist?

In this part of the literature review, the author has introduced the five analysis phases of the CWA and the EID, a design framework derived from the CWA, and reviewed the examples and extensions for modeling automated systems and financial systems. Together, this chapter provided profound knowledge that will be used in developing the theoretical and experimental works of this dissertation research.

The comparison of the two WDA approaches, the "modeling automation as constraints" approach and the "modeling automation as components" approach, called for new approaches that can handle the increasing complexity in automated systems. As the authro discussed in the introduction section of this dissertation (Part A), the increasing complexity in the automated trading domain, in particular, lies in the heavy coupling between humans and automation and the increasing flexibility in developing trading algorithms. It is typical for an established automated trading system to have a variable DOA, in which case, it is equally important to model the analysis as broadly as possible (achieved by "modeling automation as constraints"), and representing which functions and components of the system are allocated to the automation with a high resolution (achieved by "modeling automation as components"). At least in this case, there is a need to develop a hybrid approach that fulfills these requirements.

The review of the ConTA shows this analysis phase of the CWA is under-developed for modeling automation and therefore, the possibility to develop new approaches should be explored. The ConTA deals with patterns in human and automated information-processing, and naturally, this analysis phase can be correlated to the existing DOA models which describe similar problems.

The StrA, as the author reviewed, has not been adequately mutual to be used as a formative approach to achieve the author's specific modeling goal. However, the modeling stage of this dissertation research should provide useful insights into the StrA, by developing richer WDA and ConTA models that can be utilized by the StrA.

For research question 2 and 3, the WCA and the EID reviewed in this chapter guided the design of the ecological displays based on the new modeling approach. Challenges described by Dainoff et al. (2004) in the practical use of EID in the finance domain may be true in the case of automated trading. The finance industry strictly protects the confidentiality of institutional clients – much more strictly than other domains the author's worked with (e.g., aviation and healthcare) during his candidature. Careful considerations have been made to compensate for these challenges in this dissertation research.

Appendix B

Recruitment Letter

UNIVERSITY OF WATERLOO

PARTICIPANTS NEEDED FOR RESEARCH IN

Effects of Display Designs on Financial Trading Awareness and Performance

Hello,

My name is Yeti Li and I am a PhD student in Department of Systems Design Engineering at the University of Waterloo. I am supervised by Professor Catherine Burns in Department of Systems Design Engineering.

We are conducting a study on the role of display designs on financial trading awareness and performance on a financial trading simulator. As a participant in this study, you would be asked to:

- Trade financial products on a financial trading simulator;

- Complete questionnaires about how aware are you of market, portfolio and trading executions in simulated trading.

To better understand your behavior during the scenarios, a non-intrusive Gazepoint GP3 Eye Tracker (http://www.gazept.com/product/gazepoint-gp3-eye-tracker/) will be used in this study. Eye tracker uses cameras to identify where you are looking on the computer screen. Through examination of eye-tracking data, we may find the causes for your behaviour on the simulator without relying on the fallible human memory.

Recruitment Criteria

Participants should:

- be a University of Waterloo undergraduate or graduate student;

- have successfully completed at least 1 computer programming course;

- have a normal or corrected normal visual acuity (e.g. wearing glasses or contact lens);

- have a normal colour vision;

- be comfortable with using spreadsheet software (e.g. editing a Microsoft Office Excel workbook) and information graphic software (e.g. creating a chart from a provided data set).

The experiment is expected to last 2 hours. In appreciation for your time, you will receive $30 in exchange for your participation in the session. In any cases, the final decision about participation is yours.

I would like to assure you that this study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Committee. However, the final decision about participation is yours. If you have any comments or concerns resulting from your participation in this study, please contact Dr. Maureen Nummelin, the Director, Office of Research Ethics, at 1-519-888-4567. Ext. 36005 or maureen.nummelin@uwaterloo.ca.

If you are interested in participating, please contact:

Sincerely,

Yeti Li

Student Investigator

PhD Student, Department of Systems Design Engineering, University of Waterloo

Email: yeti.li@uwaterloo.ca

Appendix C

Screening Questionnaire

Dear participant,

We would like to thank you for your interest in this study. Please answer the following demographic questions. This information is used to guide us with your entry into the study. Financial trading software uses various colours for data visualizations, so you should have a normal or corrected normal visual acuity and a normal colour vision. This study involves interacting with a computer based trading algorithm. Therefore, it requires having a prior knowledge of computer programming. Since the experiment software contains financial charts and lists, you should be comfortable with using spreadsheet software and information graphic software.

1. Age: ___________________

2. Do you have normal or corrected normal visual acuity (e.g. wearing glasses or contact lens)?

____

3. Do you have a normal colour vision? ___________________

4. Your current degree program is ___________________ in department of ___________________

with a minor or option of ___________________

5. Have you successfully completed at least one computer programming course? ___________________

Indicate how much you agree or disagree with each statement, using a scale ranging from 1 = strongly disagree to 5 = strongly agree.

6. I am comfortable with using spreadsheet software (e.g. editing a Microsoft Office Excel workbook).

____

STRONGLY DISAGREE 1 2 3 4 5 STRONGLY AGREE

7. I am comfortable with using information graphics software (e.g. creating a chart from a provided data set). ___________________

STRONGLY DISAGREE 1 2 3 4 5 STRONGLY AGREE

At this stage, we only collect necessary information to identify whether you meet the criteria of this study. Answers from participants who are not invited will be disposed immediately. In the meantime, if you have any questions about the study, please do not hesitate to please contact the researcher Yeti Li at yeti.li@uwaterloo.ca or by calling the research lab at 519-999-4567 Ext. 35874. You may also contact Professor Catherine Burns at 519-888-4567 Ext. 33903 (catherine.burns@uwaterloo.ca) or the University of Waterloo Counseling Services at 519-888-4567 Ext. 32655.

As with all University of Waterloo projects involving human participants, this study was reviewed by, and received ethics clearance through a University of Waterloo Research Ethics Committee. Should you have any comments or concerns resulting from your participation in this study, please contact Dr. Maureen Nummelin, the Director, Office of Research Ethics, at 1-519-888-4567, Ext. 36005 or maureen.nummelin@uwaterloo.ca.

Sincerely,

Yeti Li

Student Investigator

PhD Student, Department of Systems Design Engineering, University of Waterloo

Email: yeti.li@uwaterloo.ca

Appendix D

Information Letter

UNIVERSITY OF WATERLOO

INFORMATION LETTER

Effects of Display Designs on Financial Trading Awareness and Performance

Faculty Supervisor: Catherine Burns (catherine.burns@uwaterloo.ca, 519-888-4567 Ext. 33903) , Systems Design Engineering.

Student Investigators: Yeti Li (yeti.li@uwaterloo.ca, 519-999-4567 Ext. 35874), Systems Design Engineering

Study Overview

You are invited to participate in a study examining the effects of display designs in a financial trading software. We are interested in understanding the role of display designs on awareness performance in a financial monitoring task. This study will involve a few questionnaires and interactions with a computer simulator in a lab setting.

What You Will Be Asked to Do

After your consent, you will be asked to complete a series of surveys about demographic information. You will then be trained to use our simulator. This will be followed by going through 4 different scenarios; while you are going through each of the scenarios, a few questions about your experience with the simulation will be provided to you. This is not a test of your knowledge or intelligence but an opportunity for us to identify display design effects in the software simulation that you will interact with.

To better understand your behavior during the scenarios, a non-intrusive Gazepoint GP3 Eye Tracker (http://www.gazept.com/product/gazepoint-gp3-eye-tracker/) will be used in this study. Eye tracker uses cameras to identify where you are looking on the computer screen. Through examination of eye-tracking data, we may find the causes for your behaviour on the simulator without relying on the fallible human memory.

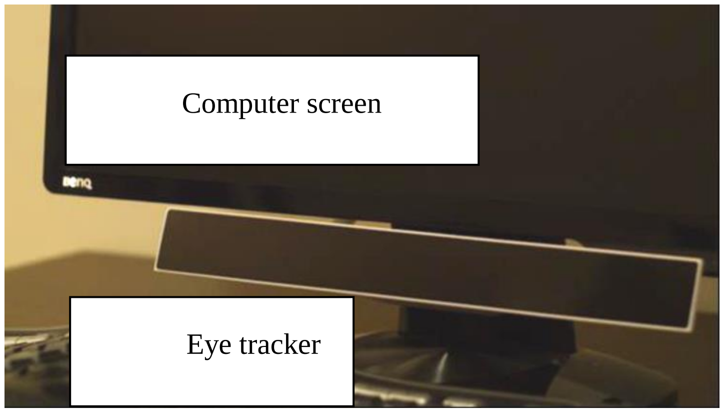


Figure 46. Gazepoint GP3 eye tracker.

Participation and Remuneration

Participation in this study is voluntary, and will take approximately 2 hours of your time. You may decline to answer any questions presented by the experimenter. Further, you may decide to withdraw from this study at any time by advising the researcher, and may do so without any penalty or loss. You will be paid $30 for your participation in this study even if you decide to withdraw your consent at any time. The amount received is taxable. It is your responsibility to report this amount for income tax purposes. According to University of Waterloo Finance polices, the investigators and participants are required to complete a University of Waterloo Research Participant’s Acknowledgement of Receipt of Remuneration and Self-Declared Income when the remuneration is provided.

Personal Benefits of the Study

You will be familiarized with financial software and basic financial terminology through our simulated task.

Risks to Participation in the Study

The risks associated to participation in this study are minimal. There are no known or anticipated risks or stressors that may be characterized as physiological, psychological, emotional, social or economic in nature other than any risks normally experienced on a day-to-day basis. The potential risks if any would not exceed that of using spreadsheet software and information graphics software (e.g. Microsoft Office Excel). In the event that any problems develop, please contact the researcher Yeti Li at yeti.li@uwaterloo.ca or by calling the research lab at 519-999-4567 Ext. 35874. You may also contact Professor Catherine Burns at 519-888-4567 Ext. 33903 (catherine.burns@uwaterloo.ca) or the University of Waterloo Counseling Services at 519-888-4567 Ext. 32655.

Confidentiality

Your data will be kept confidential. The data stored will be linked to specific participant identifiers (e.g. Participant ID1). All survey data and performance data will be linked to participant identifiers only. A document linking unique identifiers of the participants (e.g. name, major, email address) to the participant identifier will be maintained and will not be stored in the same location as the raw data thereby rendering your data anonymous; this document will be accessible only to authorized personnel. The raw data will not leave the University of Waterloo,

Questions and Research Ethics Clearance

If after receiving this letter, you have any questions about this study, or would like additional information to assist you in reaching a decision about participation, please feel free to ask the investigators or the faculty supervisor listed at the top of this sheet.

I would like to assure you that this study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Committee. However, the final decision about participation is yours. If you have any comments or concerns resulting from your participation in this study, please contact Dr. Maureen Nummelin, the Director, Office of Research Ethics, at 1-519-888-4567. Ext. 36005 or maureen.nummelin@uwaterloo.ca.

Thank you for your interest in our research and for your assistance with this project.

Appendix E
Consent Form

UNIVERSITY OF WATERLOO

INFORMED CONSENT BY SUBJECTS TO PARTICIPATE IN A RESEARCH EXPERIMENT

Effects of Display Designs on Financial Trading Awareness and Performance

I have read the information presented in the information letter about a study being conducted by Yeti Li under the supervision of Professor Catherine Burns of the Department of Systems Design Engineering at the University of Waterloo. I have had the opportunity to ask any questions related to this study, to receive satisfactory answers to my questions, and any additional details I wanted.

I am aware that I may withdraw my consent for any of the above statements or withdraw my study participation at any time without penalty by advising the researcher.

I am aware that I may be asked to provide demographic information including name, age, gender, vision, major, courses previously completed, experience in using spreadsheet software, information graphic software and in financial trading.

I am aware that my eye activity on the computer screen will be measured by a non-intrusive Gazepoint GP3 Eye Tracker.

This project has been reviewed by, and received ethics clearance through a University of Waterloo Research Ethics Committee. I was informed that if I have any comments or concerns resulting from my participation in this study, I may contact the Dr. Maureen Nummelin, Office of Research Ethics at 519-888-4567 ext. 36005, maureen.nummelin@uwaterloo.ca.

By signing this consent form, you are not waiving your legal rights or releasing the investigator(s) or involved institution(s) from their legal and professional responsibilities.

	Please Circle One		Please Initial Your Choice
With full knowledge of all foregoing, I agree, of my own free will, to participate in this study.	**YES**	**NO**	____

Participant Name: ____ (Please print)

Participant Signature: ____

Witness Name: ____ (Please print)

Witness Signature: ____

Date: ____

Appendix F

Demographic Questionnaire

Participant ID ____________________ Date ____________________

Demographics

1. Gender: ____________________________

2. Please estimate the average time that you usually spend on using computers each week (including desktop and portable computers, not including mobile electronic devices such as tablets and cellphones): _____________________ hours/day.

3. How many years have you been using computers? ___________________ years.

Experience with financial trading

1. Do you have experience with financial trading (e.g. buying a stock)? ___________________

If yes, please specify your experience is from ___________________

a. Academic Work

b. Finance Industry

c. Personal investment

d. Other (please specify) ______________________________________

Appendix G

Risk Preference Survey

Identifying emerging trends

If there is a micro trend that the market will move to one direction, I would more likely:

a. immediately place the position (buy) for the maximum profitability;

b. wait until the market direction is clear.

Responding to trend reversals

If there is a change in the market direction after a position has been placed (buy), I would more likely:

a. immediately close the position (sell) to minimize losses;

b. believe this reverse trend is only momentary.

Detecting regime shifts

If the market has been in a shock (regime shifts) for quite a while (5 minutes, in the context of this experiment), I would more likely:

a. immediately place the position (buy), as I believe the market will break the shock and the space for uptrend has been opened up;

b. wait longer, as I still believe the market is in the shock.

Taking action following sudden interruption to supply

If the market is collapsing (e.g. market crash), I would more likely:

a. immediately close out all or most positions, or do nothing;

b. promptly buy back the same financial product to lower the average portfolio price.

Appendix H

Unweighted NASA-TLX Survey

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Appendix I

Eye Calibration Criteria

From: Johnny
Sent: November 16, 2015 12:03 PM
To: Yeti Li
Cc: Catherine Burns; Dev Minotra
Subject: Re: Gazepoint GP3 calibration

Hello Yeti,

A good calibration is one in which at least 4 but preferably all (5 or 9) points are detected for both the left and right eyes. Accuracy should be around 1 degree of visual angle which depends on screen size/resolution but for common values roughly 1-2 cm or 40-80 pixels, though smaller is better.

Regards,

Johnny
Gazepoint
www.gazept.com

On 11/13/2015 1:02 PM, Yeti Li wrote:

> Dear Gazepoint,
>
> I am a researcher at the University of Waterloo, Canada. I am using GP3 Eye Tracker in a human-computer interaction study. I especially like the simplicity of the set-up and the software. Great product!
>
> After the completion of each calibration the Gazepoint software provides a result of "x of 5 points" (in a 5-point calibration). Assuming "5 of 5 points in both eyes" means a good calibration, how would you define a bad calibration and therefore a re-calibration should be done?
>
> If possible, could you kindly provide an example of paper in which the criterion of GP3 calibration is discussed?
>
> Many thanks.
>
> Yeti Li
> PhD Candidate
> Systems Design Engineering
> University of Waterloo

Appendix J

Supplementary Notes on Normality Tests

Most measures in the two experimental studies were analyzed using non-parametric tests because the assumption of normality was violated in all data groups, *ps* < .05.

According to the suggestion of Holmqvist et al. (2011), total dwell time in common AOIs was log transformed and submitted to a repeated measures ANOVA. The normality test was performed for each experiment to examine how well the normality of the eye-tracking data was improved. The repeated measures ANOVA was proceeded with the transformed data for a robust estimation. The perceived workload data were submitted to a repeated measures ANOVA in a way similar to Arrabito et al. (in review). Results of the normality tests are presented in Table 51 to 54.

Table 51. Shapiro-Wilk Normality Test Results for Transformed Total Dwell Time in the Common AOIs (Experiment 1).

Display Type	DOA	*w*	*p*
Market AOI			
Moderate-conventional	Moderate	.912	.126
Moderate-ecological	Moderate	.953	.532
High-conventional	High	.966	.763
High-ecological	High	.941	.359
Portfolio AOI			
Moderate-conventional	Moderate	.904	.092
Moderate-ecological	Moderate	.769	.001*
High-conventional	High	.870	.027*
High-ecological	High	.868	.026*
Trading History AOI			
Moderate-conventional	Moderate	.954	.556
Moderate-ecological	Moderate	.856	.017*
High-conventional	High	.898	.076
High-ecological	High	.868	.025*

Table 52. Shapiro-Wilk Normality Test Results for Transformed Total Dwell Time in the Common AOIs (Experiment 2).

Display Type	*w*	*p*
Market AOI		
Improved-high-conventional	.930	.137
Improved-high-ecological	.955	.418
Adaptive-conventional	.946	.279
Adaptive -ecological	.912	.060
Portfolio AOI		
Improved-high-conventional	.810	< .001*
Improved-high-ecological	.763	< .001*
Adaptive-conventional	.826	.002*
Adaptive -ecological	.632	< .001*
Trading History AOI		
Improved-high-conventional	.769	< .001*
Improved-high-ecological	.750	< .001*
Adaptive-conventional	.851	.004*
Adaptive -ecological	.891	.024*

Table 53. Shapiro-Wilk Normality Test Results for Perceived Workload (Experiment 1).

Scenario Type	*w*	*p*
Mental Demand		
Moderate-conventional	.920	.059
Moderate-ecological	.914	.044*
High-conventional	.962	.488
High-ecological	.934	.121
Physical Demand		
Moderate-conventional	.925	.077
Moderate-ecological	.871	.006*
High-conventional	.847	.002*
High-ecological	.852	.002*
Temporal Demand		
Moderate-conventional	.952	.294
Moderate-ecological	.946	.221
High-conventional	.921	.061
High-ecological	.944	.202
Performance		
Moderate-conventional	.843	.188
Moderate-ecological	.934	.121
High-conventional	.922	.065
High-ecological	.937	.139
Effort		
Moderate-conventional	.899	.021*
Moderate-ecological	.962	.470
High-conventional	.938	.147
High-ecological	.933	.115
Frustration		
Moderate-conventional	.946	.222
Moderate-ecological	.961	.467
High-conventional	.899	.021*
High-ecological	.912	.039*

Table 54. Shapiro-Wilk Normality Test Results for Perceived Workload (Experiment 2).

Display Type	*w*	*p*
Mental Demand		
Improved-high-conventional	.901	.023*
Improved-high-ecological	.909	.034*
Adaptive-conventional	.942	.184
Adaptive -ecological	.960	.440
Physical Demand		
Improved-high-conventional	.890	.013*
Improved-high-ecological	.919	.055
Adaptive-conventional	.831	<.001
Adaptive -ecological	.859	.003*
Temporal Demand		
Improved-high-conventional	.931	.098
Improved-high-ecological	.960	.431
Adaptive-conventional	.948	.241
Adaptive -ecological	.920	.059
Performance		
Improved-high-conventional	.942	.170
Improved-high-ecological	.962	.489
Adaptive-conventional	.935	.124
Adaptive -ecological	.942	.182
Effort		
Improved-high-conventional	.923	.068
Improved-high-ecological	.958	.397
Adaptive-conventional	.945	.215
Adaptive -ecological	.954	.328
Frustration		
Improved-high-conventional	.902	.024*
Improved-high-ecological	.936	.130
Adaptive-conventional	.962	.473
Adaptive -ecological	.947	.228