4-14-2021

Decision-Making Dilemma in Human-Automation Interaction: Who Should Grasp Authority, Human or Intelligent Systems?

Xiaocong Cui

DECISION-MAKING DILEMMA IN HUMAN-AUTOMATION INTERACTION: WHO SHOULD GRASP AUTHORITY, HUMAN OR INTELLIGENT SYSTEMS?

BY

Xiaocong Cui

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree

Of

Doctor of Philosophy

In the Robinson College of Business

Of

Georgia State University

GEORGIA STATE UNIVERSITY
ROBINSON COLLEGE OF BUSINESS
2021

ACCEPTANCE

This dissertation was prepared under the direction of the *Xiaocong Cui* Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Business Administration in the J. Mack Robinson College of Business of Georgia State University.

Richard Phillips, Dean

DISSERTATION COMMITTEE

Dr. JJ Po-An Hsieh (Co-Chair)
Dr. Mark Keil (Co-Chair)
Dr. Likoebe Mohau Maruping
Dr. Liwei Chen (External)

ABSTRACT

DECISION-MAKING DILEMMA IN HUMAN-AUTOMATION INTERACTION: WHO SHOULD GRASP AUTHORITY, HUMAN OR INTELLIGENT SYSTEMS?

BY

Xiaocong Cui

November 17, 2020

Committee Chair: *JJ Po-An Hsieh*

Major Academic Unit: *Computer Information Systems Department*

Overreliance on automation often leads to problems such as human skill decay, a loss of situation awareness, and even casualties. Human-centered automation (HCA) is proposed to keep humans in the loop so that both humans and automation work on the same task. Effective task allocation between humans and automation has attracted a great deal of attention since the introduction of automation technologies, which assist humans in task performance. Yet, the rise of intelligent systems (IntelSys) that are capable of making task-allocation judgments has inspired an urgent inquiry: who (humans or IntelSys) should have the authority to make the decision to allocate tasks between humans and automation?

The current literature proposes three decision-making approaches (DMAs) to allocate tasks between humans and automation in HCA: IntelSys-Decides, Human-Decides, and Intel-Advises approaches. In the IntelSys-Decides approach, IntelSys has full decision-making authority to allocate tasks between humans and automation. In the Human-Decides approach, humans have the sole authority to make task allocation decisions. In the IntelSys-Advises approach, humans have the freedom to make task allocation decisions at any time; meanwhile, IntelSys generates task allocation advice to humans, who can either accept or reject the advice by IntelSys. Given the fact that there are no consistent findings about which approach is the best, it is necessary to identify the best approach, and if necessary, to identify the boundary conditions for the three DMAs.

Drawing on the perspective of team-based decision-making, this study proposes four hypotheses that compare the impacts of three decision-making approaches (DMAs) on human-automation team performance. To test the hypotheses, we conducted a large-scale experiment with 881 participants playing on a gaming platform. The results suggest that the effectiveness of DMAs is contingent on the task uncertainty and human expertise. The findings reveal insights on whether humans or IntelSys should assume the role of decision-maker in human-automation teams under different scenarios.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the guidance, support, help, and encouragement of my dissertation committee, colleagues, friends, and family. First and foremost, I would like to express my most sincere acknowledgment to my advisor Dr. JJ Po-An Hsieh and co-advisor Dr. Mark Keil. I have had the great pleasure of having Dr. Hsieh as my advisor. Many thanks for his countless hours of reading, revising, reflecting, and encouraging me throughout my entire doctoral life, for teaching me how to become a professional researcher, and for giving me excellent advice as he guided me through the dissertation process. A special thanks to Dr. Mark Keil for investing so much valuable time and effort in the development of my dissertation. He has given me excellent professional advice, tireless support and help, and detailed feedback. I am also grateful for the support provided by my committee members, Dr. Likoebe Mohau Maruping and Dr. Liwei Chen. They have contributed insightful comments and suggestions to my dissertation. I must also acknowledge the support I have received from all my colleagues. Thank you Christine Abdalla Mikhaeil, Yukun Yang, Amrita George, Kambiz Saffarizadeh, Maheshwar Boodraj, Zhitao Yin, Alan Yang, Pengcheng Wang, Renzhi Zhao, Youyou Tao, Hongyu Gao, Rongen Zhang, Hyoung Yong Choi, Junyoung Park, Tawfiq Alashoor, Yumeng Miao, Zirun Qi, and Jungmin (Jeremy) Lee.

Last but not least, I would like to express my deepest gratitude to my family, for their love and support throughout my life. I want to thank my mother Nuqin Li, my father Huilong Cui, my brother Congcong Cui, and my nieces Bihua Cui and Bixin Cui. I am deeply grateful for their love and support they have provided throughout my life.

TABLE OF CONTENTS

List of Figures

List of Tables

1 Introduction

1.1 Decision-making in Human-automation Interaction

Automation technologies (hereinafter automation) have been applied to assist humans in a variety of domains such as air traffic control, military command and control, autonomous vehicles, crisis management, and industrial process control (Moray et al. 2000; Arciszewski et al. 2009). However, inappropriately allocating tasks to automation often lead to disasters (Kaber 1996). An illustrative example is the autopilot system of Tesla's cars. There have been at least eight accidents involving Tesla cars since 2016.[1,2] Most of these accidents happened because human drivers allocated too many driving tasks to the autopilot system (automation). Therefore, in the area of human-automation interaction (HAI), the essential question is how to best allocate the task between humans and automation. To answer this question, the first thing to consider is who should make the task allocation decisions. In the context of HAI, other than humans, intelligent systems (IntelSys), which refer to the integration of computer algorithms that mimic the decision-making ability of a human (Jackson 1986), can also offer task-allocation advice or can even make the task-allocation decisions (Gombolay et al. 2016; Lee 2018). Therefore, it is critical to know whether *humans* or *IntelSys* should have the decision-making authority to allocate tasks between humans and automation.

Early research on HAI has advocated technology-centered automation (TCA), aiming to automate every task whenever possible. While TCA recognizes that automation generally outperforms humans in repetitive tasks (Kaber 1996), TCA often encounters human out-of-the-loop problems, such as human skill decay, a loss of situation awareness, and failure to respond quickly during unforeseen situations (Endsley and Kiris 1995; Kaber 1996). To tackle these issues associated with TCA, some researchers have proposed human-centered automation (HCA) to keep humans in the loop so that they are engaged in the task activities (Kaber 1996; Kidwell et al. 2012).

[1] https://www.reuters.com/article/us-tesla-crash/tesla-says-crashed-vehicle-had-been-on-autopilot-prior-to-accident-idUSKBN1H70232
[2] https://en.wikipedia.org/wiki/Tesla_Autopilot

There are three important roles in HCA: human, IntelSys, and automation. *Automation* stands for the device that replaces certain functions previously carried out by humans (Groover 2018; Janssen et al. 2019). The primary function of IntelSys is to allocate tasks between humans and automation (Scerbo 2001; Scerbo 2008). Under HCA, humans and automation both work on the same task with the same goal, whereas humans, IntelSys, or the two together could decide how to allocate tasks between humans and automation (Ruff et al. 2018). When the task is carried out by humans, it is in the manual mode, and when the task is carried out by automation, it is in the automation mode.

As shown in Figure 1.1, the HCA literature centers on three decision-making approaches (DMAs): IntelSys-Decides, Human-Decides, and IntelSys-Advises approaches.[3] In the IntelSys-Decides approach, IntelSys has full decision-making authority to allocate tasks between humans and automation (Scerbo 2008). The IntelSys-Decides approach ensures that automation can be tied closely to human workloads in different situations (Scerbo 2008). For example, if the human's workload is too high, IntelSys will allocate the task to automation; in contrast, if the human's workload is low, IntelSys will allocate the task to the human to keep him or her engaged in the task (so as to keep the human in the loop). For the Human-Decides approach, humans have the sole authority to decide whether the task should be allocated to them or automation (Parasuraman and Wickens 2008). Lastly, in the IntelSys-Advises approach, humans have the freedom to make task allocation decisions at any time; meanwhile, IntelSys generates task allocation advice to the human, who can either accept or reject the advice by IntelSys (Li 2013).

Interestingly, the few existing studies that compare the performance outcomes of the above three DMAs reveal rather inconsistent results (detailed in Section 2.4). For example, Bailey et al. (2006) and

[3] In human automation interaction, the original terms for three decision-making approaches are: adaptive automation, adaptable automation, and hybrid. Adaptive automation is represented by IntelSys-Decides because in adaptive automation, it is IntelSys that makes the task allocation decisions. Adaptable automation is represented by Human-Decides because in adaptable automation, it is the human who makes the task allocation decisions. The hybrid approach is represented by IntelSys-Advises because in Hybrid, IntelSys provides task allocation advice to the human, who then decides to accept or reject the advice. To make these terms more self-explanatory and reader friendly, in this study, I call the adaptive automation, adaptable automation, and hybrid approaches as IntelSys-Decides, Human-Decides, and IntelSys-Advise modes, respectively.

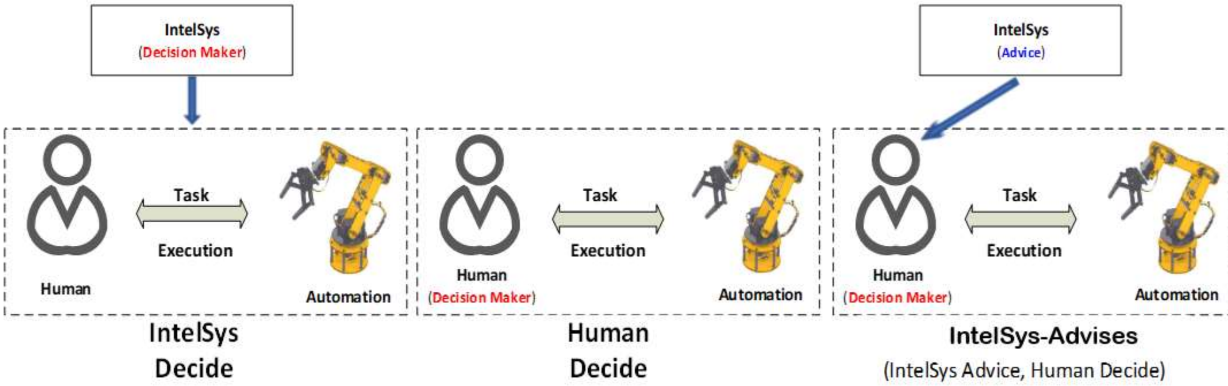


Figure 1.1 HCA Decision-making Approaches

Sauer et al. (2012) empirically verified that in terms of task performance, the IntelSys-Decides approach is better than the Human-Decides approach because the mechanism to trigger the IntelSys-Decides approach considers the underlying engagement states (i.e., workload), which cannot be recognized by humans. However, Wickens (1994) found that humans in the Human-Decides approach demonstrate better situation awareness, which positively impacts the task performance, than in the IntelSys-Decides approach. Sauer et al. (2012) and Kidwell et al. (2012) also found no significant difference between IntelSys-Decides and Human-Decides in term of task performance. Theoretically speaking, the IntelSys-Advises approach should be the best because it has been designed to combine the advantages of the IntelSys-Decides approach and the Human-Decides approach by integrating humans' knowledge and IntelSys in decision making (Li 2013; Li et al. 2013; Ruff et al. 2018). This argument is empirically supported by Li (2013) and Ruff et al. (2018). However, Cabrall et al. (2018) found that the Human-Decides approach exhibits higher task performance than the IntelSys-Advises approach.

The literature review suggests several issues that may have caused such inconsistencies in the HAI literature. First and foremost, previous research lacks a theoretically grounded effort to systematically investigate and compare the effectiveness of these DMAs, which is perhaps the key reason leading to such inconsistent findings. It is also surprising that no prior DMA studies have conceived the human operator and the automation function collectively as a team, despite the rich body of knowledge accumulated in the team-based management and decision-making literature. Moreover, while employing various tasks and different human participants for investigation, these studies have rarely considered how task characteristics (e.g., task uncertainty) and individual characteristics (e.g., human expertise) may independently and interactively affect decision making and task performance. Against this backdrop, this dissertation aims to attain the following two research objectives (ROs):

Research Objective 1: Systematically compare the three DMAs and identify the best DMA for task performance.

Research Objective 2: Examine how task uncertainty and human expertise, as two boundary conditions, influence the task performance of the three DMAs.

1.2 Team-based Decision-making Perspective

Capitalizing on the opportunity to address the knowledge gaps discussed above, I draw on the team management literature to theoretically investigate the phenomenon. In particular, this study draws on the theoretical lens of team-based decision making to (1) examine the effectiveness of the three DMAs by conceptualizing humans and automation as a team; and (2) incorporate task uncertainty and human expertise as the two core boundary conditions for the impacts of the different DMAs on team performance.

Importantly, I propose that it is necessary to study HAI from the team perspective. Humans and automation can be a team because (a) the team perspective fits well with the idea of HCA, which emphasizes that both humans and automation should be involved in the same task; and (b) humans and automation share the same goal to achieve the best team performance. Based on the team-based decision-making perspective, I identify that the team member's expertise and expertise coordination can impact team task performance. Given the fact that automation is designed to outperform the majority of humans, it is reasonable to expect that the expertise coordination principle should allocate most of the task to automation rather than to humans.

The essence of comparing the three DMAs involves comparing the expertise coordination mechanism of IntelSys and human decision makers. IntelSys makes decisions based on the human's workload. The human decision maker, relative to IntelSys, has three advantages in decision making. The human can switch the mode at any time, and therefore is more flexible than IntelSys. Moreover, humans knows themselves better (mood, mental state, emotion, physical condition, etc.) than IntelSys; hence, when making decisions, humans, as compared to IntelSys, can factor more individual characteristics (mood, mental states, physical condition, etc.) into the decision-making process. Humans can also integrate information from different sources (e.g., environment, task feedback, automation performance). However, according to my review of the personal control model, I argue that the superiority of human decisions over IntelSys decisions is not consistent; rather, the relative advantages between human decisions versus IntelSys decisions are contingent on the task uncertainty.

The personal control model indicates that humans' decision-making behavior is affected by task uncertainty. To start with, humans have no problems when making decisions under the condition of low task uncertainty. However, in the case of high task uncertainty, humans with high task expertise (i.e., experts) in the target tasks have a strong desire to seek more control, whereas humans with low task expertise (i.e., novices) are more likely to give up control. In this thesis, by integrating the team-based decision-making perspective and personal control model, I develop a theoretical model that identifies how human decision makers generate decisions under different levels of task uncertainty. In the case of low task uncertainty, human decision makers, including both experts and novices, are more likely to make flexible, informed, and personalized decisions than IntelSys. However, novices and experts will behave differently under the condition of high task uncertainty. In particular, novices are likely to give up control, and therefore allocate more tasks to automation. In contrast, experts are inclined to allocate more tasks to themselves because they prefer to have more control in their own hands.

Based on the above ideas, I develop a research model that identifies the boundary conditions of IntelSys and human decision makers in terms of decision making: human expertise and task uncertainty. I further propose hypotheses to explain how human expertise and task uncertainly collectively moderate the effects of different DMAs on team performance.

1.3 Research Design and Results

To test the proposed research model and associated hypotheses, I develop a multitask software that implements the three DMAs and different (e.g., low and high) task uncertainty levels. Multitask software is often used in HAI research as a simulation of real HAI scenarios such as a cockpit in flight (Miller and Parasuraman 2007; Visser et al. 2010). In the multitask software of this thesis, there is a shooting task, which is the main task operated by either a human or automation; there is also a gauge-monitoring task, which is used to measure the human player's workload. Following prior HAI studies (Kaber 1996; Bailey et al. 2006; Kaber et al. 2006; Kidwell et al. 2012; Cabrall et al. 2018), a lab experiment is conducted to test the research model and hypotheses.

The findings of this study support the notion that task uncertainty and expertise are the key boundary conditions for understanding the effects of the three DMAs on team performance. The results of further analyses suggest that the automation ratio, which indicates the amount of automation used during the entire task execution (ranging from 0% to 100%), positively impacts team performance. The results further reveal that, interestingly, novices, like experts, also sought more control (allocated most of the task to themselves) rather than giving up control (e.g., allocating most of the task to automation) in high task uncertainty situations, compared with low task uncertainty situations. I also discover that for human decision makers, the automation ratio mediates the impact of the task uncertainty and expertise on team performance.

1.4 Contributions

The research findings of this thesis make significant theoretical contributions and bear important practical implications. In regard to research, this dissertation contributes to the HAI literature by providing theoretical explanations for the inconsistent findings in prior DMA research, as well as identifying the boundary conditions for DMAs. This study advances this body of literature by approaching HAI and conceiving the human and automation that collectively perform the task as a "team"; this team-based view empowers me to draw on the rich body of knowledge accumulated in the area of team-based decision making, behaviors, and performance. Drawing on insights from team-based decision making and the personal control model, I develop a theoretical model to explain how task uncertainty and human expertise affect the impact of DMAs on team performance.

The findings of this study also offer insights that are instrumental for practitioners in various sectors, including aviation, autonomous vehicles, manufacturing, and service, among others. The results suggest that managers who intend to provide human-centered automation must go beyond the traditional mentality that treats humans and automation separately. Rather, practitioners should design HCA from the perspective of the team and consider how to best coordinate efforts between humans and automation. When doing so, it is critical to consider both the level of human expertise and the level of task uncertainty, as these are the two boundary conditions on which the effects of DMA on team performance vary.

The remainder of this thesis proceeds as follows. Chapter 2 "Theoretical Background" provides a detailed review involving the literature of human-automation interaction, team-based decision making, and the personal control model. Chapter 3 "Research Model and Hypotheses" elaborates on the research model and hypotheses. Chapter 4 "Experiment Design" provides and delineates the experiment platform, experiment setting, pilot study, and large-scale experiment. Chapter 5 "Results" reports and discusses the results of the main study and post-hoc analyses. Chapter 6 "Discussion" summarizes the research results, discusses the theoretical contributions and practical implications, and concludes with limitations and future research directions.

2 Theoretical Background

In this section, I review and summarize the studies in human-automation interaction and team-based decision-making areas. For the studies of human-automation interaction, I first review the problems associated with traditional technology-centered automation (TCA). Then I introduce the concept of human-centered automation (HCA). The HCA literature has identified humans and intelligent systems (IntelSys) as different decision-making authorities and has distinguished three decision-making approaches (DMAs): IntelSys-Decides, Human-Decides, and IntelSys-Advises. In this section, I review this stream of research and then compare the three DMAs in terms of their advantages and disadvantages. I then identify the underlying reasons for the inconsistent findings concerning the task performance of DMAs. Moreover, by drawing on the knowledge of team-based decision-making research, I identify the boundary conditions for DMAs. For the literature of team-based decision making, I identify four factors that may affect team performance, including the team member's expertise, recognition of expertise, team coordination, and task characteristics, which will be further discussed in the following section. The literature review and theory development are based on an important assumption, that is, the automation outperforms most humans in task execution (Kaber 1996). This is why humans want to use automation in different tasks.

2.1 Human-automation Interaction

2.1.1 Technology-centered automation

As an initial human-automation interaction method, technology-centered automation (TCA) aims to automate tasks to machines as much as possible because machines are more reliable, more highly efficient, and more productive than humans (Kaber 1996). With this approach, humans only focus on tasks that cannot be automated. However, TCA causes human out-of-the-loop (OOTL) problems, both in the short and long term. Specifically, in the short term, TCA decreases the operator's *situation awareness* due to poor monitoring, complacency, and vigilance (Endsley 1995). In the long term, the operator faces skill decay problems (Endsley 1995) caused by underutilization of skills and fragmented tasks that impede the development of the skill set (Kaber 1996). Below I elaborate on the problems of a loss of situation awareness and skill decay.

2.1.1.1 Loss of situation awareness

The short-term problem of TCA involves the loss of situation awareness (SA). Endsley and Kiris (1995, p. 382) define SA as "*the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*." There are three levels of SA (Endsley 1995a; Endsley 1995b). In level 1, the operator perceives available systematic and environmental information from the current context. In level 2, the operator integrates and comprehends the perceived information. In level 3, the operator uses the processed information to predict the future. According to Kaber (1996), in HAI, there are five reasons that lead to the loss of SA.

(1) Poor monitoring. In highly automated tasks, the operator often plays a supervisory role to monitor the state of automation. Monitoring requires the operator to keep an eye on automation and to detect errors when automation fails. In the case of poor monitoring, the operator misses the state information of the automation and, therefore develops an inaccurate perception of the current situation. As such, poor monitoring leads to decreased SA in level 1 (Endsley 1995).

(2) Complacency. Complacency means that operators place too much trust and confidence in automation (Kaber 1996). High reliability on automation often decreases the operator's suspicions about the automation. As the operators become more confident with the systems, poor monitoring also happens (Wiener 1988).

(3) Decreased vigilance. Vigilance indicates that operators maintain a state in terms of being watchful of the automation. The vigilance problem is defined as failing to respond to automation malfunction (Parasuraman et al. 1993; Warm et al. 2008). When operators experience fatigue, high workloads, or high levels of stress, they often miss the malfunction of automation (Wiener 1988). The human operator's low level of involvement in the task is more likely to result in the vigilance problem than high involvement. Empirical findings suggest that human operators could better detect the state of the system when exercising manual control (i.e., high involvement) than when playing only the supervisor role (i.e., low involvement) (Wickens and Kessel 1981).

(4) Active versus passive information processing. The style of information processing (i.e., passive or active) influences the extent to which human operators comprehend and analyze the perceived information, corresponding to level 2 of SA (Parasuraman et al. 1993; Kaber 1996). Active information processing usually indicates a high level of involvement in which operators invest more cognitive resources in processing information. The passive process is the opposite: operators allocate fewer cognitive resources to process information. Understandably, active information processing leads to a high level of SA, whereas passive information processing leads to a low level of SA.

(5) Inappropriate feedback. Another SA problem caused by OOTL is inappropriate feedback from automation, which may cause operators to miss certain critical cues from their current tasks (Kaber 1997). When in OOTL, operators often play the role of supervisor to monitor the state of automation. As supervisors, operators often lack proprioceptive information such as vibration and smell, which are informal information but important for operators to make a diagnosis (Kessel and Wickens 1982; Moray 1986). Another problem caused by inappropriate feedback is the amount of information from automation. On the one hand, information overload may distract operators from making the right decisions and taking the right actions (Endsley and Kiris 1995). On the other hand, having too much processed and integrated information, which is usually the preference of system designers, may mislead operators if one of the many information inputs is incorrect (Billings 1991; Wickens 1994).

The results of losing SA are critical and sometimes even vital (Endsely 2017). Due to the lack of SA, operators often experience “out-of-the-loop unfamiliarity” problems in which they do not know when the automation malfunctions (level 1 of SA), why it malfunctions (level 2 of SA), and what they should do once the malfunction occurs (level 3 of SA). Yet, the operators cannot enter the loop appropriately due to poor SA (Wickens 1994). For example, when the pilots of an Airbus A330 failed to discover that the autothrottle had malfunctioned, the plane crashed into a river after taking off from New York's LaGuardia Airport (National Transportation Safety Board 1990). Another air crash that happened in the ocean near Brazil suggested that even if the pilots had correctly realized that the cockpit automation malfunctioned, they could not save the flight because they were not able to diagnose why it happened (BEA 2012). Even if the operators have enough SA in levels 1 and 2, when automation malfunctions, it is difficult to seamlessly pass the current unfinished task to humans, who may need time to sufficiently understand the current system state and then act appropriately. Therefore, there is a delay in the task transition between the time when the malfunction takes place and the point at which human operators can recover the failure (Kaber and Endsley 2004). The delay, even if only several seconds, may result in fatal accidents involving tasks such as driving cars and flying aircraft.

2.1.1.2 Skill decay

Compared with the loss of SA, which is a short-term consequence of OOTL, skill decay is an undesirable long-term effect of OOTL. The long time of OOTL leads operators to lose proficiency or skills due to the underutilization of manual skills and a lack of practice. Skill decay is particularly problematic in life-critical tasks such as cockpit tasks (Wiener 1988; Kaber 1996). For example, Shiff (1983) found that operators who serve as supervisory controllers and who work on monitoring tasks (i.e., low involvement), relative to those who were in the manual mode (high involvement), performed more slowly and were less efficient.

2.1.2 Human-centered automation (HCA)

Human-centered automation (HCA) was proposed to overcome problems during TCA while maintaining the benefits of automation. The spirit of HCA is to keep humans in the loop. In this approach, the operators are actively involved in the task, even if it can be accomplished by automation. Nonetheless, the operator's workload is still reduced during HCA, as compared to the case in which the entire task is performed manually by the operator alone (Endsley and Kiris 1995). The literature suggests that there are three widely discussed methods: (1) the level of automation (LOA), (2) the IntelSys-Decides approach (adaptive automation), and (3) the Human-Decides approach (adaptable automation).

2.1.2.1 Level of automation (LOA)

LOA is a way to allocate the task assignment between the human and automation so that both are involved in the tasks (Kaber 1996; Endsley and Kaber 1999). LOA uses different numbers to indicate the extent to which a task is automated (Endsley and Kaber 1999; Kaber and Endsley 2004). For example, as one of the most widely cited LOA taxonomies (Kaber and Endsely 2004; Vagia et al. 2016), Sheridan and Verplank (1978) proposed a 10-level automation framework in the context of undersea teleoperator control (Table 2.1). Other researchers have also proposed different LOA models, such as Endsley (1987)'s five-level LOA model, Ntuen and Park (1988)'s five-level LOA model, and Endsley and Kaber (1999)'s 10-level LOA model. These papers tried to develop LOA models based on a specific task context. For example, Endsley (1987) and Ntuen and Park (1988) developed LOA models in the context of the teleoperated system.

Table 2.1 The LOA framework by Sheridan and Verplank (1978) (pp. 8-17)

Level	Task allocation
1	Human does the whole job up to the point of turning it over to the computer to implement
2	Computer helps by determining the options
3	Computer helps determine the options and suggests one that the human needs not follow
4	Computer selects action, and the human may or may not do it.
5	Computer selects action and implements it if the human approves
6	Computer selects action and informs the human in plenty of time to stop it
7	Computer does whole job and necessarily tells the human what it did
8	Computer does whole job and tells the human what if did only if the human explicitly asks
9	Computer does whole job and decides what the human should be told
10	Computer does whole job if it decides it should be done, and if so, tells the human, if it decides that the human should be told

Parasuraman et al. (2000) developed a four-stage LOA model that can be used for most tasks. In this model, the task processing procedures can be divided into four stages according to the human information processing mechanism: 1) information acquisition, 2) information analysis, 3) decision and action selection, and 4) action implementation. Within each stage, automation can be applied across a continuum from low to high. Almost all LOA models mentioned above can be transferred to this four-stage model by Parasuraman et al. (2000). The acquisition automation stage mainly focuses on information input. The lowest LOA can involve scanning data by sensors, and moderate LOA can involve organizing input data. The information analysis stage refers to analyzing the input information and then making the prediction. In this stage, a low level of LOA concerns making a prediction based on input information; a moderate level of LOA could concern integrating different inputs, while a high level of LOA plays the role of "information manager." In the decision automation stage, a low level of LOA involves recommending courses of action to humans, while a high level of LOA executes those courses of action. The action implementation stage concerns implementing the action based on the decision of the previous stage. Recently, Vagia et al. (2016) surveyed existing LOA taxonomies, compared them comprehensively, clarified the range of application, and proposed an integrated version of LOA. Vagia et al. (2016) arrived at an eight-level LOA taxonomy ranging from full manual to full automation.

Advantages of LOA

The most significant advantage of LOA is that it keeps operators in the loop, which therefore alleviates some problems of technology-centered automation (Miller and Parasuraman 2007). In addition, the operator's performance could be significantly improved by LOA (Miller and Parasuraman 2007). Empirical findings suggest that an intermediate level of automation often leads to better SA and fewer OOTL problems than other levels of LOA (Endsley and Kaber 1999). This is because low LOA often increases the operator's workload and cost, while high LOA often leads to skill degradation problems (Miller and Parasuraman 2007).

Disadvantages of LOA

One of the most critical problems of LOA is that the task allocation between the operators and automation is static. However, the states of the machine, human, environment, and task requirements keep changing, which requires continuous adjustment of task allocation between the operators and automation. The apparent disadvantage of LOA is that it cannot adapt to the dynamic context (Endsley and Kaber 1999). Another problem of LOA pertains to the task transition between the operator and automation when a malfunction occurs. After a period of running, automation may encounter failures and requires an immediate remedy that cannot be achieved by static task allocation; hence, a human-based intervention is necessary. In this case, the operators not only need to do their share of the work, but also need to monitor the automation and react to the malfunction. Under LOA, operators may not be able to undertake the automation's share of the work when the automation malfunctions (Kaber 1997; Kaber et al. 1999; Kaber and Endsley 2004).

LOA can sometimes result in low user acceptance, especially for highly skilled operators (Miller and Parasuraman 2007; Endsley 2017). As automation becomes more sophisticated and takes a larger share of the work from the human operators, these operators may feel threatened by the loss of authority (Wickens 1994). Moreover, for operators with high-level skills, when encountering automation that offers no explanation (e.g., feedback and transparent information), they may become very resistant in terms of accepting the technology (Wickens 1994; Miller and Parasuraman 2007).

Another issue of LOA is that some problems that happen in TCA can also happen in LOA, albeit to a lesser extent (Miller and Parasuraman 2007). In particular, complacency and vigilance problems may occur, given that operators put too much trust in well-designed and reliable automation (Miller and Parasuraman 2007). When operators become bored, poor monitoring will take place (Wickens 1994). Skill degradation is also a risk of LOA. For example, pilots often face problems involving the loss of psychomotor skills, such as aircraft altitude control, and decision-making skills in emergency circumstances (Billings 1997; Kaber et al. 1999).

2.1.2.2 IntelSys-Decides approach

The concept of IntelSys-Decides originates from Artificial Intelligence (AI) research in the 1970s as the form of an adaptive aid applied in cockpit automation (Rouse 1988; Scerbo 1996). It was conceptualized as a varying level of automation based on the task characteristics and operator state (Rouse 1988). In this thesis, I focus on two levels of automation: total manual and total automation. In this vein, IntelSys-Decides will allocate the task between the manual mode (total manual) and the automation mode (total automation). For example, if an operator's workload is high, IntelSys will allocate the task to automation.

On the contrary, when the operator's workload is low, in order to keep the human in the loop, IntelSys will ask the operator to perform the task. Although there is not a consistent definition, IntelSys-Decides is widely accepted as a method to enhance task performance while reducing many problems associated with OOTL (Hilburn et al. 1997; Li 2013). The key difference between LOA and IntelSys-Decides is that IntelSys-Decides is a dynamic task allocation approach, while LOA is a static one (Kaber and Endsley 2004).

Different from LOA, the core assumption of IntelSys-Decides is that when the task environment and the operator's mental state change, the task allocation mode should also be adjusted. In this vein, a more flexible method such as IntelSys-Decides is needed to dynamically allocate tasks between automation and humans based on the varying environments, task demands, and human workload levels (Scerbo 1996; Kaber and Perry 2006; Steinhauser and Hancock 2009). The goal of IntelSys-Decides is to reduce OOTL performance problems such as poor monitoring, skill decay, vigilance, and complacency. The level or type of automation in IntelSys-Decides can be modified in real-time in order to achieve better human-system collaboration (Scerbo 1996; Kaber et al. 2005; Sheridan and Parasuraman 2006; Scerbo 2008). Under IntelSys-Decides, four strategies are commonly used to trigger dynamic task allocation:

(1) Critical events. IntelSys-Decides can be triggered by the occurrence of specific events that could critically impact the system goals (Parasuraman et al. 1992; Kaber 1996; Kaber and Endsely 2004). If a critical event happens, the system switches the task from the current mode (e.g., automation) to the other mode (e.g., manual), and vice versa. For example, the Automatic Ground Collision Avoidance System will automatically pull up the airplane when it detects that the airplane is too close to the ground. However, this strategy does not consider the operator's workload and/or performance. The potential problem is that the task-switching pattern may not fit with the operator's workload or requirements for optimal task performance (Parasuraman et al. 1992).

(2) Performance measurement. IntelSys-Decides can also be triggered by the difference between the operator's current and past performance (Parasuraman 1993; Sauer et al. 2012). The key point of this approach is to choose an appropriate performance model that assesses the operator's performance (Parasuraman 1992). Parasuraman (1992) lists three performance models of IntelSys-Decides in the aviation field. In a lab experiment, Parasuraman et al. (1996) employed a straightforward way to build the performance model, in which IntelSys-Decides is triggered when individuals' performance is below a specific criterion.

(3) Psychophysiological assessment. The core idea of psychophysiological assessment is to keep monitoring the operators' workload in real-time. If the workload is too high, the IntelSys will switch to the automation mode; if the workload is too low, the IntelSys will switch to the manual mode in order to keep the human in the loop. There are generally two ways to measure the workload in real-time. The first is the physiological measures such as heart rates and electroencephalograms (EEG) (Pope et al. 1994; Byrne and Parasuraman 1996). Both indices can reflect the operators' workload level. The second is the secondary-task method introduced by Kaber and Riley (1999), in which the subject works on two tasks simultaneously: the primary task and the secondary task (Kaber and Riley 1999; Clamann et al. 2002; Kaber et al. 2006). The secondary-task method builds on Wickens' multiple-resource theory (MRT; Parasuraman et al. 1992). MRT suggests that the primary and secondary tasks compete for the same type of cognitive resources from the human. When focusing on the primary task, if the human's workload is too high, then only few cognitive resources are available to be allocated to the secondary task; on the contrary, if the human's workload is too low, then plenty of cognitive resources are available to be allocated to the secondary task. Therefore, the performance of the secondary task is inversely related to the operator's workload on the primary task (Kaber and Riley 1999). In this research, we use the secondary task method because (1) compared with physiological measures, the secondary-task method is unintrusive to the subject; (2) the EEG method often needs many sensors, which are cumbersome to wear and cannot provide real-time feedback to the adaptive system (Feigh et al. 2012); and (3) the signals of EEG sensors are susceptible to noise and cannot be read accurately (Feigh et al. 2012).

(4) Behavior modeling strategy. IntelSys-Decides can also be triggered based on a behavioral model (Kaber 1996; Kaber and Endsley 2004) that measures the operator's activities, awareness, intentions, resources, and performance (Rouse et al. 1986). These measurements are then used to estimate the operator's current state and predict the future state (Rouse et al. 1986). The state information is then used to trigger IntelSys-Decides in order to achieve a predetermined model of overall system functioning.

Among these four methods, the psychophysiological assessment and performance measurement methods are widely used in experimental settings to trigger IntelSys-Decides (e.g., Kaber and Riley 1999; Parasuraman et al. 1996). Following prior IntelSys-Decides studies conducted in the lab context, we use the psychophysiological assessment method to trigger IntelSys-Decides in this research.

Advantages of IntelSys-Decides approach

IntelSys-Decides can enhance human-system performance and can reduce many problems associated with the OOTL method (Hilburn et al. 1997; Li 2013; Table 2.2). First, by keeping humans in the loop, humans can be more engaged and can obtain real-time feedback during work. Therefore, IntelSys-Decides enhances operators' situation awareness and monitoring performance. Second, when the task is not demanding, the dynamic allocation between the manual and automation modes also eliminates the operator's tendency to be complacent (Parasuraman and Wickens 2008). Third, the dynamic task-allocation mechanism ensures that IntelSys-Decides maintains the operator's manual skills and keeps the human workload within a reasonable range (Parasuraman et al. 1992; Hancock and Scallen 1996; Steinhauser and Hancock 2009). Fourth, IntelSys-Decides is a method to retain the advantages of both humans and automation and make them complementary to each other (Parasuraman et al. 1992; Scerbo 2008). On the one hand, automation can protect operators from fatigue or a high workload. On the other hand, human involvement also reduces automation's probability of failure; even if the automation malfunctions, the operators are available and capable of intervening (Kaber 1996).

The above discussion suggests that IntelSys-Decides can overcome the limitations of LOA and adjust humans' workload by matching the state of automation with that of the operator. However, without considering the operators' psychological state, IntelSys-Decides inevitably has potential limitations.

Disadvantages of IntelSys-Decides approach

The IntelSys-Decides approach also has numerous drawbacks (Table 2.2). To begin with, because the share of humans' work is allocated by the IntelSys, humans play a passive role rather than an active one. Understandably, humans usually show a low level of acceptance toward this approach (Miller et al. 2005). In particular, operators who are confident about and who value their task expertise are reluctant to accept task allocation decisions from IntelSys-Decides (Parasuraman et al. 2007).

Another problem involves system unpredictability, which means that human operators cannot predict the behavior of IntelSys (Billings and Woods 1994; Miller and Parasuraman 2007; Parasuraman et al. 2007). There are two reasons that lead to system unpredictability. The first concerns operators' loss of control (Miller and Parasuraman 2007). Under IntelSys-Decides, IntelSys can switch from manual to automation, or vice versa, but humans have no control over such decisions; hence, for humans, it is difficult to predict when the system will switch, thereby compromising their situation awareness and performance (Parasuraman et al. 2007). The second reason is that the mechanisms (e.g., workload) used to trigger IntelSys-Decides keep changing (Miller and Parasuraman 2007). For example, the human's workload is a trigger that keeps changing unpredictably. This unpredictability is often associated with a higher workload and may increase the human's workload when rest is needed (Table 2.2; Parasuraman et al. 2007). Next, when making switching decisions in IntelSys-Decides, IntelSys does not consider individual characteristics, such as humans' mental state, emotion, personality, or physical conditions that could affect performance outcomes (Steinhauser and Hancock 2009; Dekker and Woods 2002). According to Steinhauser and Hancock (2009), the fit between individual characteristics and IntelSys is a necessary condition for optimal human-automation interaction (HAI). However, in IntelSys-Decides, humans play a passive role rather than an active one; they must accept the tasks allocated by the IntelSys-Decides system, regardless of their preference, mental states, health conditions, or habits. In addition, operators may have different levels of task expertise, which usually involves tacit knowledge and is unique to themselves. Unfortunately, IntelSys-Decides also does not consider humans' tacit knowledge, nor does it examine how humans' experience impacts overall performance (Table 2.2).

Table 2.2 Approaches to human-centered automation

	Level of Automation	**IntelSys-Decides[4]**	**Human-Decides[5]**	**IntelSys-Advises[6]**
What is it	Uses different levels of automation to indicate different degrees by which a task is automated.	Varying level of automation based on the task characteristics and operator state.	The same principle as IntelSys-Decides, but it is the operator, rather than the system, that has the authority to allocate tasks between the human and automation.	Combines IntelSys-Decides and Human-Decides. IntelSys prompts suggestions to the operator, who can decide to either accept or reject them.
Authority	IntelSys	IntelSys	Human	Human
Advantage	Keeps human operators in the loop. Improves the human operator's performance; An intermediate level of automation is the best.	Enhances human-system performance; Enhances human operators' situation awareness and monitoring performance; Maintains the human operator's manual skills.	It inherits most advantages of IntelSys-Decides, except for the increased workload; High user acceptance; Considers individual differences; Reduces unpredictability; Efficient resource management; Improves system performance.	Same advantages as Human-Decides, but with a lower workload.
Disadvantage	Static task allocation cannot adapt to the changing context; A lack of user acceptance; Low situation awareness; Skill degradation.	Low user acceptance; Increased system unpredictability; No consideration of individual characteristics, such as mental state, personality, emotion, health conditions, and habits.	Increases cognitive demand and workload; Increases interface management tasks.	Increases interface management tasks; Information overload and distraction.

[4] Also termed as Adaptive Automation in prior literature (Li 2013)
[5] Also termed as Adaptable Automation in prior literature (Li 2013)
[6] Also termed as Hybrid in prior literature (Li 2013)

2.1.2.3 Human-Decides approach

The critical difference between IntelSys-Decides and Human-Decides approaches is that, under the Human-Decides approach, the operator, rather than the system, has the authority to allocate tasks between humans and automation. Such a difference mitigates not only many issues associated with IntelSys-Decides, but also provides humans with more opportunities to implement their decision-making capabilities. Moreover, the difference between IntelSys-Decides and Human-Decides can be further understood from another perspective: IntelSys-Decides incorporates automation and operators by considering the limiting factors of operators (i.e., workload, performance); meanwhile, by giving human decision-making authority, the Human-Decides approach incorporates automation and operators by considering operators' physical and psychological states, user habits, user acceptance, and so forth.

The playbook method is one of the most famous Human-Decides strategies (Parasuraman et al. 2005). This method was first proposed by Miller et al. (2000). Its logic is borrowed from the idea of a sports team's playbook, in which the coach selects players and gives them specific instructions to achieve the team's goals (Parasuraman et al. 2005). In the playbook approach, human operators and automation share a developed vocabulary in which they can communicate through a hierarchical task model. This model simulates the human-human task delegation process in scenarios akin to a sports team by indicating how the operator (who plays the role of a coach) delegates subtasks to automation (which plays the role of a player). In this model, operators delegate plans and instructions to automation at times when they think it is necessary to do so. The playbook approach is a very effective method for implementing Human-Decides approach. It has been validated empirically as a mission-planning tool when applied in unmanned combat air vehicles (Miller et al. 2004; Parasuraman et al. 2005; Miller and Parasuraman 2007; Kidwell et al. 2012). The results suggest that this delegation approach can improve the performance of the entire system, keep the human operator's workload at a manageable level, and improve the operator's confidence when making task-related decisions. Interestingly, Sauer et al. (2012) found no clear advantages between IntelSys-Decides and Human-Decides, except that operators under Human-Decides show high self-confidence.

Advantages of Human-Decides approach

According to Miller and Parasuraman (2007), for operators, the benefits of Human-Decides include "*improved situation awareness, more accurate automation usage, more balanced mental workload, increased user acceptance, and improved overall performance*" (p. 57). Because operators have the authority to allocate tasks, they play an active role. This active role helps reduce operators' complacency and prompts more attention on the part of operators toward monitoring and task completion (Kidwell et al. 2012), thereby increasing operators' situation awareness (Kaber 1996). In addition, more control leads to improved user acceptance, confidence, and satisfaction (Kidwell et al. 2012). More importantly, operators can make decisions to delegate tasks to automation or themselves based on their workload, mental state, emotional state, working memory capacity, or even personal habits, thereby making the tasks more suitable for themselves (Visser et al. 2010; Chen and Barnes 2012). In addition, operators are more efficient at managing resources and are more effective in terms of optimal human-system performance (Scerbo 2008). They can tailor the automation mode to suit their current and future workloads (Kidwell et al. 2012).

Disadvantages of Human-Decides approach

Despite the many benefits of Human-Decides, there are some limitations as well (Table 2.2). This DMA increases operators' cognitive demands and workload by undertaking the decision-making task, above and beyond performing the main task (Bailey et al. 2006; Miller and Parasuraman 2007; Kidwell et al. 2012; Li 2013).

2.1.2.4 IntelSys-Advises approach

There are some studies that aim to solve the *IntelSys-Decides* or *Human-Decides* dilemma by exploring a third strategy – the IntelSys-Advises approach, which is proposed to integrate the advantages of IntelSys-Decides and Human-Decides (Li 2013). After reviewing both the advantages and disadvantages of the IntelSys-Decides and Human-Decides approaches, it is natural to think that IntelSys-Decides and Human-Decides each has unique strengths and limitations. The IntelSys-Advises approach is thus proposed to integrate the advantages of IntelSys-Decides and Human-Decides while minimizing disadvantages. One can expect that compared with the Human-Decides approach, the IntelSys-Advises approach inherits the advantages of Human-Decides and increases the decision-making workload at an acceptable and manageable level. Under IntelSys-Advises, the human operator has the ultimate authority. There are two IntelSys-Advises methods. First, when the IntelSys of IntelSys-Decides finds it is necessary to make a switch from human to automation or the opposite, it prompts a decision as a suggestion to the operator, who can decide to accept or reject the advice. The second mechanism is that the IntelSys under this DMA can make decisions about when to switch and then act on them; however, the operator has the authority to override or reserve the decision. These two approaches correspond to Olson and Sarter (2000)'s management-by-consent and management-by-exception strategies, which are widely used in the aviation domain. For the management-by-consent strategy, automation can only act after getting permission from the operator. For the management-by-exception strategy, IntelSys works automatically, and the operator has the right to override or reverse the automation.

Based on a simulated modern flight deck, Olson and Sarter (2000) examined the preferences of pilots with rich operational experiences in terms of three different human-automation interaction modes: management-by-consent, management-by-exception, and full automation. The results suggested that pilots have a strong preference for the management-by-consent strategy in which automation can only take effect with explicit permission from the operators (Olson and Sarter 2000). For tasks with low criticality, when the workload and time pressure are very high, pilots switch their preference to management-by-exception. This is understandable because the management-by-exception approach is easy, safe, and effective in regard to relying on IntelSys; at the same time, pilots' override or reverse authority provides a "double-check" warranty for safety and effectiveness.

Theoretically speaking, IntelSys-Advises should be the best approach because it has been designed to combine the advantages of IntelSys-Decides and Human-Decides by integrating the knowledge of human and IntelSys in decision making (Table 2.2; Li 2013; Li et al. 2013; Cabrall et al. 2018; Ruff et al. 2018).

Most empirical studies support this viewpoint. For example, Ruff et al. (2018) found that (a) Human-Decides and IntelSys-Advises have higher task performance than IntelSys-Decides; and (b) Human-Decides has a lower workload level than IntelSys-Advises, which is lower than IntelSys-Decides (Table 2.2). Li (2013) and Li et al. (2013) had similar results, except that Li et al. (2013) indicated that IntelSys-Advises and IntelSys-Decides have a lower level of situation awareness than Human-Decides. One exception is that Kaber and Riley (1999) found that IntelSys-Decides is better than IntelSys-Advises in terms of total task performance and individual performance, both of which were caused by the disadvantages of IntelSys-Advises. In the IntelSys-Advises approach, the human must interact with the task, automation, IntelSys, and the environment. Therefore, an additional interface design and management are needed (Li 2013). In addition to working on the task, the human must also make decisions, which, together with interface management, increase the human's workload. Another problem is that the additional interface design not only leads to information overload, but may also distract the human from making high-quality decisions, especially when the advice of IntelSys conflicts with the operator's decision (Kaber and Riley 1999). When inconsistent with the human's decisions, the advice of IntelSys may lead humans to spend cognitive resources to question and consider whether to automate, and thus distract the human from task processing (Harris et al. 1993; Kaber and Riley 1999).

2.1.3 Summary of literature review

From the literature review, it is easy to see that IntelSys-Decides and Human-Decides are much better than the LOA approach (Table 2.2). The difference between IntelSys-Decides and Human-Decides centers on who has the authority to allocate tasks. The detailed differences are summarized in Table 2.2. From Table 2.2, compared with IntelSys-Decides, Human-Decides reduces problems associated with user acceptance, confidence, unpredictability, and individual differences, but also increases the operator's workload. However, which approach is better, IntelSys-Decides or Human-Decides? From the literature, I find that empirical research on the benefits of IntelSys-Decides are well and widely documented (Kaber and Endsley 2004; Kaber 1996; Kaber and Perry 2006; Cosenzo et al. 2006; Calhoun et al. 2011), while relatively fewer empirical papers involve Human-Decides (Miller et al. 2005; Parasuraman et al. 2005; Squire and Parasuraman 2010). Furthermore, fewer empirical papers compare these two strategies simultaneously (Parasuraman and Wickens 2008; Kidwell et al. 2012). Therefore, based on the existing literature, it is difficult to derive consistent conclusions about which DMA is better.

2.1.3.1 Viewpoints supporting IntelSys-Decides

As mentioned above, the IntelSys-Decides approach improves human situation awareness and monitoring performance (Hilburn et al. 1997; Parasuraman and Wickens 2008; Li 2013; Table 2.2). Empirical evidence supports the idea that in terms of task performance, IntelSys-Decides is better than Human-Decides because the mechanism triggering IntelSys-Decides considers the underlying engagement states (i.e., workload) that can be recognized by IntelSys, but not by humans (Bailey et al. 2006; Steinhauser and Hancock 2009; Sauer et al. 2012). For example, humans may not accurately recognize that their performance is going down, while IntelSys can. IntelSys-Decides, triggered by the workload, "*provides a tighter coupling between an operator's available cognitive resources and the immediate demands of the task*" (Bailey et al. 2006, p. 705). Therefore, the switching decisions of IntelSys-Decides are generally more objective than those of Human-Decides.

In addition, IntelSys-Decides is better than Human-Decides because the automation is more accurate than operators, and the reaction of automation is faster and more precise than the operators' responses, especially in hazardous situations (Scerbo 2008). For example, the Ground Collision-Avoidance System (GCAS) in F-16D fighters is an IntelSys-Decides system that allocates the altitude adjustment task to automation and the pilot. In some extreme situations, when the altitude of the aircraft is too low and dangerous, and the pilot does not know or has no time to respond, the automation will take over to maneuver the plane to a safe altitude and then return control to the pilot (Scerbo 2008). Therefore, for safety considerations, IntelSys-Decides is better than Human-Decides, even if the final decision is made without permission from the operator (Parasuraman et al. 2007). Moreover, Bailey et al. (2006) found that IntelSys-

Decides, relative to Human-Decides, performs better in terms of situation awareness and results in a lower workload.

2.1.3.2 Viewpoints supporting Human-Decides

Some researchers have found evidence that prefers Human-Decides over IntelSys-Decides in terms of SA. For instance, Wickens (1994) found that, when comparing user awareness of changes in the environment and system, the operators of Human-Decides demonstrate higher situation awareness than those of IntelSys-Decides. In IntelSys-Decides, it is the IntelSys that decides when and what to automate, whereas in Human-Decides, it is the human operator who makes such decisions. Understandably, Human-Decides may increase operators' workload because they must make the switch decision.

Still, other scholars advocate Human-Decides because this approach allows operators to compensate for suboptimal working conditions by providing them with decision-making authority. For example, by switching between manual and automation modes according to their physical or mental needs, operators can counteract suboptimal working conditions such as task overload/underload and environmental stressors (Sauer and Chavaillaz 2017). Dekker and Woods (2002) suggested that LOA is a quantitative method because it uses numbers to allocate tasks between humans and automation (e.g., based on Sheridan and Verplank (1978)'s LOA framework, level 2 means that "*the computer helps by determining the options,*" and level 5 means "*The computer selects action and implements it if the human approves*"). According to Dekker and Woods (2002), the task allocation in HAI is not a quantitative problem because the relative strengths and weaknesses of humans relative to automation are dynamically changed. For example, overreliance on automation may lead to skill decay and complacency problems, which then weaken humans' advantages with regard to the task. Instead, the task-allocation problem in HAI is a qualitative problem because it forces operators to exert their capability by adapting their skills and routines, which cannot be quantified. Therefore, it is necessary to advocate for the Human-Decides approach so that tasks can be allocated in a non-quantitative manner. Other researchers argue that Human-Decides can keep its advantages while avoiding the disadvantages of IntelSys-decides approach such as a lack of user acceptance, a lack of confidence, system unpredictability, and no consideration of individual differences (Miller and Parasuraman 2007). Some argue that Human-Decides is better because humans have the ultimate responsibility for the entire system, and humans are more efficient at resource management (Billings and Woods 1994; Malin and Schreckenghost 1992). However, little or no research has been carried out to examine why Human-Decides is better than IntelSys-Decides or under what conditions Human-Decides outperforms IntelSys-Decides.

2.1.3.4 Review summary and research gap

Synthesizing the above, Table 2.3 points to several noteworthy knowledge gaps and opportunities for research. In particular, while some studies have compared IntelSys-Decides versus Human-Decides, or IntelSys-Decides versus IntelSys-Advises, only recently have scholars begun to compare all three DMAs (Kidwell et al. 2012; Li 2013; Ruff et al. 2018); yet the results of these studies reveal inconsistent findings on which DMA is the best. We attribute these mixed findings to two possible reasons. First, most of these studies lack a theoretically grounded overarching framework, and thus fail to systematically explain the antecedents, contingencies, and consequences of DMA. Second, it is surprising to observe that almost all prior DMA studies have approached this phenomenon by focusing on individual human operators without considering the human operator and the automation function collectively as a team for attaining a common objective (e.g., best team performance). Third, while management scholars have long submitted that individual characteristics (Visser et al. 2010; Chen and Barnes 2012) and task attributes (Lewis and Herndon 2011; Baumann and Bonner 2017) can affect team performance, none of these DMA studies considers how human or task properties may affect the consequences of DMA. Given the inconsistent findings about DMAs and the knowledge gaps in the existing literature, it is possible that there are boundary conditions for DMAs in team performance. To identify these possible boundary conditions, I refer to the team management literature and draw on the team-based decision-making perspective to investigate the phenomenon.

Table 2.3 Literature summary

Study	Perspective	Individual Traits	Task Characteristics	DMAs Compared	Findings
Hilburn et al. (1993)	Individual	None	Multi-Attribute Task Battery	IntelSys-Decides vs. Human-Decides	The overall task performance of Human-Decides is better than that of IntelSys-Decides. The number of switches is negatively related to performance. IntelSys-Decides made more switches than Human-Decides.
Bailey et al. (2006)	Individual	Complacency	Multi-Attribute Task Battery	IntelSys-Decides vs. Human-Decides	IntelSys-Decides has better situation awareness than Human-Decides. The workload of IntelSys-Decides is lower than that of Human-Decides.
Sauer et al. (2012)	Individual	None	A computer-based simulation environment called the Automated Cabin Air Management System	IntelSys-Decides vs. Human-Decides	No significant difference between IntelSys-Decides and Human-Decides in overall task performance. Participants in Human-Decides have higher levels of self-confidence than in IntelSys-Decides.
Kidwell et al. (2012)	Individual	None	Simulation software of multiple vehicle supervisory control testbeds	IntelSys-Decides vs. Human-Decides	No significant difference between IntelSys-Decides and Human-Decides in overall task performance. Human-Decides increases participants' confidence in their decision-making ability more than IntelSys-Decides. Participants prefer Human-Decides over IntelSys-Decides. Human-Decides has better change detection (situation awareness) than IntelSys-Decides. Human-Decides has a higher workload than IntelSys-Decides.
Kaber and Riley (1999)	Individual	None	Self-developed Multitask software, which simulated air traffic control	IntelSys-Decides vs. IntelSys-Advises	IntelSys-Decides is better than IntelSys-Advises in total performance. Individual performance in IntelSys-Decides is better than that in IntelSys-Advises.
Cabrall et al. (2018)	Individual	None	Simulated driving task	IntelSys-Decides vs. IntelSys-Advises	IntelSys-Advises (Backup automation) has higher performance than IntelSys-Decides (foreground automation). IntelSys-Advises has a lower workload than IntelSys-Decides. Participants prefer IntelSys-Advises over IntelSys-Decides.
Li (2013)	Individual	None	Teleoperation task	IntelSys-Decides vs. Human-Decides vs. IntelSys-Advises	Overall performance: IntelSys-Advises >IntelSys-Decides, Human-Decides For mental workload: IntelSys-Advises=Human-Decides > IntelSys-Decides For physical workload: IntelSys-Advises>Human-Decides>IntelSys-Decides
Li et al. (2013)	Individual	None	Simulated robotic arm control tasks	IntelSys-Decides vs. Human-Decides vs. IntelSys-Advises	Performance of Human-Decides and IntelSys-Advises is better than IntelSys-Decides. IntelSys-Advises and Human-Decides have a lower level of workload than IntelSys-Decides. IntelSys-Advises and IntelSys-Decides have a lower level of situation awareness than Human-Decides. Participants' preference: Human-Decides > IntelSys-Advises > IntelSys-Decides.
Ruff et al. (2018)	Individual	None	Surveillance tasks performed in a multi-task simulation environment	IntelSys-Decides vs. Human-Decides vs. IntelSys-Advises	Task accuracy was better in Human-Decides and IntelSys-Advises than in IntelSys-Decides. Human-Decides has the lowest workload, then IntelSys-Advises, then IntelSys-Decides. Participants preferred Human-Decides and IntelSys-Advises over IntelSys-Decides.

2.2 Team-based decision-making

In this section, I review and analyze the literature in team-based decision making. I first propose that it is necessary to study the human-automation interaction from the team perspective. Then I clarify the meaning of team performance and illustrate four factors that impact team performance, including team members' expertise, expertise recognition, team coordination, and task characteristics. Due to many terms and definitions in this section, I list all terms and definitions in Table 2.4.

Table 2.4 Terms and definitions

Terms	Definitions	Reference
Team	Collectives that exist to perform organizationally relevant tasks, share one or more common goals, interact socially, exhibit task interdependencies, maintain and manage boundaries, and which are embedded in an organizational context that sets boundaries, constrains the team and influences exchanges with other units in the broader entity.	Kozlowski and Bell (2003, p. 334)
Human-automation team	The dynamic, interdependent coupling between one or more human operators and one or more automated systems requiring collaboration and coordination to achieve successful task completion.	Langan-Fox et al. (2009, p. 896)
Expertise	The specialized skills and knowledge that an individual brings to the team's task.	Faraj and Sproull (2000, p. 1555)
Team coordination	Team coordination refers to the way team members' expertise is organized in order to complete the team task successfully.	Wittenbaum, Vaughan, and Stasser (1998, p. 177)
Expertise coordination	Team-situated interactions aimed at managing resources and expertise dependencies.	Faraj and Sproull (2000, p. 1555)
Behavioral control	A belief that one has a behavioral response available that can affect the aversiveness of an event	Thompson (1981, p. 90)
Personal control	The individual's beliefs, at a given point in time, in his or her ability to effect a change, in a desired direction, on the environment.	Greenberger and Strasser (1986, p. 164)
Learned helplessness	When individuals learn that their behaviors are independent of the desired outcomes, a state of learned helplessness occurs.	Greenberger and Strasser (1986, p. 165)

2.2.1 Automation: Teammate or not?

In this paper, I investigate human-automation interaction from the team perspective and argue that humans and automation should work as a team to achieve the desired team performance. The reasons that I adopt the human-automation team perspective include the following: (a) the team perspective fits well with the idea of human-centered automation, which emphasizes that both the human and automation should be involved in the same task; and (b) the human and automation share the same task goal.

Following Mathieu et al. (2008), I adopt the definition of team from Kozlowski and Bell (2003, p. 334) as follows:

> "*Collectives who exist to perform organizationally relevant tasks, share one or more common goals, interact socially, exhibit task interdependencies, maintain and manage boundaries, and are embedded in an organizational context that sets boundaries, constrains the team, and influences exchanges with other units in the broader entity*."

Based on the above definition of a team, the first question that needs to be answered is: can automation be a teammate for humans?

There are many studies examined human-automation interaction from the team perspective. Christoffersen and Woods (2002) proposed that humans and automation can be thought of as a team if two conditions are satisfied: observability and directability. Observability means that humans need to know the current status and future activities of the automation. Directability means that humans can re-direct machine activities in necessary situations. Viewing automation as being subordinate to humans, Christoffersen and Woods (2002) emphasize how humans control automation in order to achieve team objectives. Wijngaards et al. (2006) view humans and automation as a team because humans and automation are heterogeneous, but have potentially equal status; that is, the human can give orders to automation, and vice versa. Wijngaards et al. (2006)'s human-automation team is, in fact, an instance of an actor-agent community.

The actor-agent community is a particular organization in which multiple participants (i.e., humans and automation) collaborate to reach a common goal.

Furthermore, Shively et al. (2017) switch viewpoints of the human-automation team from the traditional view of "*automation works for us*" to "*automation works with us,*" emphasizing the communication function of the human-automation team. The key to human-autonomy teaming is to develop an interface that interacts between humans and automation. The interface should be a bidirectional communication channel in order to make sure that humans can set a goal for the team, request that the automation work toward the goal, and check and adjust the status of the automation. At the same time, the human can get feedback from the automation and update his/her behavior. The empirical evidence from Brandt et al. (2017) verified the effectiveness of the human-autonomy teaming framework of Shively et al. (2017).

Researchers have different opinions about whether automation can be a teammate for humans (Langan-Fox et al. 2009). Groom and Nass (2007) compared a human-robot team with a human-human team and concluded that the human-robot team assumption is invalid. The robot cannot gain trust from the human because the robot lacks humanlike mental models and self-awareness. Based on the above, in this research, I study the human-automation interaction from the team perspective for the following reason: automation and humans are highly heterogeneous, but they have a shared team goal. . Following Langan-Fox et al. (2009) and Cuevas et al. (2007), I define the human-automation team as "*the dynamic, interdependent coupling between one or more human operators and one or more automated systems requiring collaboration and coordination to achieve successful task completion*'' (p. 896). The three different decision-making approaches (DMAs) mentioned above (IntelSys-Decides, Human-Decides, IntelSys-Advises) correspond to three distinct human-automation teams. In the team with the IntelSys-Decides decision-making DMA, IntelSys is the decision maker and is responsible for coordinating the automation and the human according to the human's workload. In the team with the Human-Decides DMA, the human makes task-allocation decisions based on his or her workload and other individual-level factors (i.e., habits, preferences, moods) and situation awareness (Miller and Parasuraman 2007). In the team with the IntelSys-Advises DMA, the human is the decision maker with the advice of IntelSys. The goal of this research is to identify the DMA that yields the best team performance.

Successful human-automation team

Groom and Nass (2007) identified a list of qualities for a successful human-human team, including shared goals, shared mental models, subordinating individual needs to those of the team, viewing interdependence as positive, understanding and fulfilling each member's role, and trust. However, for human-automation team, it is challenging to build mutual trust between humans and automation because humans would like to trust automation (Lee and See 2004), but the automation does not have the ability to trust humans. As previously mentioned, studies such as Christoffersen and Woods (2002), Wijngaards et al. (2006), and Shively et al. (2017) did not ask for a successful team to be as perfect as Groom and Nass (2007). Instead, they adopted a more pragmatic view to study the successful human-automation team, given that automation lacks the capability to trust a human teammate. The following two arguments support this point-of-view.

First, Christoffersen and Woods (2002) suggested that in a cooperative relationship between humans and automation, automation plays the role of a subordinate for the human. The human knows how to control the automation and coordinate the human-automation team. In other words, to be a successful human-automation team, the underlying assumption by Christoffersen and Woods (2002) is a unidirectional understanding and trust from the human to automation, rather than a bidirectional understanding and trust between the human and automation. Therefore, following the logic and evidence of Christoffersen and Woods (2002), the human and automation can be a successful team even if the automation is not humanlike and cannot trust the human teammate.

Second, evidence from the existing literature suggests that a mutual understanding and trust between the human and automation are not absolutely necessary conditions for team performance. For a team that only consists of a human and automation, attaining the best team performance is an operational optimization problem, and the answer would be optimal task allocation (Bottger et al. 1988; Faraj and Sproull 2000; Bonner et al. 2002; Bonner 2004). For Human-Decides approach, optimal task allocation depends on the human, who is the decision maker. Furthermore, the human's appropriate understanding and trust in the automation are necessary conditions for optimal task allocation in the human-automation team. In contrast, the automation's understanding and trust in the human is not necessary when the automation is not the decision maker. In this context, the human needs to make high-quality task allocation decisions in order to attain optimal task allocation based on his or her understanding of the automation. Therefore, it is possible that the human-automation team can achieve optimal team performance with a unidirectional understanding and unidirectional trust from the human to automation.

By carefully examining the difference between a successful human-human team and a successful human-automation team, it is reasonable to identify factors that impact the team performance of human-automation team by drawing applicable knowledge from studies of human-human team. In the next section, I will introduce the factors that impact the successful human-automation team.

2.2.2 Team performance

When working as a team, both automation and human have the same goal - the highest team performance, which is also the core dependent variable in this research. Mathieu et al. (2008) categorized team performance as organizational-level team performance, team performance behaviors and outcomes, and role-based team performance.

- Organizational-level team performance refers to the organizational outcomes from the top management teams, such as the firm's profitability and financial ratios.
- Team performance behaviors and performance outcomes are two different but interrelated concepts. *Performance behaviors* stand for actions or behaviors relevant to desired outcomes, while *performance outcomes* refer to actual consequences of the performance behaviors (Beal et al. 2003; Mathieu et al. 2008).

- Role-based team performance emphasizes the individual member's level of expertise when working on a specific task.

In this study, the team refers to the human-automation team rather than the top management team or project team in the regular management or IS literature. There are only two members in the human-automation team: one human and one automation. As such, the unit of analysis in this thesis is the "human and automation team" rather than the human or automation alone. We only consider the team performance of the operational task instead of the individual member's performance. Therefore, the team performance here refers to the performance outcomes (rather than performance behaviors).

2.2.3 Factors impacting team performance

According to the literature on team-based decision making, there are four factors that could affect team performance: (1) the expertise of team members, (2) recognition of team members' expertise, (3) team coordination, and (4) task characteristics. Below I review each of these four factors.

(1) Expertise

Expertise is defined as "*the specialized skills and knowledge that an individual brings to the team's task*" (Faraj and Sproull 2000, p. 1555). A member's expertise is widely recognized as an important factor that impacts team performance (Littlepage et al. 1995; Littlepage et al. 1997; Littlepage and Mueller 1997; Faraj and Sproull 2000). It is reasonable to infer that the higher level of a team member's expertise, the higher the team performance will be. However, many answers remained unknown in this area, such as how to recognize an individual member's expertise and how to utilize such expertise in team research. These questions demand a greater scholarly effort, particularly in the context of human-automation interaction (Littlepage et al. 1997).

(2) Expertise Recognition

Expertise recognition refers to the extent to which the team can accurately recognize each member's level of expertise. Prior research has suggested that expertise recognition significantly impacts expertise coordination (Littlepage et al. 1995). By accurately recognizing team members' expertise, the team can design an optimal team coordination strategy (Littlepage and Mueller 1997). Both the individual member's task experience and group experience positively impact expertise recognition in the team (Littlepage et al. 1997). Task experience refers to the team member's experience on a similar task. The knowledge and skills learned from the previous similar task can be transferred to the current task (Neal and Northcraft 1986). Therefore, for team members, the more experience they have on a similar task, the higher expertise they have, which then leads to better team performance (Burke and Day 1986).

Next, group experience refers to team members' experience in working with other team members. Group experiences enhance expertise recognition accuracy. A high level of group experience helps the team member gain a better understanding of how other members contribute to the team performance in different situations (Littlepage et al. 1997).

Furthermore, team members' characteristics and behaviors also impact expertise recognition and team coordination. For example, a member who is talkative and male, and who has a high level of participation is more likely to be recognized as an expert and will be allocated more influence or authority in team decision making (Littlepage and Mueller 1997). Moreover, if the team member is confident about his or her task ability and is a dominant force in the working process, he or she is more likely to be seen as an expert by other members. The heterogeneity of team members' expertise is also found to be positively related to the accuracy of expertise recognition. The greater the expertise heterogeneity, the more likely the team members can identify the real experts in the team (Libby et al. 1987).

The impact of a member's characteristics and behaviors on expertise recognition can also be explained by the Status Characteristics Theory (SCT). SCT is used to explain how the team recognizes and utilizes the expertise in the team (Bunderson 2003). SCT posits that an individual's personal characteristics convey information about his/her expertise. The information perceived by the team members on an individual's personal characteristics can be classified into two types of cues: specific status cues and diffuse status cues. Personal characteristics that can provide information about the individual's task competence are specific status cues (e.g., expertise, knowledge, experience, skills). In contrast, personal characteristics such as gender, ethnicity, age, and appearance that do not directly indicate an individual's task competence but may affect his or her competence, are diffuse status cues. The specific and diffuse status cues indicate different types of status information, but both can be used to identify expertise. Different characteristics of one team member contribute differently to other team members' expectations of his or her expertise (Bunderson 2003). The research findings of Bunderson (2003) suggested that the effect of specific status cues is stronger than the effect of diffuse status cues in diagnosing a team member's task expertise. According to Bunderson (2003), the status information that is more relevant and closer to the task is more likely to be used to determine the expertise attributions. When comparing the specific and diffuse status cues, it is easy to find that the specific status cues are closer (i.e., more highly related to the task) than diffuse status cues in terms of the "path of task relevance." Therefore, the impact of specific status cues on expertise recognition is usually stronger than the impact of diffuse status cues on expertise recognition (Bunderson 2003).

Moreover, the recognized expert members should be granted more influence or authority in decision making and task-related interactions than others, given that the experts assigned with the corresponding influence lead to better team performance (Littlepage et al. 1995). The team member's influence refers to the weight of his or her decision in the team. The more influence the expert has, the higher the weight of his or her decision in the team. The idea of utilizing a team member's recognized expertise to reach the desired team performance depends on the alignment of the team member's expertise and the corresponding influence in his or her decision making. Therefore, it is the alignment of specific status cues, instead of diffuse status cues, that is positively associated with team performance (Bunderson 2003).

(3) Team Coordination

Team coordination refers to the way in which team members' expertise is organized in order to complete the team task (Wittenbaum et al. 1998). The key idea of team coordination is "*who among the* *members does what, as well as when, where, and how they complete their designated tasks*" (Wittenbaum et al. 1998, p. 177). In other words, the core idea of team coordination is expertise coordination (Faraj and Sproull 2000). Expertise coordination refers to *"team-situated interactions aimed at managing resources and expertise dependencies*" (Faraj and Sproull 2000, p. 1555). Expertise coordination can be effective for complex and nonroutine intellectual tasks by accessing the expertise accurately and locating the expertise as needed.

Three streams of research in expertise coordination

According to Bunderson (2003), research in expertise coordination can be categorized into three streams:

1) the transactive memory perspective, which requires members in the same team to have shared knowledge about who knows what when working on the same task (Austin 2003). A team with an effective transactive memory often has high levels of communication and coordination among team members, which often enable a mutual understanding and trust among group members and high team performance (Lewis 2004).

2) the distributed knowledge perspective, in which the team makes decisions by integrating members' shared and unshared knowledge and information in order to achieve high team performance. The distributed knowledge perspective underlines the importance of simulating and validating the unshared information in team decision making (Stasser et al. 1995; Stewart and Stasser 1995).

3) the expert influence perspective, which focuses on how members identify experts in the group and how experts influence the team decisions and outcomes.

As mentioned before, in this study, the human-automation team includes two members: the human and automation. The team members have a shared team goal. The salient feature of the human-automation team is a unidirectional understanding and trust from the human to automation, rather than a bidirectional understanding and trust between the human and automation. Therefore, the transactive memory perspective is not the most suitable to the investigative context. In this study, how to coordinate and combine each member's expertise to achieve high team performance is the key agenda of the human-automation team. Therefore, the expert influence perspective is the most relevant: different from the distributed knowledge perspective, no unshared knowledge is required in the human-automation interaction context.

(4) Task characteristics

Research that empirically studies the impact of task features on team performance is limited (Lewis and Herndon 2011; Baumann and Bonner 2017). A few papers have looked into the impacts of task type, task difficulty, and task uncertainty on team performance, as discussed below.

Task type

According to Lewis and Herndon (2011), there are three types of tasks: produce, choose, and execute tasks. A produce task refers to the task of generating ideas or images and has high demands on the team members' creativity. A choose task includes problem-solving and decision-making tasks in which the team needs to choose the best solution from different options. An execute task requires the team's actual performance such as the basketball team or the football team. Comparatively speaking, the execute task, relative to the produce and choose tasks, benefits more from team coordination because (a) the execute task relies more on the experts' professional skills; and (b) the execute task requires the team members' combined and integrated skills at both the coordinating and physical levels (Lewis and Herndon 2011).

Task difficulty

Task difficulty can impact the recognition of expertise (Bonner 2000; Bonner et al. 2002). Thus, task difficulty should be set at an appropriate level so that the experts and novices can be distinguished. A task that is too difficult leads to uniformly low performance for all team members. Moreover, a task that is too easy can result in uniformly high performance for all team members. Both of the above situations result in low accuracy of expertise recognition (Kramer 1999; Bonner 2000). Therefore, an intermediate level of task difficulty is preferable to reveal the expertise variations in the team members.

Task uncertainty

Task uncertainty is often studied as a contextual variable in team research. For example, the self-regulating team often reaches better team performance when the task uncertainty is higher than when the team is working on routine tasks (Pearce and Ravlin 1987; Cordery et al. 2010). The reason is that the team with high autonomy can make accurate and rapid decisions, especially for unexpected occurrences. Cordery et al. (2010) even found that the impact of autonomy on team performance is negative in the case of low task uncertainty. They also found that task uncertainty is a significant factor that negatively impacts team performance and positively moderates the impact of team autonomy on team performance (Cordery et al. 2010). Therefore, it is necessary to incorporate task uncertainty when studying the human-automation team.

In his seminal work, Galbraith (1974) argued that the essence of uncertainty is the information gap between the amount of information already processed and the information needed to reach the best output (Agbejule 2005; Hartmann and Maas 2011). Task uncertainty means that people lack information to figure out the optimal action strategies, while environmental uncertainty means that the factors related to the organization's performance (but are outside the organization) are unpredictable. There are many different definitions of task uncertainty. Most definitions can be categorized into two types: Galbraith (1974)'s recapitulative definition and the compositional component definition (Table 2.5). Galbraith (1974)'s recapitulative definition focuses on the lack of information from input resources to the desired outcomes. The compositional component definition elaborates the "lack of information" in different dimensions such as analyzability, variety, interdependence, and so forth (Bensaou and Venkatraman 1995).

Among the definitions of task uncertainty above, the most relevant definitions are from Wall et al. (2002) and Cordery et al. (2010) because both papers focused on the uncertainty of the operational level task. Wall et al. (2002, p. 159) defined operational task uncertainty as "*a lack of understanding about cause and effect, or action and outcome, within the system,*" and Cordery et al. (2010, p. 240) defined task uncertainty as a "*lack of prior knowledge about which operational problems will arise, and the best way of* *dealing with them*." In this thesis, in line with Wall et al. (2002) and Cordery et al. (2010), I adopt Galbraith (1974)'s "lack of information" perspective to define task uncertainty.

Table 2.5 Definition of task uncertainty

Types of Definition	Definition	Source
Galbraith (1974)'s recapitulative definition	The greater the uncertainty of the task, the greater the amount of information that must be processed between decision makers during the execution of the task to reach a given level of performance.	Galbraith (1974)
	Task uncertainty is defined in terms of a team's lack of prior knowledge about which operational problems will arise when, and the best way of dealing with them.	Cordery et al. (2010)
	Task uncertainty is defined as the difference between the amount of information required to perform the task and the amount of information already processed.	Chong, V. (1996)
	Task uncertainty is defined as the difference between the information required to solve the problem and the amount already possessed by the system user.	Lamberti and Wallace (1990)
	At a psychological level, operational uncertainty may be defined as a lack of knowledge about the production requirements, of when problems will be met, and how best to deal with them. In other words, operational uncertainty represents a lack of understanding about cause and effect, or action and outcome, within the system.	Wall et al. (2002)
Compositional component definition	We view task uncertainty as a function of three constructs: analyzability, variety, and interdependence. Task analyzability refers to the extent to which there is a known procedure that specifies the sequence of steps to be followed in performing the task. Task variety refers to the number of exceptions or the frequency of unanticipated and novel events that require different methods or procedures for doing the job, consistent with the various notions of task variability, uniformity, predictability, complexity, and sameness. Task interdependence is the extent to which unit [firm] personnel are dependent on one another to perform their individual jobs.	Bensaou and Venkatraman (1995)
	Task uncertainty includes two dimensions: task difficulty (analyzability, lack of knowledge) and task variability. Task difficulty refers to the analyzability of the work itself and the extent to which there is a known procedure that specifies the sequence of steps to be followed in performing the task. Task variability refers to the number of exceptional cases encountered in work requiring different methods or procedures for doing the work.	Van de Ven and Delbecq (1974)
	Task uncertainty is concerned with an individual's beliefs about the completeness of cause-effect knowledge associated with task performance. Two scales were used to develop the measures: task variability and task difficulty.	Hirst (1983)
	Task uncertainty refers to the variability in tasks and the analyzability of the methods to perform the tasks with high variability, and unanalyzable tasks that induce control difficulties and a need for more organic control.	Chenhall (2003)
	Task uncertainty is defined as two task attributes: complexity and interdependence. Task complexity is defined in terms of low levels of analyzability and high levels of variety.	Mani et al. (2010)

2.2.4 Applying team-based decision-making in the human-automation team

Applying the aforementioned theoretical framework to the human-automation team in the Human-Decides and IntelSys-Advises approaches, I argue that (1) the expertise of team members, (2) recognition of team members' expertise, and (3) expertise coordination are also relevant in the human-automation team for the following reasons. First, both humans and automation with high task expertise could contribute to high team performance because members with high expertise often render high team performance (Littlepage et al. 1997). Second, how the human (i.e., the decision maker) accurately recognizes the self and automation's task expertise is instrumental to team coordination in the human-automation team. Third, as mentioned before, teams with different decision-making approaches have different ways of expertise coordination, which then leads to differential team performance. Regarding the task characteristics, I focus on task uncertainty because prior literature suggests that it (a) may negatively impact team performance (Cordery et al. 2010); and (b) may affect human's control behavior when making decisions, as elaborated in the subsequent section.

2.3 Impact of Uncertainty on Personal Control

2.3.1 Uncertainty → Behavior control

According to Thompson (1981), behavioral control is defined as "*a belief that one has a behavioral response available that can affect the aversiveness of an event*" (Table 2.4). Behavioral control is an effective way to respond to an unpleasant situation because behavioral control enables people to predict and prepare for aversive events. When working on uncertain tasks, people can reduce or avoid undesired outcomes by self-administrating the task (Thompson 1981).

The literature has suggested that task uncertainty impacts humans' need for control (Hartmann 2000). Task uncertainty refers to the lack of information between *what is known* and *what needs to be known* to reach the desired outcome (Galbraith 1974). The desire to seek information caused by task uncertainty leads to an increased need for control (Chenhall 2003; Hartmann and Maas 2011).

2.3.2 Human's nature: Uncertainty → Control behavior

Having and performing control is also indicative of human nature, especially when unexpected events take place. Personal control can enhance the human's sense of mastery and personal competence; people believe they can reduce the aversive effect of uncertainty by taking more control (deCharms 1968). For example, when facing highly uncertain situations, people tend to feel low control or a lack of mastery; under this circumstance, people seek more control, which then enhances a sense of personal competence (Ashford and Black 1996).

2.3.3 Evidence from neurosciences about uncertainty and control

After a long history of evolution, humans have become very adaptive in terms of learning from uncertain situations (Mushtaq et al. 2011). Evidence from neuroscience has suggested that "*cognitive control processes are at the heart of uncertainty in decision-making contexts*" (Mushtaq et al. 2011, p. 1). According to Mushtaq et al. (2011), the relationship between cognitive control and uncertainty can be explained from three perspectives. First, there is a significant overlap between the concept of uncertainty and the concept of cognitive control (Mushtaq et al. 2011): (a) both concepts emphasize the importance of the mismatch between humans' existing schemata and the current environment; and (b) this mismatch leads to a behavioral reaction in order to adapt to the environment. Second, the neural networks associated with uncertainty are *remarkably overlapped with* the neural networks associated with cognitive control, for example, the orbitofrontal cortex (OFC). Third, the process of uncertainty often leads to the "need for control."

The orbitofrontal cortex (OFC) is a region in the frontal lobes that is linked to the cognitive process of decision making (Kringelbach 2005). Specifically, OFC controls the perception of reward and punishment feedback and facilitates the learning of stimulus–reward relationships (O'Doherty et al. 2003). Research in cognitive control has found that OFC detects information associated with valenced outcomes. The information then stimulates the need for behavioral control (O'Doherty et al. 2003; Braver and Ruge 2006; Mushtaq et al. 2011). Moreover, the level of uncertainty is associated with activation of the OFC

(Hsu et al. 2005). OFC detects and evaluates the stimuli induced by the uncertainty. The output of the OFC then motivates behavioral adaptation, which is a function of uncertainty levels (Hsu et al. 2005).

2.3.4. Personal control model

The relationship between task uncertainty and personal control can be further explained by the personal control model proposed by Greenberger and Strasser (1986). Personal control is defined as "*the individual's beliefs, at a given point in time, in his or her ability to effect a change, in a desired direction, on the environment*" (from Greenberger and Strasser 1986, p. 164). The personal control model views people's control behavior as a response to uncertain events. Specifically, people's response to task uncertainty is a function of perceived control and desired control. Based on the difference between perceived control and desired control, people perform different control behaviors.

- Based on the personal control model, there is a threshold level that depends on the difference between possessed control and desired control to ignite the individual's control behavior (Greenberger and Strasser 1986). Generally speaking, a low uncertainty task is often below the threshold; therefore, no control behavior is activated. On the contrary, a high uncertainty task is above the threshold and can activate an individual's control behavior. The specific control behavior depends on the difference between the possessed control versus the desired control (Table 2.6). This difference reflects the extent to which the individual is motivated to seek control. When the difference *approaches zero*, the amount of desired control and the amount of possessed control are at the same level, which indicates that the individual is in a balanced cognitive state. Therefore, in this state, people have no motivation to seek more control and are satisfied with the current control level. In reality, it is unlikely that the "*difference approaches zero*" (Greenberger and Strasser 1986).
- When the possessed control is significantly higher than the desired control, the individual perceives that he or she has more control than desired. Therefore, the individual is in an imbalanced stage. To achieve the balanced stage, the individual will perform cognitive reactions through cognitive adjustment. The individual will cognitively reevaluate the possessed control and desired control so

that the desired control approaches the possessed control, which indicates a balanced stage. Specifically, there are two possible ways of cognitive adjustment: (a) cognitively adjust the desired control to a higher level; and (b) cognitively decrease the possessed control to a lower level. It is worth noting that the path from the initial imbalanced stage to the ultimate balanced stage is through cognitive adjustment rather than actual control behavior (Table 2.6). In reality, *possessed control being significantly more than desired control* is unlikely to happen because usually few people would complain about having too much control (Greenberger and Strasser 1986).

- The *situation of "the possessed control is significantly less than the desired control"* means that the individual wants more control than what he or she possesses. In this situation, the individual tries to seek more control until the ratio approaches one, and then he or she enters a balanced state. Alternatively, if the individual finds that his or her control-seeking behaviors are ineffective in achieving the desired outcome, the individual will cease his or her control-seeking behaviors and will experience a state of learned helplessness. The path from the initial imbalanced stage to the ultimate balanced stage is through actual control behavior (either through more control-seeking behavior or learned helplessness), along with cognitive adjustment (Table 2.6). The *situation of "the possessed control is significantly less than the desired control"* often happens in reality (Greenberger and Strasser 1986) because it is human nature to desire more control.

Table 2.6 Individual's control behavior under high task uncertainty

Difference (desired/possessed)	Control Behavior
possessed control = desired control	Less motivated to seek more control
possessed control > desired control	Cognitive adjustments: Reevaluate control possessed and desired
possessed control < desired control	• Seek more control • Learned helplessness: Realizes no relationship between control actions and outcomes.

Learned helplessness

Learned helplessness means that the individual has a strong desire for control and then tries to seek more control; however, the individual realizes that no matter how much control he or she possesses, there is no impact on the outcome. Then the motivation to seek control will cease. Learned helplessness can also be explained by the *two-stage process model* proposed by Rothbaum et al. (1982).

The two-process model of control suggests that control can occur at two different levels: primary control and secondary control (Rothbaum et al. 1982; Sasaki and Kim 2010; Nicole et al. 2020). Primary control involves direct “seek more control” behavior, by which the individual exerts influence on the environment in order to reach the desired outcome. However, when the “seek more control” behavior is unsuccessful or useless, the individual will proceed to secondary control, wherein the individual accepts the unchangeable circumstances and adapts himself or herself to the current environment to "flow with the current" (Rothbaum et al. 1982), which indicates learned helplessness behavior.

Because most people have the general desire to be in control, situations involving “possessed control > desired control” and “possessed control = desired control” are less likely to happen (Greenberger et al. 1989; Greenberger and Strasser 1986). As such, when applying the personal control model to the human-automation interaction context, we focus more on the “possessed control < desired control” situation (Table 2.6). As aforementioned, during the “possessed control < desired control” situation, an individual will first seek more control; if the individual finds that seeking more control behavior has no bearing on the desired outcome, learned helplessness behavior will likely occur.

2.4 Research Gap and Possible Solutions

The above review of the human-automation interaction literature suggests that there are two major research gaps to be addressed: (1) What is the best decision-making approach in human-automation interaction? (2) If there is no best approach, what is the possible boundary condition to explain the impacts of DMAs on team performance? To this end, this thesis approaches these issues from the team perspective by viewing human and automation collectively as a team.

The advantage of the team perspective is that we can access a rich body of literature and established theoretical frameworks from the team research, especially the team-based decision-making area. In this vein, four factors are found to impact team performance significantly: members' expertise, expertise recognition, expertise coordination, and task characteristics. The core difference in the IntelSys-Decides, Human-Decides, and IntelSys-Advises approaches involves expertise coordination; that is, these three DMAs use different expertise coordination strategies. Therefore, when comparing IntelSys-Decides, Human-Decides, and IntelSys-Advises, the key is to compare which expertise coordination strategy of the DMA leads to the best team performance.

For the IntelSys-Decides approach, IntelSys makes task allocation decisions based on humans' workload, performance, behavior, and critical events. For the Human-Decides and IntelSys-Advises approaches, the human makes decisions based on a broader range of factors than those considered by IntelSys. The difference between Human-Decides and IntelSys-Advises is that in the IntelSys-Advises approach, the human gets advice from IntelSys. Taking one step further, the decision-making capabilities of IntelSys and the human determine their task allocation strategies. Based on the personal control model, task uncertainty is an important factor affecting humans' task allocation decisions.

In the next section, I first compare the relative advantage of IntelSys and humans in terms of decision-making mechanisms. I then compare the relative advantage of IntelSys-Decides, Human-Decides, and IntelSys-Advises in decision-making quality. Finally, I identify the boundary conditions for the three DMAs.

3 Research Model and Hypotheses

3.1 Apply the Personal Control Model in Human-automation Interaction

As aforementioned, while the personal control model implies that low task uncertainty has little influence on humans' task allocation decisions in human-automation teams, high task uncertainty could result in high variation with regard to humans' task allocation decisions. Yet, this model does not specify what types of individuals are more prone to seek additional control or experience learned helplessness. To this end, we argue that expertise is the key differentiating factor such that human control behaviors will differ between the experienced versus the inexperienced in the case of high task uncertainty.

In particular, Miller and Parasuraman (2007) suggest that humans with lower/higher task expertise are less/more confident about their expertise in achieving high task performance. Following this line, novices under high task uncertainty are more likely to develop a sense of helplessness after trying to perform the task, realizing that automation can perform the task better than they can, and hence allocating more tasks to automation for task execution. In contrast, due to their confidence in their expertise, experts under high task uncertainty are more likely to seek additional control by allocating more tasks to themselves, even if the automation is at first designed to outperform human operators. Based on our discussion thus far, Table 3.1 illustrates how task uncertainty and human expertise jointly impact humans' task allocation decisions when humans have the decision-making authority in Human-Decides and IntelSys-Advises.

Table 3.1 Human's control behavior in HAI

Task Uncertainty	Control Behavior	Impact on the task-allocation decision (Human-Decides and IntelSys-Advises)
Low	No impact	No impact on human's task allocation decision
High	Possessed control = Desired control (less likely to happen)	No impact on human's task allocation decision
	Possessed control > Desired control (less likely to happen)	No action on behavior control and no impact on human's task allocation decision
	Possessed control < Desired control - Learned helplessness	***Novices**:* Allocate more tasks to automation
	Possessed control < Desired control - Seek more control	***Experts**:* Allocate more tasks to the self

3.2 Human vs. IntelSys in Decision Making

In human-automation interaction, both humans and automation involve in the task execution, while both humans and IntelSys have the capability to make task allocation decisions. These two different decision-makers make it possible for three different DMAs. For IntelSys-Decides, the IntelSys makes task-allocation decisions based on human's workload. For Human-Decides, the human has decision-making authority. For IntelSys-Advises, both human and IntelSys work together to make decisions, but the human has the final decision-making authority. Given the theoretical premise that automation outperforms most humans, according to the expertise coordination principle, the team performance could be guaranteed by allocating most tasks to automation. Therefore, the comparison of three DMAs depends on the extent to which the decisions of DMA follow the expertise coordination principle.

An effective comparison of IntelSys-Decides, Human-Decides, and IntelSys-Advises requires an understanding of the decision-making nature of humans versus IntelSys. There are at least three advantages of humans over IntelSys in terms of decision making for the switch between the manual mode and the automation mode. First, humans are more flexible than IntelSys, as humans can make a switching decision (i.e., manual to automation or vice versa) whenever they feel it is necessary or desirable. In contrast, IntelSys can only switch the mode under specific predefined circumstances (Sheridan and Parasuraman 2006). Second, humans can make more personalized decisions than IntelSys because humans take more individual characteristics (mood, mental state, physical condition, etc.) into consideration, whereas IntelSys only considers factors such as workload and critical events that are prespecified in its algorithms (Miller and Parasuraman 2007; Li 2013). Third, by integrating different sources of information such as environmental conditions and task uncertainty, humans can make more informed decisions than IntelSys (Sheridan and Parasuraman 2006).

Importantly, the above assertion that humans' advantages over IntelSys is contingent on the assumption that humans' decisions are more likely to follow the expertise coordination principle. Most prior HAI studies, however, do not discuss this assumption and implicitly take it as a given (e.g., Kaber 1996; Parasuraman and Wickens 2008; Li 2013). Should this assumption hold true, human decision makers should always allocate more tasks to the team member with higher expertise. Yet, our discussion of the personal control model suggests that humans do not always make decisions that follow the expertise coordination principle, which thus challenges this assumption that has long been taken for granted.

3.3 Human's Control Behavior

Prior team research offers little insight into the impact of task characteristics on team coordination (Lewis and Herndon 2011; Ren and Argote 2011; Baumann and Bonner 2017). To bridge this knowledge gap, this study draws on knowledge from the personal control model and examines how task uncertainty impacts team coordination in the context of human-automation interaction.

According to the personal control model (Greenberger and Strasser 1986), low task uncertainty does not affect humans' control behavior. In this situation, generally speaking, humans would make better decisions than IntelSys because humans make more flexible, personalized and informed decisions (Visser et al. 2010; Chen and Barnes 2012). Thus, in Human-Decides and IntelSys-Advises DMAs where humans have the decision-making authority, humans are more likely to make rational task-allocation decisions based on the expertise coordination principle (i.e., humans would allocate a larger share of the task to team members who have better task expertise).

However, under high task uncertainty situations, humans' control behavior will differ between those with low expertise versus those with high expertise (Table 3.1). First, let us consider the case when the human has low expertise. When the human has low expertise, he or she is more likely to realize that no matter how much of the task he or she wants to do, his or her endeavor has a very limited impact on the performance outcome. Therefore, the human's control behavior is more likely to indicate learned helplessness, which means that he or she will allocate a greater share of the task to the automation and will allocate a lower share of the task to himself or herself. However, learned helplessness behavior, in essence, fits with the "rational" expertise coordination principle because the expert (i.e., automation) in this circumstance has more influence than the other members (i.e., humans).

Next, when the human has high expertise under high task uncertainty, he or she is more likely to demonstrate "seek more control" behavior (Table 3.1). This is because (1) experts are more likely to take more control (Parasuraman et al. 2007); and (2) it is human nature to grasp more control under high task uncertainty conditions (Thompson 1981; Greenberger and Strasser 1986; Ashford and Black 1996; Chenhall 2003; O'Doherty et al. 2003; Braver and Ruge 2006; Hartmann and Maas 2011; Mushtaq et al. 2011). Under this circumstance (i.e., high task uncertainty and high human expertise), the human will allocate a greater share of the task to himself or herself while allocating a relatively lower share of the task to the automation, even if the automation could outperform the human in task execution. In this vein, the human is more likely to make relative "irrational" decisions that violate the expertise coordination principle.

3.4 Hypothesis Formulation

The aforementioned three decision-making approaches (i.e., IntelSys-Decides, Human-Decides, and IntelSys-Advises) have different expertise coordination principles in the human-automation team context. For the IntelSys-Decides DMA, IntelSys allocates the share of tasks based on the human's workload (Kaber et al. 2006; Table 3.2). For the Human-Decides DMA, the human has the full authority to make task allocation decisions. For the System-advice DMA, while the human also has full decision-making authority, he or she is also under the influence of the advice by IntelSys (Li 2013). It is worth noting that the comparison of three DMAs is built upon the assumption that the automation outperforms most humans (both novices and experts). This is also the assumption of most research in human-automation interaction field (Kaber 1996). The rational of this assumption is that only the automation performs better than most humans, it is useful and meaningful to use automation in task execution.

Synthesizing team-based decision making and the personal control model, I also identify two boundary conditions (i.e., task uncertainty and human expertise) that may affect the effectiveness of DMAs in human-automation teams. Next, I develop four comparative hypotheses of IntelSys-Decides, Human-Decides, and IntelSys-Advises across four different scenarios, as shown in Table 3.2 and Table 3.3.

Table 3.2 IntelSys vs. Human in decision making

	IntelSys	**Human**	**Team Performance**
Low Task Uncertainty	**Only** consider limited factors (critical events, workload, behavior, performance, and others) **Does not** consider individual factors	More flexible; Consider more individual factors; Better situation awareness	For novices: Human-Decides > IntelSys-Advises > IntelSys-Decides For experts: Human-Decides > IntelSys-Advises > IntelSys-Decides
High Task Uncertainty Low Expertise	Same as above	Learned helplessness: Allocate a greater share of the task to automation	For novices: Human-Decides > IntelSys-Advises > IntelSys-Decides
High Task Uncertainty High Expertise	Same as above	Seek more control: Allocate a greater share of the task to the self	For experts: IntelSys-Decides > IntelSys-Advises > Human-Decides

Table 3.3 Research hypotheses for team performance across four scenarios

Task Uncertainty / **Human Expertise**	**Low**	**High**
Low	Cell 1 Human-Decides > IntelSys-Advises > IntelSys-Decides	Cell 3 Human-Decides > IntelSys-Advises > IntelSys-Decides
High	Cell 2 Human-Decides > IntelSys-Advises > IntelSys-Decides	Cell 4 IntelSys-Decides > IntelSys-Advises > Human-Decides

First, under the condition of low task uncertainty, humans are more likely to make task allocation decisions that follow the expertise coordination principle (i.e., allocating more of the task to either the self or automation that has better task expertise; Cells 1 and 2 of Table 3.3). Under this premise, humans' decisions are more likely to follow the expertise coordination principle. As mentioned before, the human has at least three advantages over IntelSys in decision-making: (a) the human's decision is more flexible; (b) the human considers more individual-level factors; and (c) the human has better situation awareness (Table 3.2). As such, humans' decisions, relative to decisions by IntelSys, could allow for better team performance (Bonner et al. 2002). Following this logic, in the case of low task uncertainty, the team performance of the Human-Decides DMA and the IntelSys-Advises DMA, in which the human makes task allocation decisions, would be better than the team performance of IntelSys-Decides, in which IntelSys makes the task allocation decisions.

In addition, between the Human-Decides and IntelSys-Advises DMAs, the performance will be better for the former approach. Different from the human who has full authority in the Human-Decides approach, the human will receive advice from IntelSys in the IntelSys-Advise approach. Because humans' decisions, as discussed above, would be better than decisions of IntelSys under the condition of low task uncertainty, IntelSys' advice is more akin to a distraction for the human. As such, the team performance of Human-Decides would be better than the team performance of IntelSys-Advises (Table 3.2).

It is worth noting that I do not distinguish between experts and novices because both groups would follow the same logic. The above discussion collectively suggests that, in terms of team performance, Human-Decides will outperform IntelSys-Advises, and IntelSys-Advises will outperform IntelSys-Decides, leading to our first and second hypotheses:

Hypothesis 1*: The ranking order of team performance for* ***low*** *task uncertainty task with* ***low*** *human expertise will be: Human-Decides > IntelSys-Advises > IntelSys-Decides.*

Hypothesis 2*: The ranking order of team performance for* **low** *task uncertainty task with* ***high*** *human expertise will be: Human-Decides > IntelSys-Advises > IntelSys-Decides.*

Next, in the case of high task uncertainty and low human expertise (Cell 3 of Table 3.3), the human is more likely to demonstrate learned helplessness. As elaborated earlier (per Table 3.2), the novice is more likely to realize that he or she cannot reach the desired results, no matter how hard he or she tries to carry out the task under high uncertainty (Greenberger and Strasser 1986); in this vein, because the automation performs better than the human, the novice will be willing to allocate a greater share of the task to the automation. Such an inclination is actually in line with the team-based decision-making expertise coordination principle, which allocates a greater share of the task to team members with better task expertise, i.e., automation in the investigative context of this study (Faraj and Sproull 2000). Following this logic, a team could perform better if more task-allocation decisions are made by the human operator, without the inputs from IntelSys. In other words, team performance would be the best for the Human-Decides approach wherein the human has full control, followed by the IntelSys-Advices approach wherein the human has full control, but is under the influence of IntelSys' advice; and finally, the IntelSys-Advises approach wherein the human does not have any decision-making authority. The above discussion leads to the third hypothesis:

Hypothesis 3: *The ranking order of team performance for* **high** *task uncertainty task with* ***low*** *human expertise will be: Human-Decides > IntelSys-Advises > IntelSys-Decides.*

Finally, for experts in the case of high task uncertainty (Cell 4 of Table 3.3), as discussed earlier, they would prefer to stick to seeking more control, and even experiencing undesirable performance. Under this circumstance, the high-expertise human is more likely to "seek more control" and allocate a greater share of the task to oneself, even though he or she knows that the automation could outperform him or her

(Table 3.2); however, such an inclination would violate the expertise coordination principle, given that the human does not allocate the task to the more competent team member (i.e., automation). In contrast, the decisions by IntelSys in this situation would be more consistent with the expertise coordination principle. Following this line, because IntelSys-Decides involves no humans in the decision-making process, it logical to expect that the IntelSys-Decides approach could outperform the Human-Decides and IntelSys-Advises approaches.

Next, between the Human-Decides and IntelSys-Advises DMAs, the advice by IntelSys under the IntelSys approach can play a positive role in correcting the human's inclination to seek more control, because, as aforementioned, the decision quality of IntelSys would be better than that of the experts in the case of high task uncertainty. The above discussion collectively suggests that team performance will be the best in IntelSys-Decides, followed by IntelSys-Advises, and finally Human-Decides, leading to the fourth hypothesis:

Hypothesis 4: *The ranking order of team performance for* ***high*** *task uncertainty task with* ***high*** *human expertise will be: IntelSys-Decides > IntelSys-Advises > Human-Decides.*

4. Experimental Design

I conducted a multitask lab experiment to test the hypotheses. The multitask experiment has been widely used in HAI research (e.g., Parasuraman et al. 1993; Parasuraman et al. 1996; Bliss and Dunn 2000; Bailey et al. 2006). One of the most popular multitasks platform is NASA's Multi-Attribute Task Battery II (MATB-II) Software (Santiago-Espada et al. 2011). MATB is a simulation of pilot tasks. It includes system monitoring tasks, tracking tasks, communication tasks, and resource management tasks (Santiago-Espada et al. 2011). Some researchers have also developed specific programs or software for multitask experiments. For example, the ***Multitask*** software developed by Endsley and Kaber (1997) has been used in works by Endsley and Kaber (1999), Clamann et al. (2002), and Kaber and Endsley (2006). The ***Multitask*** software includes two tasks: the primary task is a simulated air traffic control task, and the secondary task is a gauge-monitoring task. In the primary task, there are different levels of automation, from total manual to total automation. The gauge-monitoring task is used to measure the operator's workload. ***Multitask*** has been used to evaluate the performance of LOA and IntelSys-Decides (Kaber 1996; Kaber and Endsley 2004).

4.1 Experimental Platform – Multitask

Following Endsley and her associates, I developed a two-task program (Figure 4.1). The primary task is a shooting game (middle section of Figure 4.1), and the secondary task (left side of Figure 4.1) is a gauge-monitoring task, which is similar to the monitoring task in ***Multitask***. The shooting task has two modes: (a) the automation mode, in which the computer serves as the automation and plays the shooting game automatically; and (b) the manual mode, in which the human participants play the shooting game and shoot down hazards.

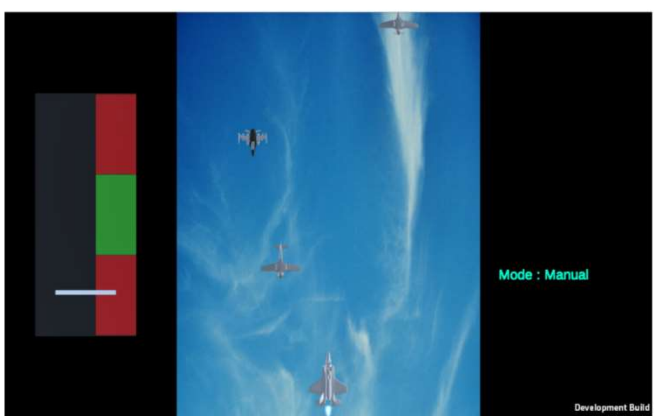


Figure 4.1 Two-tasks program

The team score is incremented every time a hazard is shot down, either in manual mode or automation mode. IntelSys often plays the role of adaptive aid, which aims to help or replace humans with regard to making decisions (Land et al. 1995; Scerbo 1996; Miller and Hannen 1999; Scerbo 2008). The goal of IntelSys is to provide human operators with the right information at the right time (Miller and Hannen 1999). The basic functions of IntelSys include monitoring the human workload and allocating tasks between humans and automation (Scerbo 2008). In this study, IntelSys is a program embedded in the gaming platform. By recording the player's performance of the gauge-monitoring task, IntelSys measures the participant's workload and makes switching decisions (e.g., switching from the automation mode to the manual mode, or vice versa) based on some criteria, which will be elaborated in the following section.

4.1.1 Shooting task

The primary task is a shooting game (middle section of Figure 4.2). For the primary task, the player is a pilot of a jet fighter at the bottom of the screen. The goal of the player is to shoot down as many enemy hazards as possible. The player's fighter will never die. The player can move the fighter to the left by pressing "←" and to the right by pressing "→"; by pressing the space key, the player's fighter can shoot bullets with the frequency of one press-one bullet.

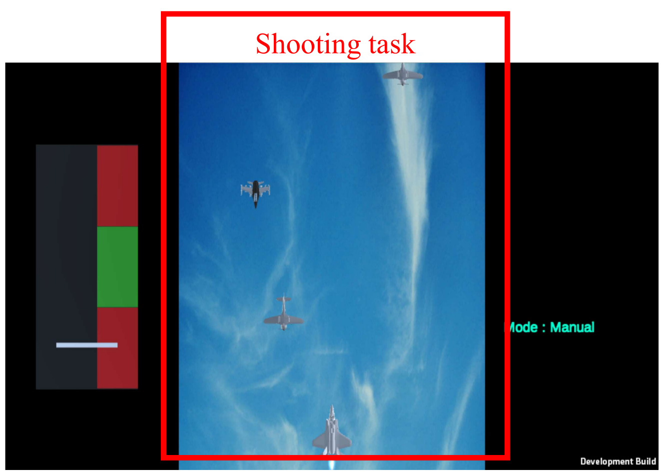


Figure 4.2 Shooting task

4.1.2 Gauge-monitoring task

In the gauge-monitoring task, there is a fixed-scale display and a moving pointer (left side of Figure 4.3). The pointer can move between the green area and the red area. The green area is the "safe" region, whereas the red areas located above and below the green area are the "unsafe" regions (Figure 4.3). When the pointer (i.e., the white bar in Figure 4.3) moves randomly from the green area to either the upper or the lower red area, IntelSys will record this move as one deviation (from the safe region). When a deviation appears, the participants need to correct the deviation by pressing the upper key "↑" when the pointer moves below the green area and by pressing the down key ''↓" when the pointer moves above the green area. The player must correct the deviation within three seconds because if the pointer stays in the red area or unsafe regions for more than 3 seconds without any correction, the pointer will automatically return to the center of the green area. Once the pointer returns to the center of the green area without any human correction, IntelSys will record this event as missing one deviation by the player. In the same vein, if the player successfully corrects the deviation, IntelSys will record this correction made by the players. On average, there are 10 deviations per 30 seconds.

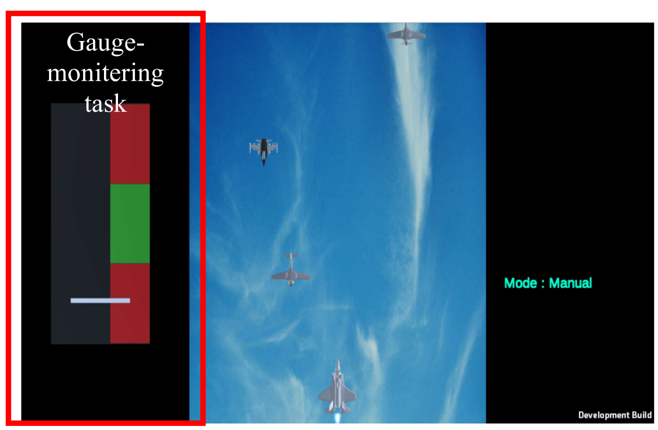


Figure 4.3 Gauge-monitoring task

4.2 Factor Operationalization

I manipulate two factors in the shooting game: task uncertainty (high versus low) and decision-making approaches (i.e., IntelSys-Decides, Human-Decides, and IntelSys-Advises). In total, there are (a) three versions of the shooting game in terms of decision-making approaches (i.e., IntelSys-Decides, Human-Decides, and IntelSys-Advises) and (b) two versions of the shooting game in terms of task uncertainty (i.e., low and high). For all versions of the shooting game, there are two playing modes: the manual mode and the automation mode. In the manual mode, it is the player who plays the shooting game; in the automation mode, it is the automation (or the automatic shooting function of this program) that plays the shooting game. The player is required to always work on the gauge-monitoring task, no matter whether he or she is in the manual mode or automaton mode.

4.2.1 Task uncertainty implementation

The task uncertainty setting is described in Table 4.1. For both low and high task uncertainty settings, there are two types of enemy hazards: white and black. The white hazard moves at a lower speed and only takes one bullet to shoot down. The black hazard moves at a higher speed and needs two bullets to shoot down. In the instructions of the shooting game, the participants are informed that they will score 10 and 30 points when shooting down a white and a black hazard, respectively.

Table 4.1 Manipulation of task uncertainty

Task Uncertainty	Settings
Low	With a predictable straight path, the hazards come out from the top of the screen to the bottom of the screen.
High	The hazards have an unpredictable path with 2*16*16 = 512 variations. 2: Up and down directions 16: The hazard can deviate from the straight path with 16 different angles toward the left. 16: The hazard can deviate from the straight path with 16 different angles toward the right.

For low task uncertainty, each hazard has a predictable path, which starts from a fixed position at the top of the screen, moving straight forward to the bottom of the screen. The participants can easily identify where a hazard comes out and can easily predict its path.

For the high task uncertainty, the path of each hazard, however, would be unpredictable. The hazard is not moving straight to the bottom of the screen. Instead, there are in general, three types of variation in terms of the path: a hazard may (a) move upward or downward; (b) move to the left with 16 different angles; and (c) move to the right with 16 different angles. Therefore, in total, there are 2*16*16 = 512 possible variations in terms of hazard paths. Moreover, each hazard changes its path three times, and each change is randomized based on the aforementioned 512 variations. Thus, the hazard path under high uncertainty is highly unpredictable from the players' point-of-view.

4.2.2. Decision-making approach implementation

4.2.2.1 Implementation of the IntelSys-Decides approach

As discussed in Section 2.1, there are four methods to trigger the IntelSys-Decides mode: critical events, performance measurement, psychophysiological assessment, and the behavior modeling strategy. The psychophysiological assessment approach is one of the most widely used methods in lab experiments when studying human-automation interaction. Hence, following Kaber and Riley (1999), Clamann et al. (2002), Kaber and Endlesy (2004), Kaber et al. (2005), and Kaber et al. (2006), I use the psychophysiological assessment method in this experiment (Table 4.2).

When applying the psychophysiological assessment in the IntelSys-Decides approach, the workload is typically measured as the indicator that triggers the mode switching. Specifically, in the automation mode, if the player's workload is too low, the shooting task should be in the manual mode so that the player is involved in the task rather than being out of the loop (Kaber and Endlesy 2004). In the manual mode, if the player's workload is too high, the shooting task should be in the automation mode so that the player is not too tired.

Measure workload in Multitask

Kaber and Riley (1999) were among the first that proposed using the secondary task (e.g., gauge-monitoring task) as an objective way to measure the subject's workload. According to the Multiple Resource Theory (Parasuraman et al. 1992; Wickens 2002; Wickens 2008), both the gauge-monitoring task and shooting task consume the same type of cognitive resources (Wickens 2002; Clamann et al. 2002; Kaber and Endlesy 2004). When a player performs both tasks at the same time, the two tasks compete for the player's cognitive resources. If the player focuses on the shooting task, then only few cognitive resources will be allocated to the gauge-monitoring task, and vice versa (Parasuraman et al. 1992).

Before the experiment, subjects are instructed to focus mainly on the shooting task and allocate additional cognitive resources to the gauge-monitoring task. When the player's workload is high, the player mainly focuses on the shooting task, and only a few or even no resources will be allocated to the gauge-monitoring task. The result is that the performance of the gauge-monitoring task goes down. As such, the performance of the gauge-monitoring task could negatively affect the player's workload. A low level of gauge-monitoring task performance suggests that the player's workload is high; in this situation, it would be better if the shooting game is operated in the automation mode. In contrast, a high level of gauge-monitoring task performance suggests that the player's workload is low; in this case, it would be better if the shooting game is operated in the manual mode so that the player is involved in the task rather than being out of the loop. Note that the performance of the gauge-monitoring task is calculated as a ratio of the number of deviations corrected by the player versus the total number of deviations that takes place within every 30 seconds; in other words, this index of the secondary task is calculated once every 30 seconds.

Using the secondary task to measure the workload is an effective method and has been extensively used in prior literature, including work by Kaber and Riley (1999), Clamann et al. (2002), and Kaber et al. (2006). In these papers, the core idea of the IntelSys-Decides mechanism is to find a baseline level of the workload (Kaber and Riley 1999) that determines whether the task should be operated in the manual or automation mode. Following this line of research, I also calculate the baseline level of the gauge-monitoring task to trigger the mode switch under the IntelSys-Decides approach.

Baseline level: Threshold to make a switch

As indicated above, the players need to play both tasks during the experiment. When the shooting task is in the automation mode, the player only needs to engage in the gauge-monitoring task, which is very easy to play; therefore, the workload level will be the lowest. When the shooting task is in the manual mode, the player needs to play both tasks, and the workload will be the maximum level in the playing process. I calculated the baseline level in the pilot study, which will be introduced later. In the pilot study, there was a five-minute stage in which the automation mode and manual mode appeared alternately in a 30-second cycle. The 30-second manual mode measures the overload level (the maximum workload level), and the 30-second automation measures the underload level (the minimum workload level). There are 10 workload values in total, including five overload values and five underload values, which are then used to determine the workload baseline level (Kaber and Riley 1999; Clamann et al. 2002; Kaber et al. 2006). There are two baseline level values: a lower bound baseline and a higher bound baseline. The values of the lower and higher bound baselines were calculated in the pilot study, which will be described later (section 4.4).

IntelSys keeps monitoring participants' performance of the gauge-monitoring task in each 30-second cycle. When the shooting task is in the manual mode, if the player's performance on the gauge-monitoring task decreases below the lower bound baseline (meaning a high workload), IntelSys will switch to the automation mode. The participants will see the notification "*now switch to automation mode*" on the screen. When the shooting task is in the automation mode, if the participants' performance on the gauge-monitoring task reaches the higher bound baseline (meaning a low workload), IntelSys will switch to the manual mode (Clamann et al. 2002; Kaber et al. 2006). The player will see the notification "*now switch to manual mode*" on the screen.

The goal of the baseline is to keep the participants' workload between the minimum and maximum acceptable range. At the minimum workload level (automation mode), participants' situation awareness may be compromised due to a lack of involvement (Kaber and Riley 1999). At the maximum workload level (manual mode), participants may experience increased frustration and may lose confidence in their ability to perform the task functions (Kaber and Riley 1999). According to Kaber and Riley (1999), who designed the ***Multitask platform***, the participants are instructed to focus on the shooting task and allocate any remaining attentional resources to perform the gauge-monitoring task to as best as they can (Kaber and Riley 1999). In the IntelSys-Decides approach, the participants were told that a decrease in their performance on the gauge-monitoring task would result in a mandatory switch from manual to automation in the shooting task. In the IntelSys-Advises approach, participants will be informed that if their performance on the gauge-monitoring task decreases, IntelSys will provide switching *suggestions,* and they can decide at will whether they would like to switch or not.

4.2.2.2 Implementation of the Human-Decides approach

In the Human-Decides approach, it is the player, rather than IntelSys, who decides when to make a switch between the manual mode and the automation mode (Table 4.2). By simultaneously pressing the left "Shift" and the letter "o" on the keyboard, the player can make a switch at any time when he or she feels like doing so.

Table 4.2 Manipulation of decision-making approaches

DMAs	Manipulation
IntelSys-Decides	IntelSys monitors the participant's workload and makes all mode switch decisions.
Human-Decides	The participant makes all mode switch decisions by pressing left "shift" + "o".
IntelSys-Advises	The participant switches modes by pressing the left "shift" + "o"; IntelSys provides mode-switching advice; the participant can accept or reject the advice.

4.2.2.3 Implementation of the IntelSys-Advises Approach

In the IntelSys-Advises approach, the player can make a switching decision at any time by pressing the left "Shift" and the letter "o" (the same as in the Human-Decides approach). Meanwhile, IntelSys also monitors the performance of the gauge-monitoring task in a 30-seconds cycle (Table 4.2).

Under the manual model, when IntelSys detects that the performance of the gauge-monitoring task is lower than the lower bound baseline, a prompt with the switching suggestion "*advice: switch to automation mode. Press 'y' to accept, 'n' to reject*" will show up on the screen. The player can then decide to accept or reject the advice. If the player presses the letter "y", IntelSys will switch to the automation mode; otherwise, the player remains in the current manual mode. In contrast, under the automation model, when IntelSys detects that the performance of the gauge-monitoring task is above the higher bound baseline, a prompt with the switching suggestion "*advice: switch to manual mode. Press 'y' to accept, 'n' to reject*" shows up on the screen. The player then can decide to accept or reject this advice. If the player presses the letter "y", IntelSys will switch to the manual mode; otherwise, the player remains in the automation mode.

4.2.3 Expertise

Expertise is measured as the participants' game experience. An expert gamer refers to a participant who has high task expertise, while a novice gamer refers to a participant who has low task expertise. The expert and novice gamers are different in terms of their cognitive skills (Boot et al. 2008). Compared to novices, expert gamers can better track fast-moving objects, capture changes in the objects stored in visual short-term memory, act very quickly, and be more effective in mental object rotation (Boot et al. 2008). More importantly, these cognitive skills can transfer to other similar tasks (Boot et al. 2008). Prior literature suggests that expert gamers, relative to non-gamers, will perform better in new games (Boot et al. 2008; Karle et al. 2010).

In this research, I adopt the measure of game expertise from Green and Bavelier (2003). To be qualified as an expert, a participant needs to play action games for at least four sessions/week, at least one hour per session, and for a period lasting at least six months. A participant who has very little (or even no) game playing experience is recognized as a novice gamer with low expertise (Karle et al. 2010). This measure of game expertise has been widely applied in prior studies (e.g., Karle et al. 2010; Feng et al. 2007). The shooting task in my experiment belongs to the action game genre, and I measure the player's action game experience before the formal experiment.

4.3 Experimental Process

A 3 (IntelSys-Decides, Human-Decides, and IntelSys-Advises) by 2 (low and high task uncertainty) factorial design was conducted. The participants first answered questions about their demographic backgrounds and prior game-playing experience. Before the experiment, the participants were told that they should primarily focus on the shooting task while allocating their remaining cognitive energy to the gauge-monitoring task. Participants were also told that they should work with the computer as a team, and only the total team score (rather than the human effort score) will be considered. The goal of this game is to maximize the overall team score, which is the sum of the human participants' shooting score and the shooting score of the automation during the seven-minute testing stage.

The experiment included a five-minute training stage and a seven-minute testing stage (Parasuraman et al. 1996; Kaber and Riley 1999; Bailey et al. 2006; and Kaber et al. 2006, Table 4.2). The goals of the training stage included (a) familiarizing the participants with both the shooting and gauge-monitoring tasks; (b) helping the participants understand how to play both games simultaneously; (c) allowing the participants to experience the mode-switching decision. After the training stage, the participants experienced a seven-minute formal testing stage. After the experiment, as a manipulation check, I measured participants' perceived task uncertainty. The measures of perceived task uncertainty were adapted from Van de Ven and Delbecq (1974), Power et al. (2017), and Power et al. (2018).

Table 4.3 Experimental process

Stage	IntelSys-Decides	Human-Decides	IntelSys-Advises
Training	Five minutes for the two tasks (IntelSys makes the switching decision).	Five minutes for the two tasks (participants make the switching decision).	Five minutes for the two tasks (IntelSys provides switching advice, and players make the switching decision).
Actual Experiment	Seven minutes for the two tasks (IntelSys makes the switching decision).	Seven minutes for the two tasks (participants make switch decision).	Seven minutes for the two tasks (participants make the switching decision and IntelSys provides the switching suggestion).

Participants

The experiment was approved by the Institutional Review Board (IRB) protocol of Georgia State University. I recruited 900 undergraduates from the Business School of a major university in China. According to Compeau et al. (2012), recruiting students as participants in this experiment is appropriate for this line of research because students are representative of game players. All participants were randomly assigned to one of six groups. The experiment lasted about 20 minutes, and each participant received 35 RMB (about 5 U.S. dollars) as compensation. Participants were randomly assigned to one of the six groups (low/high task uncertainty * three DMAs). The final sample included 435 males and 448 females.

4.4 Variables and Measurement

The dependent variable in this research is team performance, which is calculated as the ratio of the team score to the total score (Table 4.4). The team score refers to the summation of the shooting scores earned by the player (in the manual mode) and by the computer (in the automation mode) during the seven-minute formal testing stage. The total score consists of the score of the total hazards that appeared during the seven-minute formal testing stage. Importantly, whether in low or high task uncertainty scenarios, the total score that could be achieved is the same for all participants.

The independent variables include expertise, task uncertainty, and DMAs (Table 4.4). The measures of task uncertainty used in the manipulation check were adapted from Van de Ven and Delbecq (1974) and Power et al. (2018). I made some wording adjustments (e.g., game uncertainty) to fit with the game setting (Table 7). Expertise is measured by the player's game experience as low (novice) or high (expert) levels (Green and Bavelier 2003; Karle et al. 2010).

The covariates include gender and expertise recognition (Table 4.4). Gender is controlled because prior studies show its impact on game performance (Brown et al. 1997; Quaiser-Pohl et al. 2006). For example, Brown et al. (1997) found that males played significantly better than females in video games. Quaiser-Pohl et al. (2006) indicated that studies in game research converged to the consistent conclusion that males are significantly better than females in mental rotation, which is found to have a significant positive impact on game playing performance. Expertise recognition refers to the extent to which the team can accurately recognize each member's level of expertise. Prior studies have also found that expertise recognition may affect expertise coordination (Littlepage et al. 1995). A team could design an optimal team coordination strategy if the team members could accurately recognize one another's expertise (Littlepage and Mueller 1997). According to existing literature, the measure of expertise recognition is based on a comparison between perceived and actual expertise (Libby et al. 1987; Littlepage et al. 1995; Bazarova and Yuan, 2013). For example, Bazarova and Yuan (2013) measured expertise recognition as the difference between an expert's actual expertise ranking and his/her expertise rankings as perceived by other team members. Therefore, in Bazarova and Yuan (2013), a ***negative value of*** expertise recognition means that the expert's expertise is underestimated; a ***positive value of*** expertise recognition means that the expert's expertise is overestimated; a ***zero*** for expertise recognition means that the expert's expertise is accurately evaluated. In this study, I adopt the measurement of expertise recognition from Bazarova and Yuan (2013). The measurement and coding of expertise recognition can be found in Table 4.4.

Table 4.4 Variables and measures

Variable Type	Variables	Measurement/Coding	References
Dependent variable	Team performance	the ratio of the team score to the total score on the shooting game	Kaber and Riley (1999)
Independent Variables	Task Uncertainty	A seven-point Likert scale with six items was adapted from the literature to the current context. Item 1: Unpredictable random hazards were influencing my performance. Item 2: I was relying on chance in the game. Item 3: I could not predict where is the hazard comes out. Item 4: I could not predict which direction the hazard flies over. Item 5: To what extent do you think you have adequate info about the coming hazards to support you to make the left-or-right movement to shoot the hazards? Item 6: How difficult is it for you to get the necessary information about the direction of incoming hazards for your actions?	Van de Ven and Delbecq (1974); Power et al. (2018)
	Expertise	High expertise: Play action games at least four sessions/week, at least one hour per session, and for a period lasting at least six months. Low expertise: Has very little (or even no) game playing experience.	Green and Bavelier (2003)
	DMAs	0 - IntelSys-Decides, 1 - Human-Decides, 2 - IntelSys-Advises	NA
Covariates	Gender	0 - Female, 1 - Male	NA
	Expertise Recognition (ER)	The common measurement of expertise recognition is based on the comparison between perceived expertise (PE) and actual expertise (AE). A negative value of ER means that the participants underestimate their expertise, and a positive value means that participants overestimate their expertise. PE: Perceived Expertise, AE: Actual Expertise. ER = -2 if PE = Computer and AE = Self ER = -1 if PE = Computer and AE = Equal ER = 0 if PE = AE ER = 1 if PE = Self and AE = Equal ER = 2 if PE = Self and AE = Computer	Libby et al. (1987); Littlepage et al. (1995); Bonner (2004); Bazarova and Yuan (2013)

4.5 Pilot Study

To measure the baseline level for the IntelSys-Decides and IntelSys-Advises approaches, I conducted two pilot studies. The first pilot study aimed to measure the baseline-level parameters (average and standard deviation value) for both low and high task uncertainty conditions. The second pilot study aimed to determine the lower and higher bound baseline levels based on the parameters determined by the first pilot study. Both pilot studies included a five-minute training stage and a five-minute testing stage. The data collected from the testing stage were used to measure the baseline workload level.

4.5.1 Pilot study one

As discussed earlier, when engaging in both tasks simultaneously, the performance of the gauge-monitoring task will be inversely related to the participants' workload. IntelSys monitors the participants' performance on the gauge-monitoring task in a 30-seconds cycle. During the manual mode, if IntelSys detects that the participant's workload is too high (the performance of the gauge-monitoring task is too low), IntelSys will switch from the manual mode to the automation mode. During the automation mode, if IntelSys detects that the participant's workload is too low (the performance of the gauge-monitoring task is too high), IntelSys will switch from the automation mode to the manual mode (Table 4.5). To evaluate whether the workload is high or low, it is necessary to compare the performance of the secondary task against its baseline, which includes the lower bound and higher bound baselines. Because the experiment includes low task uncertainty and high task uncertainty settings, I measured the baseline for both versions.

Table 4.5 Triggering conditions for mode switching

Current Mode	Switch Criterion	IntelSys-Decides	IntelSys-Advises
Manual	Gauge performance < lower bound baseline	Manual → Automation	System advice: Manual → Automation
Automation	Gauge performance > upper bound baseline	Automation → Manual	System advice: Automation → Manual

In the five-minute testing stage, there are five manual automation mode cycles, in which I recorded and calculated 10 workload values, including five overload values (from the manual mode) and five underload values (from the automation mode). The mean and standard deviation calculated from these 10 workload values were then used as two parameters to calculate the lower bound and higher bound baselines

(Kaber and Riley, 1999; Clamann et al., 2002; Kaber, 2006). The results of the first pilot study are shown in Table 4.6.

Table 4.6 Pilot study one results

	N	Mean	SD
Low Uncertainty	28	0.62	0.40
High Uncertainty	28	0.72	0.35

SD: The standard deviation of the 10 workload values from the five-minute testing stage.
Mean: The mean value of the 10 workload values from the five-minute testing stage.

4.5.1 Pilot study two

Following the procedure by Kaber (2002) and Kaber et al. (2006), after I obtained the baseline parameters (mean and SD of the performance of the gauge-monitoring task) for the low and high task uncertainty games, I conducted the following steps to determine the lower and higher bound baseline levels (N indicates how many times for SD. The workload criterion is identified by setting different values for N):

- Lower bound baseline = Mean – N * SD
- Higher bound baseline = Mean + N * SD
- Set three different values for N: 0.33, 0.67, and 1.0. The goal is to investigate the participant's workload under the different values of N and then choose the N that leads to the lowest level of workload for the whole group (Clamann et al. 2002). The workload is measured by the NASA-TLX scale, which has been used extensively to measure individuals' workload (Hart 2006).
- Each value of N corresponds with a pair of lower bound and higher bound baselines. Therefore, both low task uncertainty and high task uncertainty games have three different sets of baseline levels, respectively. In total, six groups (3 (values of N) * 2 (low and high uncertainty)) are needed to test the overall workload.
- The pilot study was conducted with six groups.

- For each group, the participant's workload was measured by the NASA-TLX scale.
- For both low and high task uncertainty games, the N with the lowest workload was chosen respectively.

Once I had the parameters for the baseline (Table 4.6), I calculated the lower bound and higher bound baselines for the six groups (Table 4.7).

Table 4.7 Pilot study two settings

Uncertainty	N	Low Bound	High Bound
Low	0.33	0.49	0.75
	0.67	0.35	0.89
	1.0	0.22	1.0
High	0.33	0.61	0.84
	0.67	0.49	0.95
	1.0	0.38	1.0

The goal of the second pilot study was to choose the best N that corresponds to the lowest workload. The results of pilot study two are shown in Table 4.8. I conducted ANOVA for low and high task uncertainty settings. The results suggested that for both the low and high task uncertainty settings, there was no significant difference in terms of the workload for the three different values of N. Theoretically speaking, any values of N are valid in determining the lower bound and higher bound baselines for both low and high task uncertainty settings. Therefore, for the low task uncertainty game, I chose N = 0.67. For the high task uncertainty game, I chose N = 0.67.

Table 4.8 Pilot study two results

Uncertainty	N	Sample Size	Workload	ANOVA
Low	0.33	31	59.6	Non-sig
	0.67	31	63.4	
	1.0	30	65.0	
High	0.33	30	64.7	Non-sig
	0.67	31	59.3	
	1.0	29	56.0	
Total	184			

Based on the results in Table 4.8, I developed the switch criteria for IntelSys in Table 4.9. For the low task uncertainty game, the lower bound baseline is 0.35, and the higher bound baseline is 0.89. In the formal experiment, if the participant is in the manual mode, and the performance of the gauge-monitoring task (30-second record) is lower than 0.35, IntelSys will switch to the automation mode (in the IntelSys-Decides approach) or will prompt the following text: "*advice: switch to automation mode. Press 'y' to accept, 'n' to reject*" (in the IntelSys-Advises approach). When the shooting task is in the automation mode, if the participant's performance on the gauge-monitoring task (30-second record) is higher than 0.89, IntelSys will switch to the manual mode (in the IntelSys-Decides approach) or will prompt the following text: "*advice: switch to manual mode. Press 'y' to accept, 'n' to reject*" (in the IntelSys-Advises approach).

For the high task uncertainty game, the lower bound baseline is 0.49, and the higher bound baseline is 0.95. In the formal experiment, when the shooting task is in the manual mode, if the participant's performance on the gauge-monitoring task (30-second record) is lower than 0.49, IntelSys will switch to the automation mode (in the IntelSys-Decides approach) or will prompt the following text: "*advice: switch to automation mode. Press 'y' to accept, 'n' to reject*" (in the IntelSys-Advises approach). When the shooting task is in the automation mode, if the player's performance on the gauge-monitoring task (30-second record) is higher than 0.95, IntelSys will switch to the manual mode (in the IntelSys-Decides approach) or will prompt the following text: "*advice: switch to manual mode. Press 'y' to accept, 'n' to reject*" (in the IntelSys-Advises approach).

Table 4.9 Switch criteria for IntelSys

Uncertainty	Baseline Levels	IntelSys-Decides	IntelSys-Advises
Low	Gauge performance < 0.35	Manual -> Auto	System suggestion: Manual -> Auto
	Gauge performance > 0.89	Auto - > Manual	System suggestion: Auto - > Manual
High	Gauge performance < 0.49	Manual -> Auto	System suggestion: Manual -> Auto
	Gauge performance > 0.95	Auto - > Manual	System suggestion: Auto - > Manual

5 Results

In this chapter, before showing the results, I introduce the dataset of the experiment and conduct descriptive analyses about the dataset. For the results, I first present the main findings based on the research model and the proposed hypotheses; that is, I compare the team performance of three DMAs under each boundary condition (e.g., low task uncertainty and high expertise, in Section 5.1). To explore more findings from the data, I conduct additional analyses on team performance, that is, for each DMA, comparing the team performance under different boundary conditions (in Section 5.2).

I also conduct a series of post-hoc analyses. The automation ratio, as a measure of the amount of automation used during the testing stage of the experiment, is found to have an impact on team performance. Therefore, I conduct two ANCOVA analyses in terms of the automation ratio, that is (a) for each boundary condition, comparing the automation ratio of the three DMAs (in section 5.3.1); and (b) for each DMA, I compare the automation ratio under different boundary conditions (in Section 5.3.2). Due to the possible mediation role of the automation ratio, I then conduct moderated-mediation analyses for each DMA (in

Section 5.3.4). At the end of this chapter, I investigate more to study how the participants responded to the advice of IntelSys in the IntelSys-Advises approach (section 5.3.5). All analyses are listed in the following Table 5.1.

Table 5.1 Structure of Chapter 5

	Section	Analyses
Main results	5.1	For each boundary condition, compare the **team performance** of three DMAs.
Additional analysis	5.2	For each DMA, compare the **team performance** under different boundary conditions.
Post-hoc analyses	5.3.1	For each boundary condition, compare the **automation ratio** of the three DMAs.
	5.3.2	For each DMA, compare the **automation ratio** under different boundary conditions.
	5.3.3	Summary of the post-hoc analysis: For each DMA, examine the automation ratio vs. team performance. For each boundary condition, examine the automation ratio vs. team performance.
	5.3.4	Moderated mediation analysis
	5.3.5	Participants' advice acceptance behavior in the IntelSys-Advises approach

I first conducted two pilot studies, which involved 240 participants. Then I conducted the formal experiment, in which 900 participants were randomly assigned to one of the six groups (2 [high vs. low] task uncertainty * 3 decision-making [IntelSys-Decides, Human-Decides, or IntelSys-Advises approaches]). Nineteen participants in the IntelSys-Advises group did not answer questions after the experiment and were therefore removed from the dataset. In total, I collected a sample of 881 valid records with 307 participants in the IntelSys-Decides group, 300 participants in the Human-Decides group, and 274 participants in the IntelSys-Advises group (Table 5.2). Table 5.3 lists the distribution of the sample across different levels of task uncertainty and decision-making approaches (DMAs). As shown in Table 5.3, for each DMA under different task uncertainty, the distribution of gender is quite similar, approximately 50% female and 50% male, which suggests effective randomization.

Table 5.2 Summary of the dataset

Dataset	Pilot	IntelSys-Decides	Human-Decides	IntelSys-Advises	Total
Sample Size	240	307	300	274	1121

Table 5.3 Sample distribution across task uncertainty and DMAs

Task Uncertainty	IntelSys-Decides		Human-Decides		IntelSys-Advises	
	Female	Male	Female	Male	Female	Male
Low	78 (50.0%)	78 (50.0%)	72 (50.7%)	70 (49.3%)	68 (48.9%)	71 (51.1%)
High	71 (47.0%)	80 (53.0%)	91 (57.6%)	67 (42.4%)	68 (50.4%)	67 (49.6%)

As discussed earlier in the experimental design (Section 4.2.3), I measured the participants' level of expertise through the survey rather than manipulating their expertise. Specifically, before the experiment, the participants were asked to answer questions about their prior experience in playing action-oriented games. Only participants who responded that they played action games (1) at least four sessions per week, (2) at least one hour per session, and (3) lasting for a period of at least six months were categorized as high expertise participants or expert gamers (Green and Bavelier 2003). Participants who played action games less than two sessions per week, less than or equal to 30 minutes per session, and lasting for a period of less than one month were categorized as low expertise participants or novice gamers. In the end, 662 valid records with low and high expertise were selected and used for the following analyses.

Table 5.4 lists the distribution of the sample across different levels of expertise, task uncertainty, and DMAs. For each group, there is a noteworthy difference between the number of female and male participants in terms of their experience in playing action games. Players with high expertise include more males than females, whereas players with low expertise include more females than males. In the low expertise groups, the number of female participants is approximately three times the number of male participants. In the high expertise groups, the number of male participants is approximately three times the number of female participants. Next, the ratio of female and male participants is relatively consistent across the IntelSys-Decides, Human-Decides, and IntelSys-Advises conditions. This distribution pattern is consistent with our expectation and prior literature that males generally have more experience than females in game playing (Brown et al. 1997; Quaiser-Pohl et al. 2006). The dataset suggests that (a) the randomization is effective, and (b) male and female participants have different but consistent distributions of action game expertise across different situations.

Table 5.4 Sample distribution across expertise, task uncertainty, and DMAs

Task Uncertainty	Expertise	IntelSys-Decides			Human-Decides			IntelSys-Advises		
		Total	Female	Male	Total	Female	Male	Total	Female	Male
Low	Low	63	49 (77.8%)	14 (22.2%)	59	47 (79.7%)	12 (20.3%)	62	42 (67.7%)	20 (32.3%)
Low	High	66	15 (22.7%)	51 (77.3%)	60	14 (23.3%)	46 (76.7%)	58	16 (27.6%)	42 (72.4%)
High	Low	52	39 (75.0%)	13 (25%)	53	44 (83.0%)	9 (17.0%)	44	38 (86.4%)	6 (13.6%)
High	High	51	5 (9.8%)	46 (90.2%)	45	9 (20.0%)	36 (80.0%)	49	12 (24.5%)	37 (75.5%)

I then conducted manipulation checks by measuring the perceived task uncertainty under low and high task uncertainty conditions. The results (3.465 vs. 5.220, $p < 0.000$) suggest that the manipulation of task uncertainty in this experiment was effective.

5.1 Main Results

To analyze the behavioral and performance data, I employ Analysis of Covariance (ANCOVA) to test if there are statistical differences as hypothesized. As mentioned in Section 4.4 of Chapter 4, the dependent variable is team performance, the independent variables include task uncertainty and expertise, and the covariates include gender and expertise recognition. Before delving into the detailed results below, I first briefly describe the overall results, as displayed in Table 5.5.

Table 5.5 Team performance results

Task Uncertainty / Human Expertise	Low	High

Low	Cell 1 (Full Support) ***H1:*** Human-Decides > IntelSys-Advises > IntelSys-Decides ***Results:*** Human-Decides > IntelSys-Advises > IntelSys-Decides	Cell 3 (Not Supported) ***H3:*** Human-Decides > IntelSys-Advises > IntelSys-Decides ***Results:*** No significant difference
High	Cell 2 (Full Support) ***H2:*** Human-Decides > IntelSys-Advises > IntelSys-Decides ***Results:*** Human-Decides > IntelSys-Advises > IntelSys-Decides	Cell 4 (Partial Support) ***H4:*** IntelSys-Decides > IntelSys-Advises > Human-Decides ***Results:*** IntelSys-Decides > Human-Decides > IntelSys-Advises

To begin with, Hypotheses 1 and 2 proposed that, under the condition of low task uncertainty (including both Cell 1 and Cell 2 of Table 5.5), team performance will be the highest for Human-Decides, followed by IntelSys-Advises, and the lowest for IntelSys-Decides. Consistent with this expectation, the results suggest that the rank of performance is Human-Decides > IntelSys-Advises > IntelSys-Decides, thereby supporting Hypotheses 1 and 2 (Table 5.5).

Next, Hypothesis 3 proposes that for low expertise participants engaged in highly uncertain tasks, team performance will be the highest for IntelSys-Decides, followed by IntelSys-Advises, and the lowest for Human-Decides. However, the results reveal no significant difference in team performance across the three DMAs in this cell (Cell 3 in Table 5.5), thereby rejecting Hypothesis 3.

Finally, Hypothesis 4 proposes that for high expertise participants engaged in high uncertainty tasks (Cell 4 in Table 5.5), team performance will be the highest for IntelSys-Decides, followed by Human-Decides, and the lowest for IntelSys-Advises. The results in Cell 4 suggest that, consistent with my expectation, the team performance of IntelSys-Decides is indeed higher that of Human-Decides and IntelSys-Advises (as shown in Table 5.5); yet, contrary to my expectation, the performance of the Human-

Decide is higher than that of the IntelSys-Advise. The results, as a whole, provide partial support to Hypothesis 4. Below I discuss the results of each cell in further detail.

5.1.1 Results of cell one: Low task uncertainty low expertise

Figure 5.1a and Table 5.6a both display the results of team performance in the situation of low task uncertainty and low expertise (i.e., Cell 1 in Table 5.5). The results in Figure 5.1a, Table 5.6a, and Table 5.6b suggest a significant difference in team performance across the IntelSys-Decides, Human-Decides, and IntelSys-Advises approaches ($F = 17.048$, $p < 0.001$). Consistent with H1, the Human-Decides approach has the highest level of team performance (0.522), which is also significantly higher than that of the IntelSys-Advises approach (0.445) (Table 5.6a and Table 5.6c). The IntelSys-Decides approach has the lowest team performance (0.364) among all three DMAs (Table 5.6a and Table 5.6c). The results support Hypothesis 1 and suggest that under the condition of low task uncertainty and low human expertise, the ranking for the team performance of DMAs is: Human-Decides > IntelSys-Advises > IntelSys-Decides.

Table 5.6a Descriptive statistics of team performance in low task uncertainty low expertise (Cell 1)

Descriptive	N	Mean	SD
IntelSys-Decides	63	0.364	0.107
Human-Decides	59	0.522	0.105
IntelSys-Advises	62	0.445	0.115

Table 5.6b ANCOVA results of low task uncertainty and low expertise

ANCOVA	df	F	SS	R^2 (adj)	sig	Power	Effect Size (partial η^2)
	4	17.048	0.808	0.260	0.000	1.000	0.259
		Covariates: Gender, Expertise Recognition					

Table 5.6c Post-hoc results of low task uncertainty and low expertise

post-hoc (LSD)	Mean Difference	p-value

IntelSys-Decides vs. Human-Decides	-0.159	0.000***
IntelSys-Decides vs. IntelSys-Advises	-0.080	0.000***
Human-Decides vs. IntelSys-Advises	0.078	0.000***

*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)

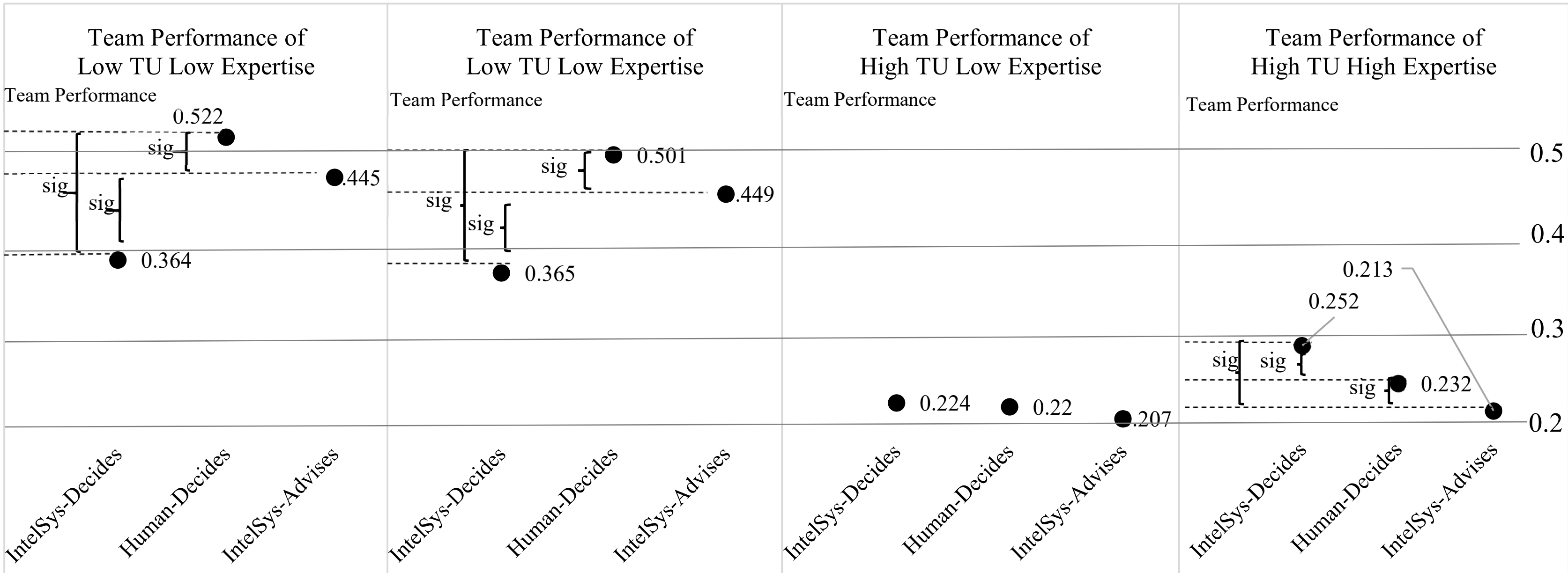


Figure 5.1a Team Performance of Low TU Low Expertise

Figure 5.1b Team Performance of Low TU High Expertise

Figure 5.1c Team Performance of High TU Low Expertise

Figure 5.1d Team Performance of High TU High Expertise

Figure 5.1 Within scenario comparison for team performance

Power refers to the probability of rejecting the null hypothesis. The power of this analysis is 1.00 (larger than 0.8), which indicates a 100% probability of finding a statistically significant difference when the difference actually exists (Table 5.6b). The effect size measures the magnitude of the experimental effect. The value of the effect size indicates the importance or meaningfulness of the observed difference among IntelSys-Decides, Human-Decides, and IntelSys-Advises. In this thesis, I adopt the effect size criterion of partial η^2 from Cohen et al. (2003, p 95). The effect sizes of 0.01, 0.09, and 0.25, respectively, represent a small, medium, and large effect (Cohen et al. 2003). The effect size of 0.259 in Model 1 suggests that the decision-making approach in the low task uncertainty and low expertise situation has a large effect on team performance (Table 5.6b). Regarding the covariates, neither gender nor expertise recognition has a significant impact on team performance (Table 5.6d).

Table 5.6d Effects of covariates on team performance

Covariates	**Coefficient**	**p-value**
Gender	0.000	0.993
Expertise Recognition	-0.044	0.056

5.1.2 Results of cell two: Low task uncertainty high expertise

Figure 5.1b and Table 5.7a illustrate the results of team performance in the scenario of low task uncertainty and high expertise (i.e., Cell 2 of Table 5.5). The results of Table 5.7b suggest that there is a significant difference in team performance across the IntelSys-Decides, Human-Decides, and IntelSys-Advises approaches ($F = 13.949$, $p < 0.001$). Consistent with H2, the Human-Decides approach has the highest level of team performance (0.501), followed by that of the IntelSys-Advises approach (0.449), which is then significantly higher than the team performance of the IntelSys-Decides approach (0.365) (Table 5.7a, 5.7c). These results collectively suggest that under the condition of low task uncertainty and high expertise, the ranking for the team performance of DMAs is: Human-Decides > IntelSys-Advises > IntelSys-Decides.

Table 5.7a Descriptive statistics of team performance in low task uncertainty high expertise

Descriptive	N	Mean	SD
IntelSys-Decides	66	0.365	0.102
Human-Decides	60	0.501	0.101
IntelSys-Advises	58	0.449	0.110

Meanwhile, the power of the above analyses is also 1.00 (larger than 0.8), suggesting a 100% probability of finding a statistically significant difference when there is one (Table 5.7b). The effect size of 0.233 suggests that the decision-making approach in the low task uncertainty and high expertise situation has a medium to large level effect on team performance (Table 5.7b). The effects of the covariates shown in Table 5.7d suggest that neither gender nor expertise recognition has a significant effect on team performance.

Table 5.7b ANCOVA results of low task uncertainty and high expertise

ANCOVA	df	F	SS	R^2 (adj)	sig	Power	Effect (partial η^2)	Size
	4	13.949	0.611	0.221	0.000	1.000	0.233	
		Covariates: Gender, Expertise Recognition						

Table 5.7c Post-hoc results of low task uncertainty and high expertise

Post-hoc (LSD)	Mean Difference	p-value
IntelSys-Decides vs. Human-Decides	-0.138	0.000^{***}
IntelSys-Decides vs. IntelSys-Advises	-0.084	0.000^{***}
Human-Decides vs. IntelSys-Advises	0.053	0.007^{**}

*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)

Table 5.7d Effects of covariates on team performance

Covariates	Coefficient	p-value
Gender	-0.020	0.279
Expertise Recognition	0.007	0.583

5.1.3 Results of cell three: High task uncertainty low expertise

Figure 5.1c, Table 5.8a, and 5.8b display the results of team performance in the high task uncertainty and low expertise situation (i.e., Cell 3 in Table 5.5). The results in Table 5.8b suggest that the ANCOVA model is statistically significant as a whole. However, the results in Figure 5.1c and Table 5.8c suggest that there is no significant difference across IntelSys-Decides, Human-Decides, and IntelSys-Advises, which is not consistent with Hypothesis 3.

Table 5.8a Descriptive statistics of team performance in high task uncertainty low expertise

Descriptive	N	Mean	SD
IntelSys-Decides	52	0.224	0.027
Human-Decides	53	0.220	0.059
IntelSys-Advises	44	0.207	0.053

Table 5.8b ANCOVA results of high task uncertainty and low expertise

ANCOVA	df	F	SS	R^2 (adj)	Sig	Power	Effect Size (partial η^2)
	4	2.821	0.025	0.047	0.027	0.759	0.015
Covariates: Gender, Expertise Recognition							

Table 5.8c Post-hoc results of high task uncertainty low expertise

Post-hoc (LSD)	Mean Difference	p-value
IntelSys-Decides vs. Human-Decides	0.001	0.952
IntelSys-Decides vs. IntelSys-Advises	0.013	0.176
Human-Decides vs. IntelSys-Advises	0.013	0.187

The power is 0.759 (lower than 0.8), which indicates a 75.9% probability of finding a statistically significant difference when there is one (Table 5.8b). The effect size of 0.015 is very small and suggests that under high task uncertainty and low expertise, the decision-making approach is not a key factor impacting team performance (Table 5.8b). The effects of the covariates shown in Table 5.8d suggest that neither gender nor expertise recognition has a significant effect on team performance.

Table 5.8d Effects of covariates on team performance

Covariates	Coefficient	p-value
Gender	0.012	0.145
Expertise Recognition	-0.001	0.887

5.1.4 Results of cell four: High task uncertainty high expertise

Figure 5.1d, Table 5.9a, and 5.9b show the results of team performance in the case of high task uncertainty and high expertise (i.e., Cell 4 in Table 5.5). The results suggest that there is a significant difference across the IntelSys-Decides, Human-Decides, and IntelSys-Advises approaches ($F = 7.148$, $p < 0.001$) in terms of team performance. As expected in Hypothesis 4, the IntelSys-Decides approach has the highest level of team performance (0.252), which is significantly higher than that of the Human-Decides approach (0.232) (p-value = 0.023, Table 5.9c) and the IntelSys-Advises approach (0.213) (p-value < 0.001, Table 5.9c). Yet, contrary to Hypothesis 4, Human-Decides is significantly higher than IntelSys-Advises in terms of team performance (p-value = 0.022, Table 5.9c). Hypothesis 4 is thus only partially supported.

Table 5.9a Descriptive statistics of team performance in high task uncertainty and high expertise

Descriptive	N	Mean	SD
IntelSys-Decides	51	0.252	0.023
Human-Decides	45	0.232	0.048
IntelSys-Advises	49	0.213	0.040

Table 5.9b ANCOVA results of high task uncertainty and high expertise

ANCOVA	df	F	SS	R^2 (adj)	sig	Power	Effect Size (partial η^2)
	4	7.148	0.041	0.146	0.000	0.995	0.142
		Covariates: Gender, Expertise Recognition					

Table 5.9c Post-hoc results of high task uncertainty and high expertise

Post-hoc (LSD)	Mean Difference	p-value
IntelSys-Decides vs. Human-Decides	0.018	0.023*
IntelSys-Decides vs. IntelSys-Advises	0.037	0.000***
Human-Decides vs. IntelSys-Advises	0.018	0.022*

*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)

The above results provide partial support to Hypothesis 4, which indicates that under high task uncertainty situations, experts would “seek more control,” and therefore take on a greater share of the task into his or her own hands rather than allocating the task to automation. Under this circumstance, experts would make an irrational decision because the decision is not based on the expertise coordination principle. In this vein, the decision made by IntelSys in the IntelSys-Decides approach would be better than the decision made by an expert in Human-Decides and IntelSys-Advises approaches. Therefore, the results seem to support my earlier argument that the team performance of IntelSys-Decides is higher than that of the Human-Decides or IntelSys-Advises approach.

On the other hand, Hypothesis 4 also expects that the team performance of Human-Decides will be higher than that of IntelSys-Advises. The results do not support this part of Hypothesis 4. I expected that in the IntelSys-Advises approach, the advice from the IntelSys would be beneficial to the participants; as such, the team performance of IntelSys-Advises should be higher than that of Human-Decides. However, the results show the opposite pattern. To identify a possible explanation, I conducted further exploration in terms of post-hoc analyses (section 5.3.2).

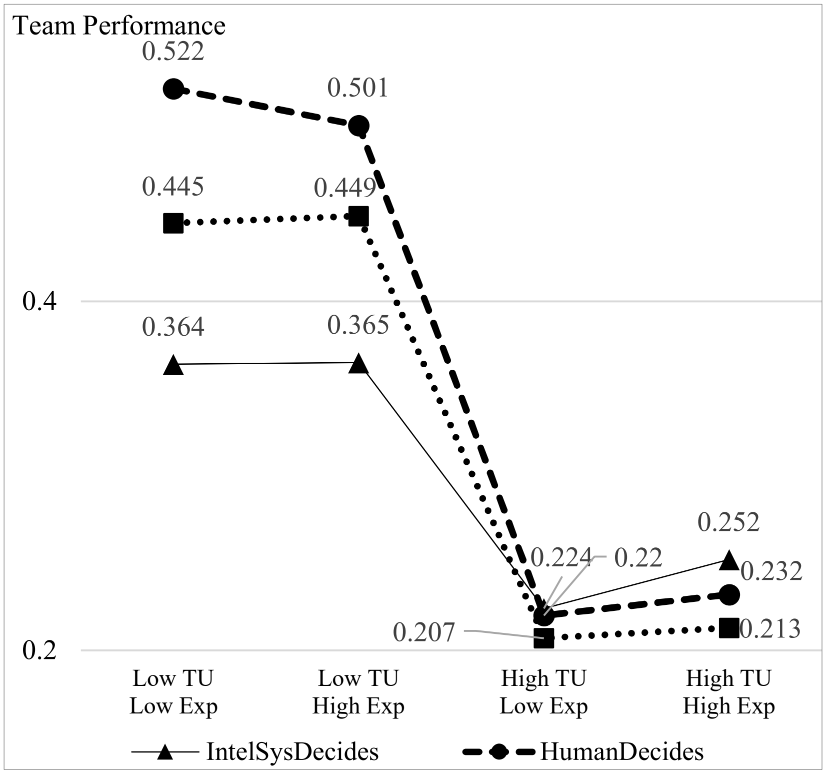
The power of this analysis is 0.995 (larger than 0.8), which indicates a 99.5% probability of finding a statistically significant difference when there is one (Table 5.9b). The effect size of 0.142 suggests that the decision-making approach in the high task uncertainty and high expertise situation has a medium to large level effect on the team performance (Table 5.9b). The effects of the covariates in Table 5.9d suggest that neither gender nor expertise recognition has a significant effect on team performance.

Table 5.9d Effects of covariates on team performance

Covariates	Coefficient	p-value
Gender	0.012	0.145
Expertise Recognition	-0.001	0.887

Figure 5.2 shows the overall results of team performance across different situations. The rankings of team performance in low task uncertainty situations are consistent for participants with different levels

of expertise. On the contrary, the rankings of team performance in high task uncertainty situations are very different for participants with different levels of expertise. In order to explore additional findings with respect to team performance, I then compare the team performance of each DMA under different situations.



(TU = Task Uncertainty; Exp = Expertise)

Figure 5.2 Overall results of team performance across different scenarios

5.2 Additional Analysis of Team Performance (within DMA Comparison)

Next, I assess the team performance of each DMA under different boundary conditions by conducting a series of ANCOVA analyses. The goal of these additional analyses is to study the possible impacts of task uncertainty and expertise on team performance for each DMA. The results are presented in Figure 5.3.

Team performance of the IntelSys-Decides approach. For the IntelSys-Decides approach, the ANCOVA results show that the whole model is significant (Table 5.10, Figure 5.3a; $F = 29.376$, $df = 5$, $p < 0.001$). The post-hoc comparisons suggest that under high task uncertainty situations, participants with high expertise attain higher team performance than participants with low expertise (Comparison 2, p =

0.058; albeit marginally significant). However, under the low task uncertainty situations, there is no significant difference in team performance between participants with high and low expertise (Comparison l, $p = 0.657$). The results of Comparison 3 ($p < 0.001$) and Comparison 4 ($p < 0.001$) of the IntelSys-Decides approach in Table 5.10 suggest that for all participants, no matter if they are experts or novices, high task uncertainty always leads to lower team performance. One possible reason is that task uncertainty negatively impacts team performance.

Table 5.10 ANCOVA results of team performance (within DMAs Comparison)

		IntelSys-Decides		Human-Decides		IntelSys-Advises	
Compare	**ANCOVA**	**MD**	**p-value**	**MD**	**p-value**	**MD**	**p-value**
1	low TU low expertise vs. low TU high expertise	-0.007	0.657	0.031	0.080^{+}	-0.006	0.747
2	high TU low expertise vs. high TU high expertise	-0.035	0.058^{+}	0.001	0.953	-0.009	0.681
3	low TU low expertise vs. high TU low expertise	0.140	0.000^{***}	0.300	0.000^{***}	0.238	0.000^{***}
4	low TU high expertise vs. high TU high expertise	0.112	0.000^{***}	0.270	0.000^{***}	0.235	0.000^{***}
DMA	**F-value**	Df	SS	R^2 (adj)	p-value	power	Effect size (partial η^2)
IntelSys-Decides	29.376	5	0.946	0.381	0.000^{***}	1.000	0.394
Human-Decides	123.827	5	4.438	0.740	0.000^{***}	1.000	0.746
IntelSys-Advises	72.176	5	2.947	0.627	0.000^{***}	1.000	0.635

Note: MD = Mean Difference; TU = Task Uncertainty; SS = Sum of Squares
*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)

Table 5.11 Effects of covariates on team performance

	IntelSys-Decides		Human-Decides		IntelSys-Advises	
Covariates	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Gender	-0.010	0.986	0.020	0.163	-0.004	0.773
Expertise Recognition	0.000	0.457	-0.009	0.371	-0.003	0.733

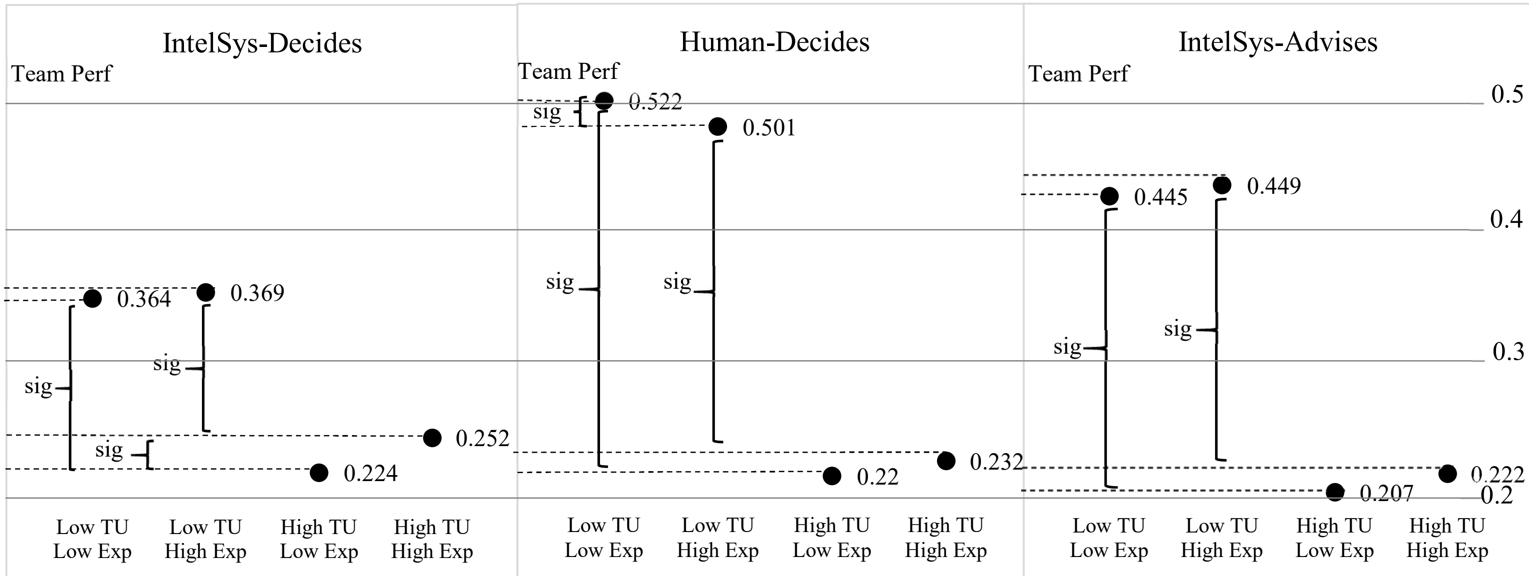


Figure 5.3a Team Performance of IntelSys-Decides

Figure 5.3b Team Performance of IntelSys-Decides

Figure 5.3c Team Performance of IntelSys-Advises

(TU = Task Uncertainty; Exp = Expertise; Perf = Performance; sig = Significant Difference)

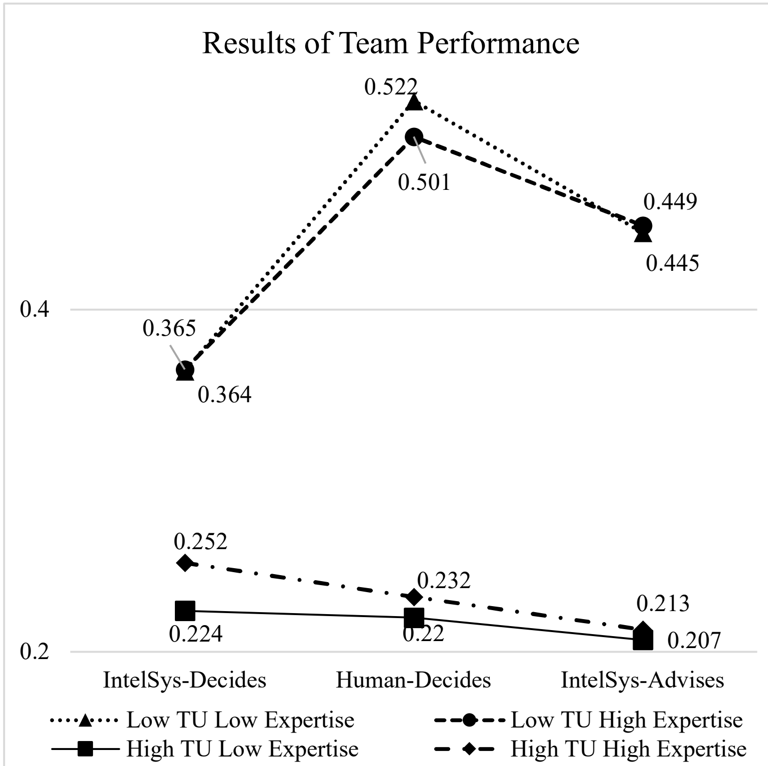
Figure 5.3 Within DMA comparison for team performance

Team performance of the Human-Decides approach. For the Human-Decides approach, the ANCOVA results show that the whole model is significant (Table 5.10, figure 5.3b; F = 123.827, df = 5, p < 0.001). The result of Comparison 1 in Table 5.10 suggests that in the situation of low uncertainty, participants with low expertise attain higher team performance than participants with high expertise (p = 0.080, albeit marginally significant). However, in the situation of high task uncertainty, there is no significant difference in team performance between participants with low and high expertise (Comparison 2, p = 0.953). Similar to the results of the IntelSys-Decides approach, the results of Comparisons 3 (p < 0.001) and 4 (p < 0.001) of the Human-Decides approach in Table 5.10 suggest that for all participants, no matter if they are experts or novices, high task uncertainty always leads to low team performance. Again, task uncertainty seems to negatively affect team performance.

Team performance of the IntelSys-Advises approach. For the IntelSys-Advises approach, the ANCOVA results show that the whole model is significant (Table 5.10, figure 5.3c; F = 72.176, df = 5, p < 0.001). The results of Comparison 1 (p = 0.747) and Comparison 2 (p = 0.681) suggest that there is no significant impact from expertise on team performance. Consistent with the results of the IntelSys-Decides and Human-Decides approaches, the results of Comparison 3 (p = 0.000) and Comparison 4 (p < 0.001) of the IntelSys-Advises approach in Table 5.10 suggest that task uncertainty negatively impacts team performance.

Finally, the results of Table 5.11 suggest that for all three DMAs, the covariates have no significant impact on team performance.

Summary of findings. To recap, Figure 5.4 shows the overall results of each DMA under different situations. The aforementioned results suggest that: (1) the impact of expertise on team performance is very different across different levels of task uncertainty and DMAs (Comparisons 1 and 2 in Table 5.10), suggesting the possible interaction effect between expertise and task uncertainty, or the possible interaction effect between expertise and DMAs; and (2) the impact of task uncertainty on team performance is consistently negative across the three DMAs (Comparisons 3 and 4 in Table 5.10).



(TU = task uncertainty)

Figure 5.4 The overall results of team performance across different DMAs

5.3 Post-hoc Analyses

I further conduct numerous post-hoc analyses to explore how participants and IntelSys, respectively, make their task-allocation decisions. Given that the task-allocation decision refers to how tasks are allotted between humans and automation, I examine the automation ratio (or manual ratio, which is the opposite of the automation ratio) to reflect the task-allocation decision. In this dissertation, the automation ratio is defined as the ratio of the amount of time in the automation mode during the formal seven-minute testing stage. For instance, if a participant spends three minutes and 30 seconds using the automation mode during the formal testing stage, the automation ratio equals [(3 minutes x 60 seconds/per minute) +30 seconds] divided by [(7 minutes x 60 seconds/per minute)] = 210 seconds / 420 seconds = 0.5. In this vein, the manual ratio is one minus the automation ratio because the sum of these two ratios should be 100%.

In the Human-Decides and IntelSys-Advises approaches, the automation ratio thus indicates the participant's control behavior. A high automation ratio means that the participant prefers to use automation, while a low automation ratio suggests that the participant prefers to rely more on himself or herself for task execution. Probing into the variation of the automation ratio of each DMA under different situations offers another way to observe how the participants (in the Human-Decides and IntelSys-Advises approaches) and IntelSys make decisions (in the IntelSys-Decides approach).

Impact of the automation ratio on team performance

It is worth noting that for studies in human-automation interaction, the assumption is that the task performance of automation is higher than the performance of most humans; this is because automation is designed to outperform the majority of humans. Consistent with this assumption, for both low and high task uncertainty situations, we observe in Figure 5.5a (Low Task Uncertainty) and 5.5b (High Task Uncertainty) that the performance of automation is indeed higher than the majority of participants' performance. Hence, the automation ratio illustrated in the formal testing stage should be positively related to team performance.

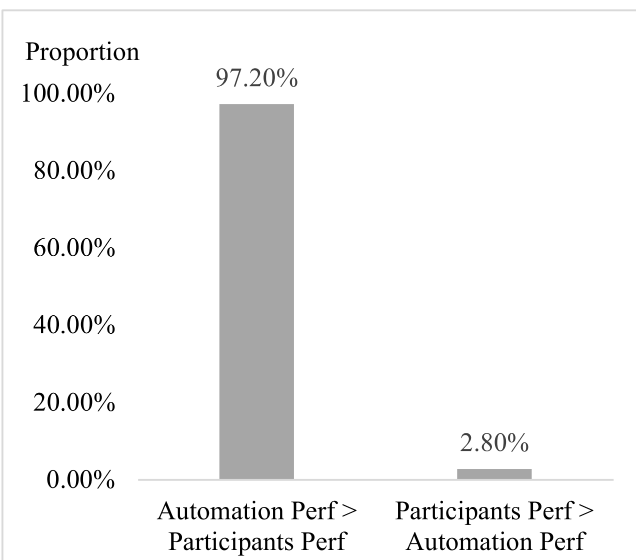


Figure 5.5a Distribution of automation/participants performance (Perf = Performance; Low Task Uncertainty)

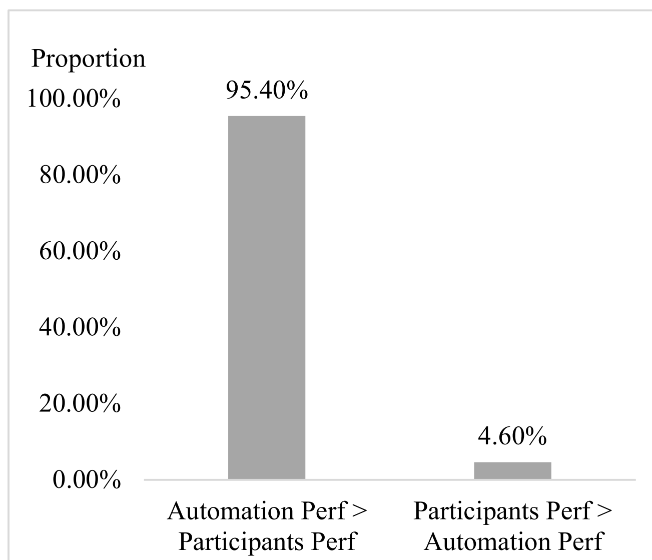


Figure 5.5b Distribution of automation/participants performance (Perf = Performance; High Task Uncertainty)

Figure 5.5 Distribution of automation/participants' performance under low/high task uncertainty

Impact of task uncertainty and expertise on the automation ratio

Given the importance of the automation ratio on team performance, I further explore the factors that may impact the automation ratio. As indicated earlier, the automation ratio represents the participants' control behavior. A high automation ratio reflects participants allocating a lower amount of the task to the manual mode, suggesting that they do not want to seek more control. In contrast, a low automation ratio means that participants who allocate more tasks to the manual mode are more willing to seek additional control.

According to Parasuraman et al. (2007), compared with novices, experts are more confident about their expertise, and therefore are likely to rely more on themselves. As a result, expertise would negatively impact the automation ratio. Moreover, based on the personal control model mentioned in Section 2.3.4, task uncertainty has a negative impact on the automation ratio because (1) a task with low uncertainty has no impact on the human's control behavior (Table 5.12); and (2) humans are more likely to "seek more control" under high task uncertainty situations (Greenberger and Strasser 1986).

Table 5.12 Moderating effect of expertise on task uncertainty

Task Uncertainty	Personal Control Model	Expected Control Behavior
Low	No impact on the individual's control behavior	Use more automation mode than manual mode
High	Possessed control < Desired control - *Learned helplessness*	***Novice****:* Learned helplessness→ Use more automation than under low task uncertainty situations
	Possessed control < Desired control - *Seek more control*	***Expert****:* Seek more control → Use less automation than under low task uncertainty situations

Note: The personal control model is adopted from Greenberger and Strasser (1986).

Based on the personal control model, low task uncertainty may have no impact on humans' control behavior. As mentioned in Section 3.3, with low task uncertainty, humans are more likely to make task-allocation decisions based on the expertise coordination principle, which means that humans would allocate a greater share of the task to team members who have better task expertise (i.e., automation in the human-automation team). Because automation plays better than most participants in the experiment, it is arguable that the participants, no matter if they are novices or experts, are more likely to use the automation mode to a greater extent than the manual mode (i.e., high automation ratio, Table 5.12) in the case of low task uncertainty. For the situation of high task uncertainty, novices would tend to experience learned helplessness, realizing that the automation performs much better than they do. Therefore, novices are more likely to exhibit a high automation ratio (Table 5.12). In contrast, experts under high task uncertainty situations are more likely to seek more control and are less likely to experience learned helplessness; thus, they are more likely to exhibit a low automation ratio (Table 5.12).

To investigate the automation ratio of DMAs under different situations, below I first compare the automation ratio of the three DMAs across different scenarios (high or low task uncertainty versus high or low expertise) (Section 5.3.1), followed by the assessment of the impact of task uncertainty and expertise on the automation ratio for each DMA (Section 5.3.2).

5.3.1 Automation ratio results of DMAs (within scenario comparison)

As mentioned in Section 3, there are two situations under which the human's decisions would be better than the decisions of IntelSys: (a) the condition of a low task uncertainty situation in which humans can make flexible, personalized, and informed decisions; and (b) the condition of high task uncertainty in which low expertise participants are more likely to realize that the automation plays much better than they do; therefore, they are more likely to experience learned helplessness (Table 5.13). In both conditions, humans' decisions are more likely to follow the expertise coordination principle than the decisions of IntelSys. Given that automation plays better than most participants, the decisions that follow the expertise coordination principle would tend to use more automation than playing by oneself.

As discussed in Section 3, the only situation in which IntelSys would make a better decision than the human is under the condition of high task uncertainty and high human expertise. In this situation (i.e., high task uncertainty and high expertise), humans have a strong desire to seek more control; therefore, they would tend to allocate a greater share of the tasks to themselves and would use the automation mode to a lesser extent than IntelSys (Table 5.12). It appears that the interaction between task uncertainty and expertise may have influenced how the participants used automation during the experiment. Given the above, I compare the automation ratio of the DMAs in each scenario to assess the human- and IntelSys-made decisions during the formal test of the experiment.

Table 5.13 IntelSys vs. Human in the automation ratio

	Human vs. IntelSys	Actual Ranking of the Automation Ratio
Low TU	Humans make rational decisions and use more automation than IntelSys.	Novice: Human-Decides > IntelSys-Advises > IntelSys-Decides Expert: Human-Decides > IntelSys-Decides For experts, no significant difference exists between Human-Decides and IntelSys-Advises; For experts, no significant difference exists between IntelSys-Decides and IntelSys-Advises.
High TU Low Expertise	Learned helplessness: Humans allocate a greater share of the tasks to automation; Humans make rational decisions and have a ***higher*** automation ratio than IntelSys.	No significant difference exists between IntelSys-Decides, Human-Decides, and IntelSys-Advises.
High TU High Expertise	Seek more control: Humans allocate a great share of the tasks to themselves; Humans make irrational decisions and have a ***lower*** automation ratio than IntelSys.	IntelSys-Decides > IntelSys-Advises No difference exists between IntelSys-Decides and Human-Decides; Human-Decides > IntelSys-Advises

TU = Task Uncertainty

In particular, I conduct a series of ANCOVA analyses to explore the effect of task uncertainty and expertise on the automation ratio. Table 5.14 shows the descriptive results of the automation ratio. Figure 5.6 and Table 5.15 show the ANCOVA results. The results of pairwise comparisons in Model 1 (ANCOVA of low task uncertainty and low expertise) suggest that the rank of the automation ratio is: Human-Decides > IntelSys-Advises > IntelSys-Decides (Figure 5.6 and Table 5.15). In the IntelSys-Decides approach, 48.3% of the time was allocated to automation, whereas in the Human-Decides approach, 78.2% of the testing stage was dominated by automation. Finally, in the IntelSys-Advises approach, 63.3% of the testing stage was in the automation mode (Table 5.13). In short, participants with low expertise under low task uncertainty allocated a greater share of the tasks to automation than IntelSys did.

Table 5.14 Descriptive statistics of the automation ratio

Different Situations	IntelSys-Decides			Human-Decides			IntelSys-Advises		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**
Low TU Low Expertise	63	0.483	0.280	59	0.782	0.296	62	0.633	0.299
Low TU High Expertise	66	0.463	0.318	60	0.654	0.337	58	0.563	0.328
High TU Low Expertise	52	0.505	0.040	53	0.528	0.363	44	0.518	0.299
High TU High Expertise	53	0.500	0.000	45	0.468	0.399	49	0.344	0.265

Table 5.15 Automation ratio of DMAs across different situations

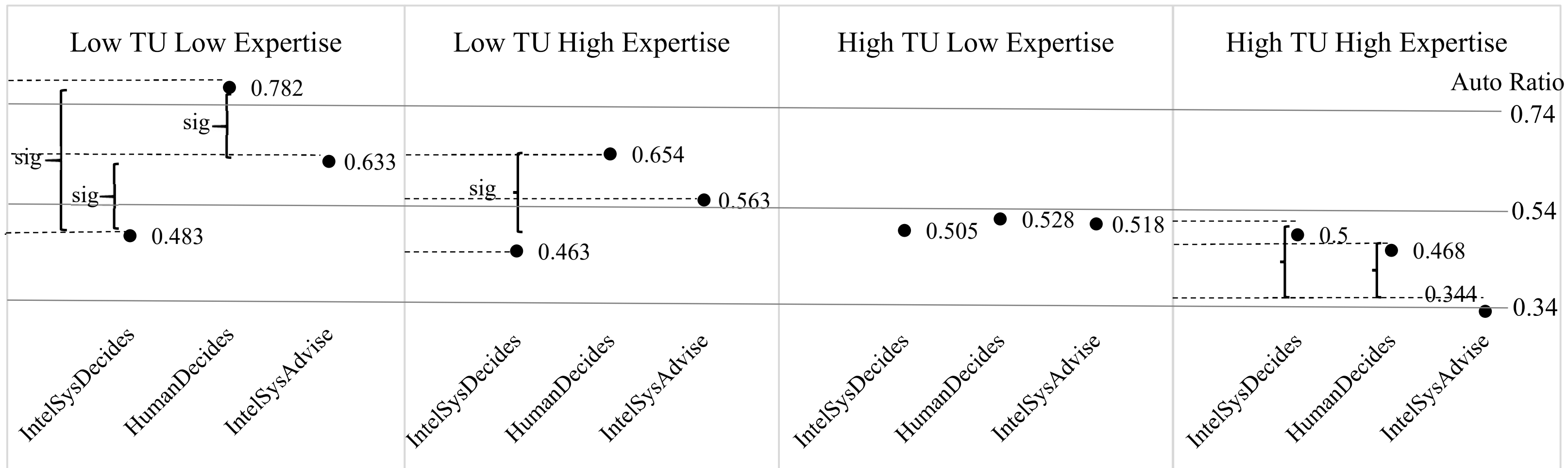
Model	Comparison	MD	p-value
Low TU Low Expertise (Model 1)	IntelSys-Decides vs. Human-Decides	-0.291	0.000***
	IntelSys-Decides vs. IntelSys-Advises	-0.141	0.007**
	Human-Decides vs. IntelSys-Advises	0.150	0.005**
F = 10.233, df = 4, SS = 3.364, R^2 (adj) = 0.168, p = 0.000 power = 1.000, effect size = 0.149			
Low TU High Expertise (Model 2)	IntelSys-Decides vs. Human-Decides	-0.178	0.003**
	IntelSys-Decides vs. IntelSys-Advises	-0.081	0.169
	Human-Decides vs. IntelSys-Advises	0.096	0.107
F = 4.510, df = 4, SS = 1.880, R^2 (adj) = 0.071, p = 0.002 power = 0.937, effect size = 0.050			
High TU Low Expertise (Model 3)	IntelSys-Decides vs. Human-Decides	-0.019	0.721
	IntelSys-Decides vs. IntelSys-Advises	-0.018	0.747
	Human-Decides vs. IntelSys-Advises	0.001	0.985
F = 1.972, df = 4, SS = 0.560, R^2 (adj) = 0.026, p = 0.102, power = 0.582, effect size = 0.001			
High TU High Expertise (Model 4)	IntelSys-Decides vs. Human-Decides	0.029	0.602
	IntelSys-Decides vs. IntelSys-Advises	0.143	0.010*
	Human-Decides vs. IntelSys-Advises	0.114	0.045*
F = 2.874, df = 4, SS = 0.837, R^2 (adj) = 0.049, p = 0.025, power = 0.767, effect size = 0.050			

Note: MD = Mean Difference; TU = Task Uncertainty; SS = Sum of Squares; effect size is partial η^2
For all models, the covariates include gender and expertise recognition.
*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)

From Model 2, the results suggest that in low task uncertainty and high expertise, the automation ratio is lower under the IntelSys-Decides approach (0.463) than the Human-Decides approach (0.654). Yet, the difference in the automation ratio between IntelSys-Decides and IntelSys-Advises (Δ = -0.081, p = 0.169) and the difference in the automation ratio between Human-Decides and IntelSys-Advises (Δ = 0.096, p = 0.107) are not significant. In the IntelSys-Decides approach, 46.3% of the time was allocated to automation, while in the Human-Decides approach, 65.4% of the time was allocated to automation; and in the IntelSys-Advises approach, 56.3% of the time was allocated to automation (Table 5.14). These results suggest that under the condition of low task uncertainty, participants with high expertise in the Human-Decides approach are more likely to make decisions that follow the expertise coordination principle than IntelSys.

The results of Model 3 suggest that in the situation of high task uncertainty and low expertise, the automation ratios in IntelSys-Decides (50.5%), Human-Decides (52.8%), and IntelSys-Advises (51.8%) are not significantly different (Table 5.14, Table 5.15). All DMAs used approximately 50% automation. It is very interesting to see that the participants in both the Human-Decides and IntelSys-Advises approaches used a similar amount of automation as IntelSys-Decides.

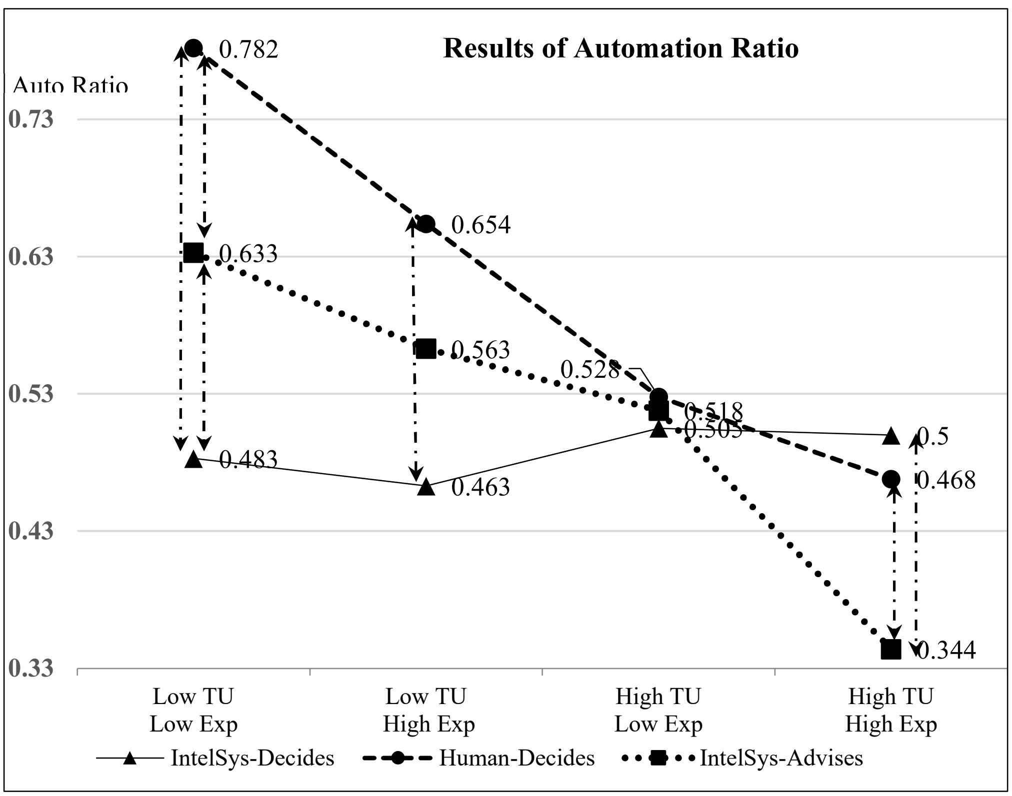
The results of Model 4 (Table 5.15) suggest that both the IntelSys-Decides (50.0%) and Human-Decides (46.8%) approaches used more automation than the IntelSys-Advises approach (34.4%). However, the difference in the automation ratio between IntelSys-Decides (50.0%) and Human-Decides (46.8%) is not significant (($\Delta = 0.029$, $p = 0.602$). The results of Model 4 suggest that the Human-Decides approach used significantly more automation than the IntelSys-Advises approach ($\Delta = 0.114$, $p = 0.045$) (Table 5.14). A possible explanation is that participants with high expertise in the IntelSys-Advises approach accepted many "switch to manual mode" advice prompts from IntelSys in the condition of high task uncertainty; it is also possible that these experts rejected many "switch to automation mode" advice prompts from IntelSys in the condition of high task uncertainty. In the subsequent Section 5.3.4, I further analyze exactly how participants accept or reject the advice from IntelSys.



(TU = Task Uncertainty; sig = Significant Difference)

Figure 5.6 Within scenario comparison for the automation ratio

Summary of results. Figure 5.7 shows the trend of the automation ratio across all situations. The automation ratio of IntelSys-Decides seems to be quite consistent across different situations, around 50.0%. Under the low task uncertainty situations, participants in both the Human-Decides and IntelSys-Advises approaches used significantly more automation than their counterparts in the IntelSys-Decides approach. Next, under the condition of high task uncertainty, participants with either low or high expertise used automation at a similar level as IntelSys in IntelSys-Decides (around 50.0%). However, in the case of high task uncertainty, participants with high expertise in the IntelSys-Advises approach show a significantly lower automation ratio (34.4%) than their counterparts in the IntelSys-Decides (50.0%) and Human-Decides approaches (46.8%) approaches. It appears that participants with high expertise in the IntelSys-Advises approach were more radical in taking control into their own hands than participants with high expertise in the Human-Decides and System-Decides approaches. A possible explanation is that participants with high expertise adopted many "switch to manual mode" advice prompts from IntelSys.



(TU = Task Uncertainty, Exp = Expertise; The dashed line with two arrows indicates the significant difference.)

Figure 5.7 The overall results of the automation ratio across different scenarios

5.3.2 Automation ratio results of DMAs (within DMA comparison)

Results of the IntelSys-Decides approach. In the IntelSys-Decides approach, all switching decisions were made by IntelSys. I thus ran an ANCOVA model to investigate how IntelSys makes switching decisions under different situations. The ANCOVA results show that the whole model is significant (Table 5.16; F = 2.316, df = 5, p = 0.044). The results of the post-hoc comparisons (Table 5.16) suggest that there is no significant difference in the automation ratio across different situations, indicating that task uncertainty and expertise have no impact on the automation ratio (Figure 5.8a).

Table 5.16 Automation ratio results of DMAs

		IntelSys-Decides		Human-Decides		IntelSys-Advises	
Comparison	**ANCOVA**	**MD**	**p-value**	**MD**	**p-value**	**MD**	**p-value**
1	Low TU Low Expertise vs. Low TU High Expertise	-0.021	0.635	0.143	0.048*	0.045	0.435
2	High TU Low Expertise vs. High TU High Expertise	-0.022	0.660	0.084	0.299	0.136	0.048*
3	Low TU Low Expertise vs. High TU Low Expertise	-0.034	0.412	0.243	0.000***	0.106	0.075^{+}
4	Low TU High Expertise vs. High TU High Expertise	-0.035	0.397	0.184	0.007**	0.197	0.001**

DMA	**F-value**	df	SS	R^2 (adj)	p-value	Power	Effect Size
IntelSys-Decides	2.316	5	0.534	0.025	0.044*	0.711	0.007
Human-Decides	5.865	5	3.507	0.101	0.000***	0.994	0.104
IntelSys-Advises	6.752	5	2.967	0.119	0.000***	0.998	0.081

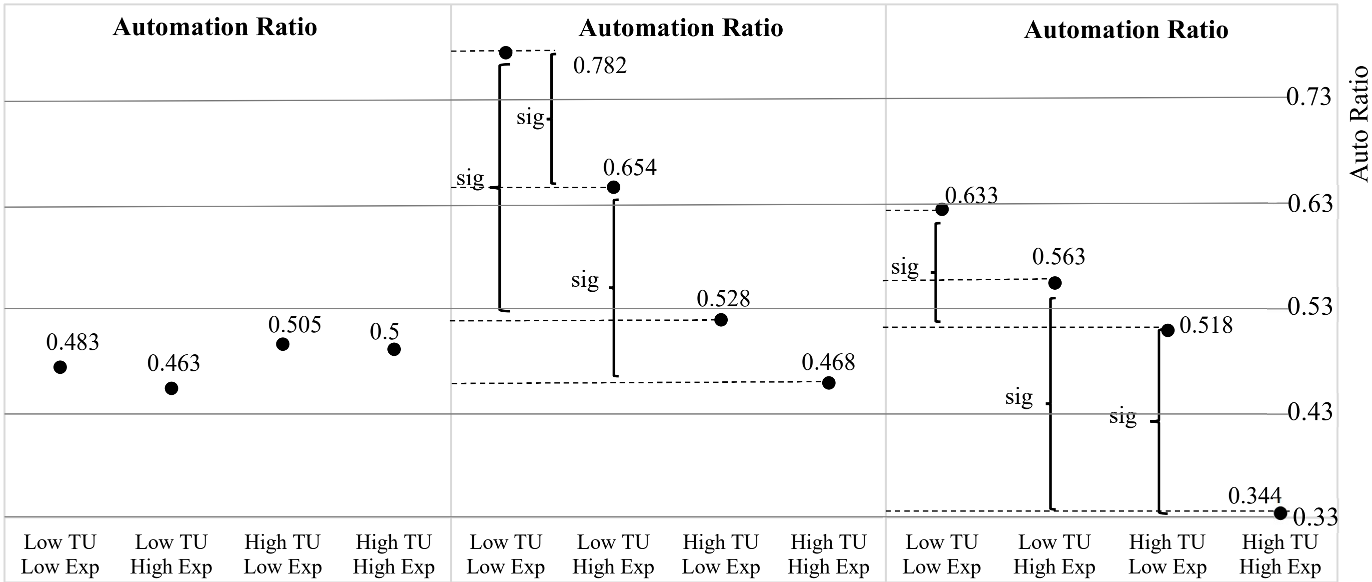
Note: MD = Mean Difference is based on estimated marginal means; TU = Task Uncertainty; SS = Sum of Squares; effect size= partial η2; *: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)

In Table 5.17, the covariate gender is not significant (coefficient = -0.046, p =0.209). However, the effect of expertise recognition on the automation ratio is significant (coefficient = -0.057, p =0.009). In this study, expertise recognition is measured by the difference between perceived and actual expertise (Bazarova and Yuan 2013). A negative value of expertise recognition means that participants underestimate their expertise when they actually play better than automation. In contrast, a positive value of expertise recognition means that participants overestimate their expertise when automation actually plays better than they do. A zero value of expertise recognition means that the participants accurately evaluate their own expertise relative to automation. Therefore, the value of expertise recognition represents how a participant perceives his or her expertise in relation to the expertise of automation. Put simply, a higher value of expertise recognition suggests a higher likelihood that the participant is positive about his or her expertise, as compared to automation.

Table 5.17 Effects of covariates on the automation ratio

	IntelSys-Decides		Human-Decides		IntelSys-Advises	
Covariates	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Gender	-0.046	0.212	0.041	0.497	-0.041	0.388
Expertise Recognition	-0.057	0.012	-0.074	0.081	-0.072	0.025

For the IntelSys-Decides approach, expertise recognition has a negative impact on the automation ratio (coefficient = -0.057, p =0.009, Table 5.17). One possible explanation is that participants who see their expertise as being equal to or better than automation (i.e., higher expertise recognition) would be more confident, and thus would devote themselves to the task, including both the gauge-monitoring and shooting tasks (i.e., manual mode) rather than relinquishing the task to automation. Given such devotion, the performance on the gauge-monitoring task is more likely to be high and reflect a low workload level in the manual mode. The low workload level can then cause IntelSys to keep the current manual mode rather than switch to the automation mode. For the IntelSys-Advises approach, expertise recognition has a negative impact on the automation ratio (coefficient = -0.072, p =0.025, Table 5.17), indicating that participants who are more positive about their own expertise are more likely to engage in the manual mode.



a. IntelSys-Decides b. Human-Decides c. IntelSys-Advises

(TU = Task Uncertainty, Exp = Expertise; sig = Significant Difference)

Figure 5.8 Within DMAs comparison for automation ratio

Results of the Human-Decides approach. I performed ANCOVA to assess the impact of task uncertainty and expertise on the automation ratio (F = 5.865, df = 5, p = 0.000, Table 5.16) in the Human-Decides approach. Table 5.16 and Figure 5.8b show the results of the automation ratio of the Human-Decides approach under different situations.

The results of Comparison 1 of Table 5.16 suggest that in the situation of low uncertainty, participants with low expertise (automation ratio = 0.782) exhibited a higher automation ratio than participants with high expertise (automation ratio = 0.654). Therefore, the experts took more control than novices (Δ= 0.143, p = 0.048, Comparison 1 of Table 5.16) in the case of low task uncertainty. This result is consistent with the expectation that experts, relative to novices, are more likely to work by themselves (Miller and Parasuraman 2007). However, under the condition of high task uncertainty, novices (automation ratio = 0.528) and experts (automation ratio = 0.468) did not show a significant difference in the automation ratio (Δ= 0.084, p = 0.299, comparison 2 of Table 5.16). Also as expected, experts took significantly more control (Δ= 0.184, p = 0.007, Comparison 4 of Table 5.16) in the high task uncertainty situations (automation ratio = 0.468) than in the low task uncertainty situations (automation ratio = 0.654). An interesting finding is that novices used less automation in high task uncertainty situations (automation ratio = 0.528) than in low task uncertainty situations (automation ratio = 0.782; p = 0.000, Comparison 3). This result is contrary to the expectation, which suggested that the control behavior of novices in high task uncertainty situations should exhibit "learned helplessness," and that they would most likely allocate a greater share of the task to automation (Table 5.13). However, it turns out that novices sought more control, as opposed to experiencing learned helplessness (Table 5.13).

As elaborated in Section 2.3, prior literature suggests that seeking more control in high uncertainty situations is indicative of human nature (Ashford and Black 1996; Mushtaq et al. 2011). According to the personal control model (Greenberger and Strasser 1986), under the situation of high task uncertainty, a human's control behavior is first to "seek more control." When the individual realizes that having more control has no significant impact on the task outcome, her or his control behavior will then evolve to behavior demonstrating "learned helplessness." Following this logic, one possible explanation regarding the results of Comparison 3 in Table 5.16 is that the novices' control behaviors were still in the first stage ("seek more control") and did not move to the second stage ("learned helplessness").

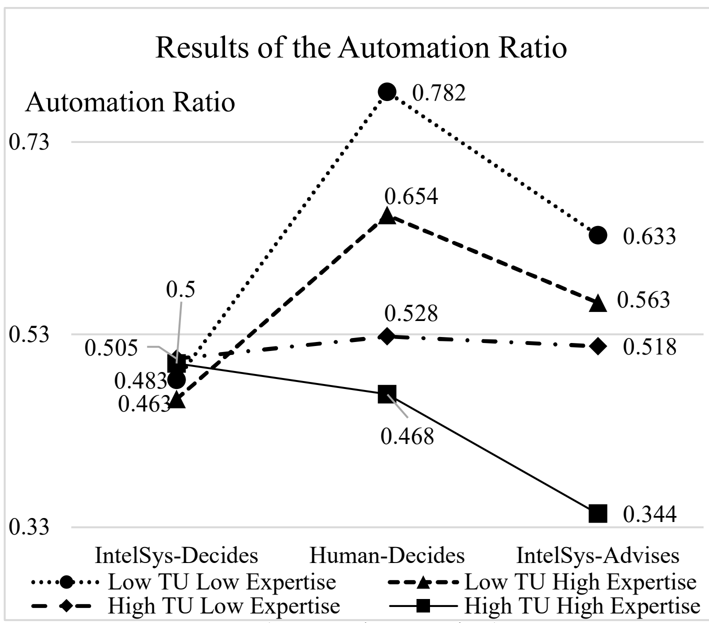
Based on the results of Comparisons 3 and 4, the control behaviors of novices under high task uncertainty are similar to the control behaviors of experts under the high task uncertainty situation. Both experts and novices made irrational decisions under high task uncertainty situations. The desire for more control, stimulated by high task uncertainty, seems to have dominated in all participants' decisions.

The results of Table 5.17 suggest that neither gender nor expertise recognition had a significant effect on control behavior in the Human-Decides approach.

Results of the IntelSys-Advises approach. As expected, because the advice is from IntelSys, the results of the automation ratio in the IntelSys-Advises approach are different from the automation ratio in the Human-Decides approach. Table 5.16 and Table 5.7c show the ANCOVA results of the IntelSys-Advises approach from different perspectives. To begin with, the whole ANCOVA model is significant (F = 6.752, df = 5, p = 0.000, Table 5.16). The results of Comparison 1 in Table 5.16 suggest that under low task uncertainty, expertise is not a significant factor that impacts the automation ratio (p = 0.435). However, the results of Comparison 2 suggest that under high task uncertainty, novices (automation ratio = 0.518) used significantly more automation than experts (automation ratio = 0.344) (Δ= 0.136, p = 0.048, Table 5.16). The results of Comparison 3 (albeit marginally significant, Δ= 0.106, p = 0.075) and 4 (Δ= 0.197, p = 0.001) suggest that both novices and experts used more automation under low task uncertainty than under high task uncertainty. The above results collectively suggest that in the IntelSys-Advices approach, task uncertainty has a damping effect on the automation ratio for both experts and novices. In other words, task uncertainty has a negative effect on the automation ratio in the Intel-Sys Advises approach.

The results of Table 5.17 suggest that the male and female participants did not show a significant difference in their automation ratio. The impact of expertise recognition showed a negative effect on the automation ratio, suggesting that the more optimistic participants were regarding their own skills, the less likely they were to allocate the task to automation in the Intel-Sys Advices DMA.

Summary of findings. To achieve a more holistic understanding, we integrate the results discussed above in Figure 5.9. As can be seen in Figure 5.9, the discussion thus far arrives at the following three main observations (Figure 5.9). First, for the IntelSys-Decides approach, the automation ratios are approximately 0.5 and bear no significant difference across different conditions. Second, for the Human-Decides approach, only expertise impacts the automation ratio in low task uncertainty situations. In contrast, for the IntelSys-Advises approach, expertise negatively impacts the automation ratio in high task uncertainty situations. Third, for both the Human-Decides and IntelSys-Advises approaches, task uncertainty negatively impacts the automation ratio for all participants, regardless of their level of expertise.



(TU = Task Uncertainty)

Figure 5.9 The overall results of the automation ratio across different DMAs

In the next section, I go one step further and compare the automation ratio in IntelSys-Decides, Human-Decides, and IntelSys-Advises under each situation. The goal is to assess (a) the difference between IntelSys and humans when making task allocation decisions under different situations; and (b) the difference between Human-Decides and IntelSys-Advises under different situations.

5.3.3 Summary of the post-hoc analysis

5.3.3.1 Automation ratio → Team performance (within DMAs)

Figures 5.10a and 5.10b show the overall results of team performance and the automation ratio, respectively. When comparing Figure 5.10a and Figure 5.10b, I find that the overall distribution of the automation ratio of the three DMAs under different scenarios is consistent with the overall distribution of team performance across different scenarios. To elaborate, in the low task uncertainty situations (including both low and high expertise participants), both team performance (Figure 5.10a) and the automation ratio (Figure 5.10b) show an inverted V shape. Next, under the situation of high task uncertainty and low expertise, both team performance (Figure 5.10a) and automation time (Figure 5.10b) show a relatively flat trend. Finally, under the high task uncertainty and high expertise situation, both team performance (Figure 5.10a) and the automation ratio (Figure 5.10a) show a descending trend, but to a different extent: team performance in Figure 5.10a displays a downward shape, while the automation ratio in Figure 5.10b shows a much sharper decrease from Human-Decides to InteSys-Advises. Such consistency implies a positive relationship between the automation ratio and team performance, which will be further explored in Section 5.3.4.

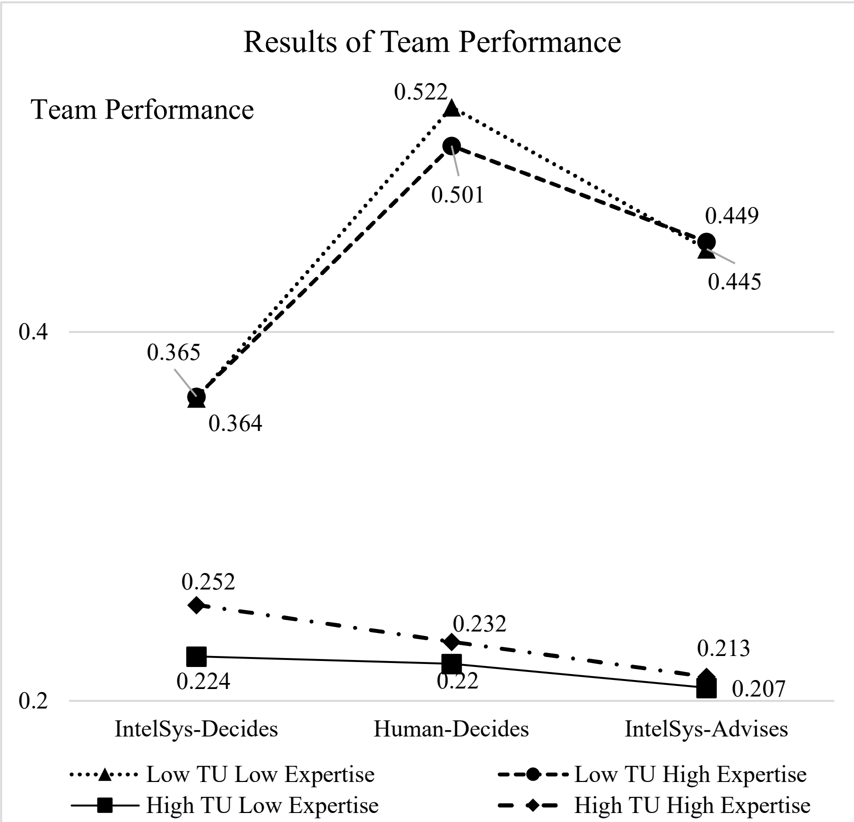


Figure 5.10a Overall results of team performance (within DMAs comparison)
(TU = Task Uncertainty)

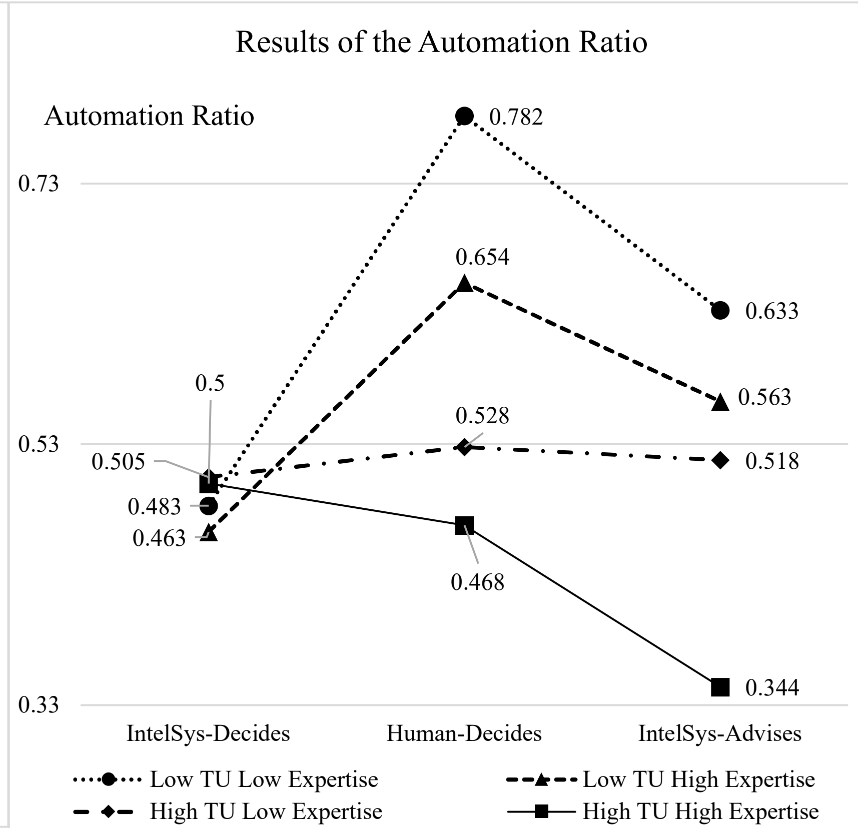


Figure 5.10b Overall results of automation ratio (within DMAs comparison)
(TU = Task Uncertainty)

Figure 5.10 Overall distribution of team performance and the automation ratio (within DMA comparison)

5.3.3.2 Automation ratio → Team performance (within scenarios)

I also investigate the variations of team performance and the automation ratio across different situations. Figure 5.11a and Figure 5.11b show the overall results of team performance and the automation ratio, respectively.

From Figure 5.11a and Figure 5.11b, I find that for almost all DMAs, there is a clear and consistent trend: both the automation ratio and team performance decrease from low task uncertainty conditions to high task uncertainty conditions. The only exception is the automation ratio of the IntelSys-Decides approach, which stays at around 0.5 across different situations. In this connection, the variation of team performance can be attributed to task uncertainty, except for the IntelSys-Decides approach.

For the Human-Decides approach, in the case of low task uncertainty, novices used more automation and attained higher team performance than the experts. In the case of high task uncertainty, there is no significant difference between novices and experts in terms of the automation ratio and team performance. These results collectively suggest that both experts and novices used more automation and achieved higher team performance in low task uncertainty situations than in high task uncertainty situations, rendering the following two general observations: (1) the trend of team performance and the trend of the automation ratio are similar; therefore, the automation ratio may positively affect team performance; and (2) the variations of team performance and the automation ratio suggest a possible interaction effect of task uncertainty and expertise on both team performance and the automation ratio.

The most interesting findings pertain to the IntelSys-Advises approach. In the low task uncertainty situations, there is no significant difference between novices and experts in terms of the distribution of the automation ratio and team performance. However, in the high task uncertainty situations, novices (0.518) used more automation than experts (0.344); yet there is no significant difference between their team performance. Both novices and experts used more automation and achieved higher team performance under low task uncertainty than high task uncertainty. The variations of team performance and the automation ratio in this approach also suggest a potential interaction effect between task uncertainty and expertise on team performance and the automation ratio.

Moreover, the variations of team performance and the automation ratio are very different across the three DMAs. It is, therefore, possible that the interaction effect between task uncertainty and expertise on the automation ratio and team performance, if any, might vary across different DMAs. In order to verify this conjecture, I further conduct a moderated mediation analysis in the next section.

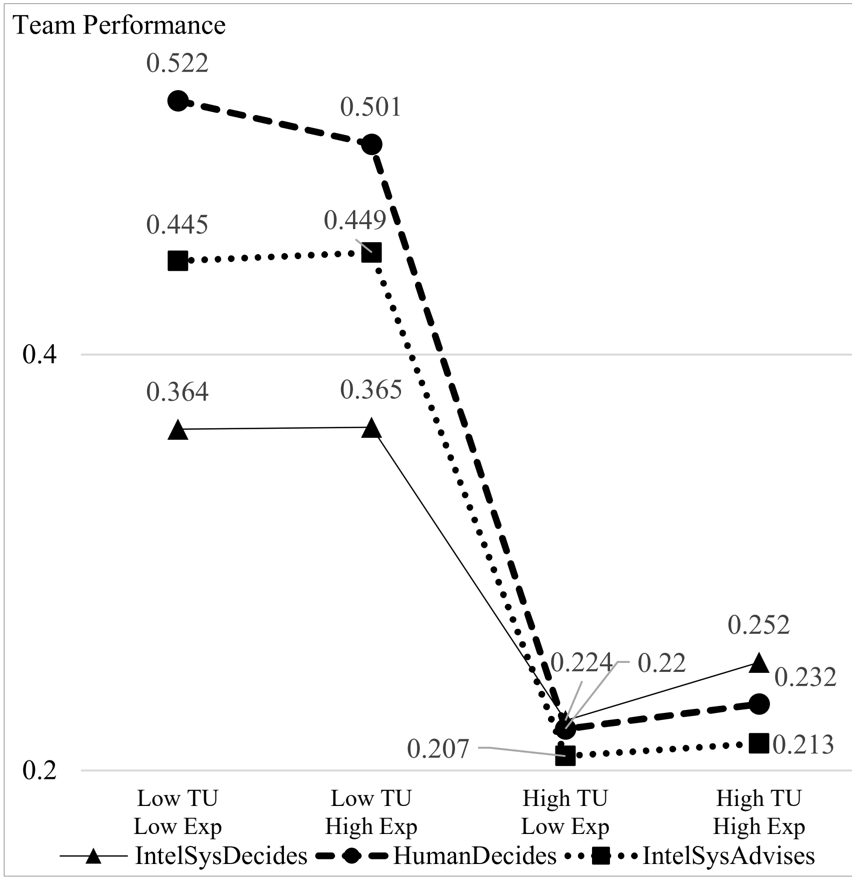


Figure 5.11a Results of team performance
(within senario comparison)
TU = Task Uncertainty, Exp = Expertise

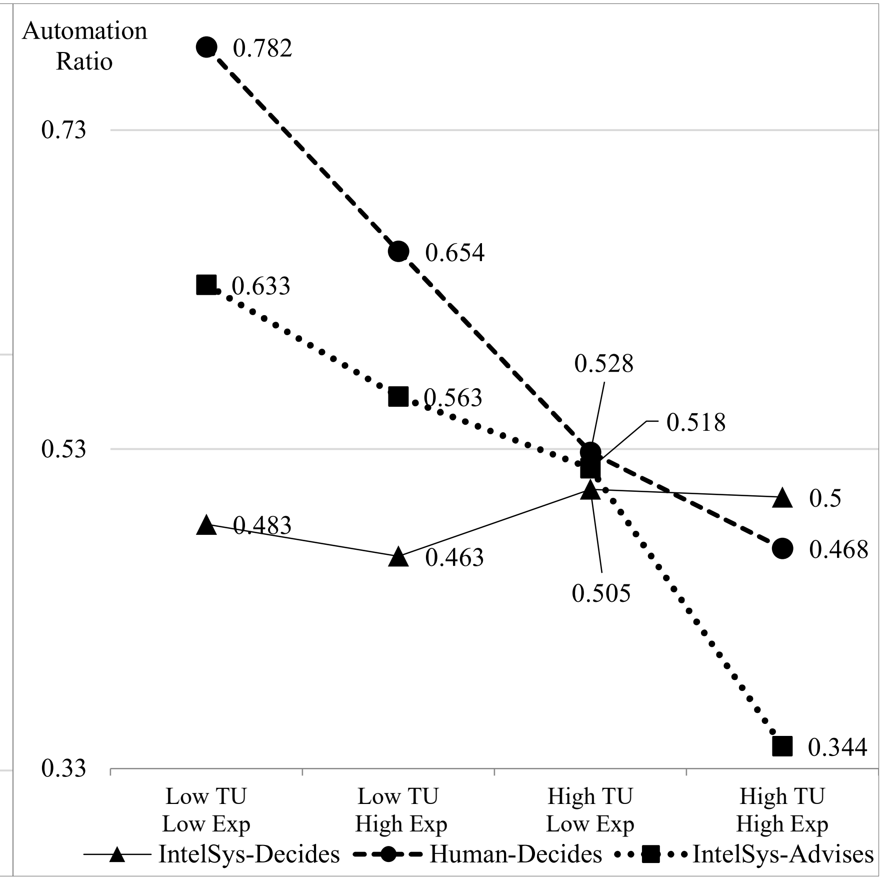


Figure 5.11b Results of the automation ratio
(within senario comparison)
TU = Task Uncertainty, Exp = Expertise

Figure 5.11 Overall distribution of team performance and the automation ratio (within scenario comparison)

5.3.4 Moderated mediation analysis

Based on the results discussed so far, I find that (1) when humans have the decision-making authority, task uncertainty and expertise can impact the automation ratio and team performance; and (2) for all DMAs, the automation ratio seems positively related to team performance. Moreover, based on the personal control model elaborated in Section 2.3.4, the impact of task uncertainty on the automation ratio is contingent on human expertise. Specifically, when facing high task uncertainty, novices could experience learned helplessness when realizing that the automation performs much better than they do; thus, they rely more on automation. In contrast, when encountering tasks with high uncertainty, experts are more likely to seek more control, which will be reflected in a low automation ratio (Table 5.13). Furthermore, according to Parasuraman et al. (2007) and Graham et al. (2009), compared with novices, experts are more confident about their expertise, and therefore tend to rely more on themselves. The above discussion suggests that expertise could negatively impact the automation ratio.

Given the above, I develop a moderated mediation model, as shown in Figure 5.12. It is noteworthy that the possible relations of Figure 5.12 are only suitable for the Human-Decides and IntelSys-Advises approaches, where humans have the decision-making authority. For the IntelSys-Decides approach wherein humans have no decision-making authority, I only assess whether (1) task uncertainty has a negative impact on team performance (Cordery et al. 2010); and whether (2) expertise and the automation ratio have a positive impact on team performance.

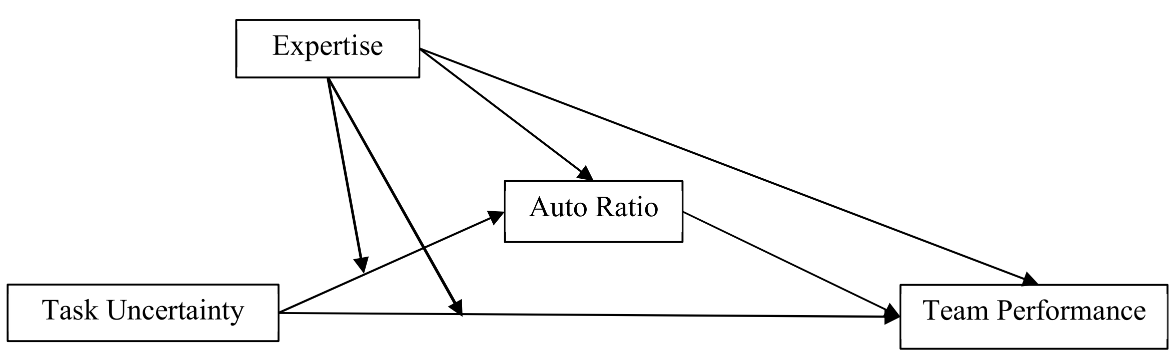


Figure 5.12 Moderated mediation model

Following the instruction by Hayes (2018), I then conduct three separate moderated mediation analyses using PROCESS. Table 5.18 shows the results of the direct effects. Tables 5.19, 5.20, and 5.21 show the results of the indirect effects.

Table 5.18 Mediation analysis results (direct effect)

Direct Effect	**IntelSys-Decides**		**Human-Decides**		**IntelSys-Advises**	
DV	Auto Ratio	Team Perf	Auto Ratio	Team Perf	Auto Ratio	Team Perf
Covariates						
Gender	-0.104 (0.083)	-0.015 (0.061)	0.056 (0.082)	0.037 (0.024)	-0.065 (0.076)	0.022 (0.026)
Expertise Recognition	-0.170 (0.067) *	0.057 (0.050)	*-0.114 (0.065)* +	0.020 (0.019)	-0.150 (0.067) *	0.068 (0.023) **
IV						
Task Uncertainty	0.077 (0.065)	-0.640 (0.048) ***	-0.293 (0.065) ***	-0.725 (0.020) ***	-0.239 (0.066) ***	-0.664 (0.024) ***
Expertise	0.048 (0.081)	0.079 (0.059)	-0.159 (0.082) *	0.019 (0.024)	*-0.134 (0.075)* +	0.098 (0.026) ***
Interaction						
TU x Expertise	0.001 (0.066)	0.067 (0.048)	0.040 (0.065)	0.027 (0.019)	-0.072 (0.065)	0.044 (0.023) *
Mediator						
Auto Ratio		0.327 (0.049) ***		0.450 (0.020) ***		0.551 (0.024) ***
R-Sqr	0.046	0.496	0.122	0.923	0.1402	0.896
MSE	0.975	0.518	0.900	0.079	0.881	0.107
F	2.188	36.848	5.865	422.495	6.752	296.591

DV=Dependent variables; IV=Independent variables; Auto Ratio=Automation Ratio; Team Perf=Team Performance
*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)

IntelSys-Decides approach. For the IntelSys-Decides approach, IntelSys makes task allocation decisions based on the participants' workload (i.e., the performance of the gauge-monitoring task). The results in Table 5.18 suggest that expertise and task uncertainty have no direct effect on the automation ratio. For task performance, while expertise exerts no impact, task uncertainty has a negative impact on team performance ($\beta = -0.640$, $p < 0.001$), which is consistent with existing research (Cordery et al. 2010). The automation ratio has a positive effect on team performance ($\beta = 0.327$, $p < 0.001$). Table 5.19 suggests that the indirect effects of task uncertainty and expertise on team performance are not significant. The moderated mediation effect is also not significant.

Table 5.19 Results of the indirect effect of the System-Decides approach

Conditional indirect effect(s) of Task Uncertainty on Team Performance						Significance
Mediator	Expertise	Effect	Boot SE	BootLLCI	BootULCI	
Auto Ratio	Low	0.025	0.028	-0.019	0.096	Insignificant
Auto Ratio	High	0.026	0.031	-0.029	0.094	Insignificant
Conditional indirect effect(s) of Expertise on Team Performance						
Mediator	TU	Effect	Boot SE	BootLLCI	BootULCI	
Auto Ratio	Low	0.015	0.040	-0.053	0.1	Insignificant
Auto Ratio	High	0.016	0.022	-0.020	0.071	Insignificant
IMM	Index	Boot SE	BootLLCI	BootULCI		
Auto Ratio	0.001	0.039	-0.086	0.070		Insignificant

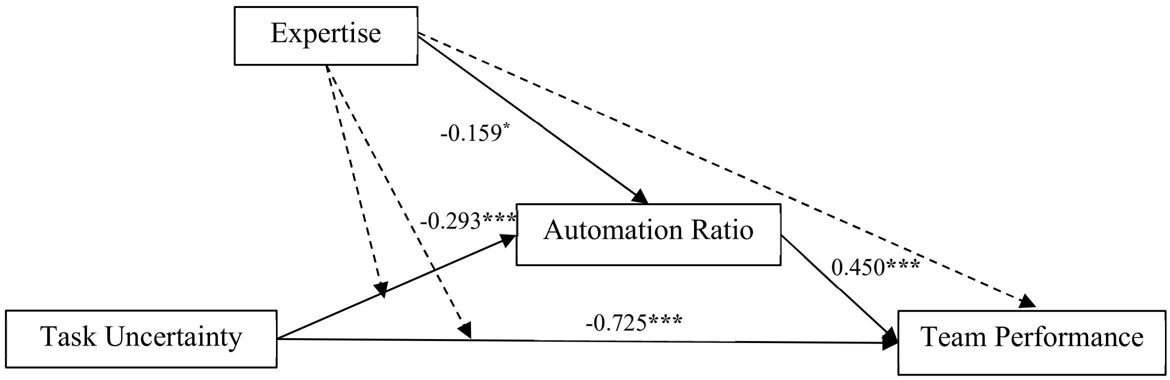
TU: Task Uncertainty; Auto Ratio: Automation Ratio; Confidence Level: 95%; IMM: Index of Moderated Mediation

Human-Decides approach. The results of Table 5.18 suggest that (1) expertise (β= -0.159, $p<0.05$) has a negative impact on the automation ratio, suggesting that participants with high expertise used less automation than participants with low expertise; (2) task uncertainty (β= -0.293, $p<0.001$) has a negative impact on the automation ratio, suggesting that participants in high task uncertainty situations used less automation than in low task uncertainty situations; (3) task uncertainty has a negative impact (β= -0.725, $p<0.001$) on team performance, suggesting that due to a lack of sufficient information about the paths of the hazards, participants' shooting performance went down; and (4) the automation ratio has a positive effect on team performance (β= 0.450, $p<0.001$). The indirect effect results of Table 5.20 suggest that, as expected, (1) for both novice and expert participants, the automation ratio negatively mediated the relation between task uncertainty and team performance (for novices, β= -0.149, CI: [-0.224, -0.08]; for experts, β= -0.113, CI: [-0.203, -0.023]); and (2) the automation ratio negatively mediated the relation between expertise and team performance, but only in the low task uncertainty situation (β= -0.088, CI: [-0.166, -0.014]). Figure 5.12 shows the overall results of the Human-Decides approach.

Table 5.20 Results of the indirect effect of the Human-Decides approach

Conditional indirect effect(s) of Task Uncertainty on Team Performance						Significance
Mediator	Expertise	Effect	Boot SE	BootLLCI	BootULCI	
Auto Ratio	Low	**-0.149**	0.036	-0.224	-0.08	Significant
Auto Ratio	High	**-0.113**	0.045	-0.203	-0.023	Significant
Conditional indirect effect(s) of Expertise on Team Performance						
Mediator	TU	Effect	Boot SE	BootLLCI	BootULCI	Significance
Auto Ratio	Low	**-0.088**	0.038	-0.166	-0.014	Significant
Auto Ratio	High	-0.052	0.05	-0.145	0.045	Insignificant
	Index	Boot SE	BootLLCI	BootULCI		
Auto Ratio	0.036	0.058	-0.076	0.154		Insignificant

TU: Task Uncertainty; Auto Ratio: Automation Ratio; Confidence Level: 95%; IMM: Index of Moderated Mediation



*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)

Figure 5.13 Overall results of the Human-Decides approach

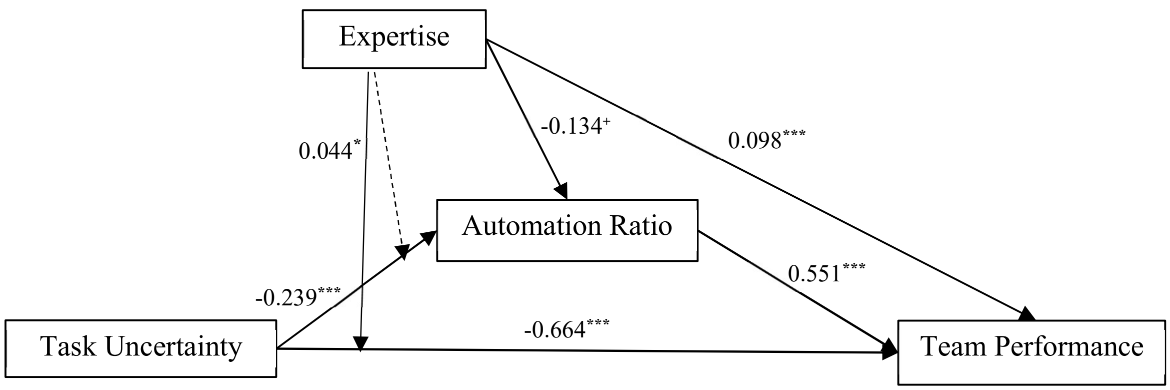
The results of Table 5.20 suggest that the interaction of task uncertainty and expertise has no impact on team performance, indicating that expertise does not moderate the impact of task uncertainty on team performance. Second, the index of moderated mediation is insignificant, suggesting that the difference between the indirect effect of task uncertainty on team performance for novices and the indirect effect of task uncertainty on team performance for experts is not significant (Hayes 2015). As mentioned in Section 5.3.1, under the condition of high task uncertainty, participants with low expertise did not exhibit learned helplessness, which is the second stage of humans' control behavior (Greenberger and Strasser 1986). Still, participants with low expertise sought more control, which is the first stage of humans' control behavior (Greenberger and Strasser 1986). Figure 5.13 shows the overall results of the Human-Decides approach.

IntelSys-Advises approach. For the IntelSys-Advises approach, from Table 5.18, I see that (1) task uncertainty has a negative effect on team performance (β= -0.664, p <0.001) and the automation ratio (β= -0.239, p<0.001); (2) expertise has a negative (albeit marginally) impact on the automation ratio (β= -0.134, p<0.1); (3) expertise has a positive impact on team performance (β= 0.098, p<0.001); (4) the automation ratio has a positive effect on team performance (β= 0.551, p<0.001); and (5) the interaction of task uncertainty and expertise has a positive impact on team performance (β= 0.044, p<0.05), indicating that expertise negatively moderates the negative relationship between task uncertainty and team performance.

Table 5.21 Results of the indirect effect of the System-Advises approach

Conditional indirect effect(s) of Task Uncertainty on Team Performance						Significance
Mediator	Expertise	Effect	Boot SE	BootLLCI	BootULCI	
Auto Ratio	Low	-0.092	0.051	-0.201	0.0003	Insignificant
Auto Ratio	High	**-0.171**	0.052	-0.28	-0.084	Significant
Conditional indirect effect(s) of Expertise on Team Performance						
Mediator	TU	Effect	Boot SE	BootLLCI	BootULCI	Significance
Auto Ratio	Low	-0.04	0.053	-0.145	0.064	Insignificant
Auto Ratio	High	**-0.119**	0.061	-0.243	-0.001	Significant
Index of Moderated Mediation	Index	Boot SE	BootLLCI	BootULCI		
Auto Ratio	-0.079	0.073	-0.218	0.074		Insignificant

Confidence Level: 95%



*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)
TU = Task Uncertainty

Figure 5.14 Overall results of IntelSys-Advises approach

For the indirect impact, based on Table 5.21, I found that (1) task uncertainty has indirect effects on team performance through the automation ratio, but only for participants with high expertise (β= -0.171, CI: [-0.28, -0.084]); (2) the automation ratio mediated the relation between expertise and team performance, but only in high task uncertainty situations (β= -0.119, CI: [-0.243, -0.001]); (3) the interaction of task uncertainty and expertise has no impact on the automation ratio, indicating that expertise does not moderate the impact of task uncertainty on the automation ratio; and (4) the moderated mediation effect is not significant. One possible reason is that under the condition of high task uncertainty, participants with low expertise did not experience learned helplessness, but rather sought more control (Greenberger and Strasser 1986).

The overall results of Figure 5.14 suggest that for the IntelSys-Advises approach, the factors that impact team performance are: task uncertainty, expertise, the automation ratio, the indirect effect of expertise on team performance, the indirect effect of task uncertainty on team performance, and the interaction of task uncertainty and expertise.

For the covariates, gender has no significant effect on either the automation ratio or team performance for all DMAs (Table 5.18). Similar to the findings in Section 5.3.1, for the IntelSys-Decides and IntelSys-Advises approaches, expertise recognition has a negative impact on the automation ratio (for IntelSys-Decides, $\beta = -0.170$, $p<0.05$; for IntelSys-Advises, $\beta = -0.150$, $p<0.01$).

There are two additional findings on the effect of expertise recognition. First, for the Human-Decides approach, expertise recognition negatively (albeit marginally, $\beta = -0.114$, $p=0.065$, Table 5.18) impacts the automation ratio, indicating that participants who have a favorable view of their own expertise are inclined to rely more on themselves. Second, for the IntelSys-Advises approach, expertise recognition positively ($\beta = 0.068$, $p=0.023$, Table 5.18) impacts team performance. After checking the distribution of expertise recognition, I find that 3.7% of participants underestimate their expertise, 80.3% of participants accurately estimate their expertise, and 17.4% of participants overestimate their expertise. A possible explanation for the positive impact of expertise recognition on team performance in the IntelSys-Advises approach could be that most participants (80.3%) accurately recognize their expertise, compared to the expertise of automation. According to Littlepage et al. (1995) and Littlepage and Mueller (1997), accurate expertise recognition has a positive impact on expertise coordination, which then leads to high team performance.

Summary

For all DMAs, (1) task uncertainty has a negative impact on team performance, and (2) the automation ratio has a positive impact on team performance. For both the Human-Decides and IntelSys-Advises approaches, in which participants had the decision-making authority, task allocation decisions are impacted by task uncertainty. Specifically, participants used less automation under the condition of high task uncertainty than under that of low task uncertainty. Also, compared to participants with low expertise, participants with high expertise relied more on themselves and used automation less often. However, in the Human-Decides approach, expertise has no significant impact on team performance, whereas in the IntelSys-Advises approach, expertise has a significant positive impact on team performance.

For the Human-Decides approach, the automation ratio mediated the relationship between task uncertainty and team performance for both novices and experts. For the IntelSys-Advises approach, the automation ratio only mediated the relationship between task uncertainty and team performance for experts, but not for novices. The mediating effect of the automation ratio on the relation between expertise and team performance also differs between the Human-Decides and IntelSys-Advises approaches. The mediation effect is only significant for novices in the Human-Decides approach, while the mediation effect is only significant for experts in the IntelSys-Advises approach. For both approaches, the moderated mediation effect is not significant, which suggests that for both novices and experts, the difference between the mediation effects of the automation ratio is not significant. One possible reason is that participants with low expertise remain in the state of seeking more control rather than moving to the state of learned helplessness.

5.3.5 Participants' advice acceptance behavior in the IntelSys-Advises approach

In the IntelSys-Advises approach, participants obtained advice from IntelSys. For this approach, there are two unexpected findings mentioned above. First, in Section 5.2, under the condition of high task uncertainty and high expertise, team performance in the Human-Decides approach is higher than that in the IntselSys-Advises approach. Second, in Section 5.3.2, under the condition of high task uncertainty and high expertise, participants in the Human-Decides approach used more automation than participants in the IntelSys-Advises approach.

Suspecting whether the particular advice offered by IntelSys plays a role in this approach, I further examine how the participants responded to the advice from IntelSys. In particular, IntelSys provides two types of advice to participants: "switch to manual mode" or "switch to automation mode." Participants can either accept or reject the advice within seven seconds. I define the manual acceptance rate as the proportion of the number of “switch to manual mode” advice prompts accepted by participants to the total number of “switch to manual mode” advice prompts. I define the automation acceptance rate as the proportion of the number of “switch to automation mode” advice prompts accepted by participants to the total number of “switch to automation mode” advice prompts. Table 5.22 shows the descriptive results of the automation acceptance rate and manual acceptance rate in the IntelSys-Advises approach. I then conduct a one-way ANOVA analysis to assess the automation acceptance rate and manual acceptance rate, respectively.

Table 5.22 Descriptive results of the acceptance rate in the IntelSys-Advises approach

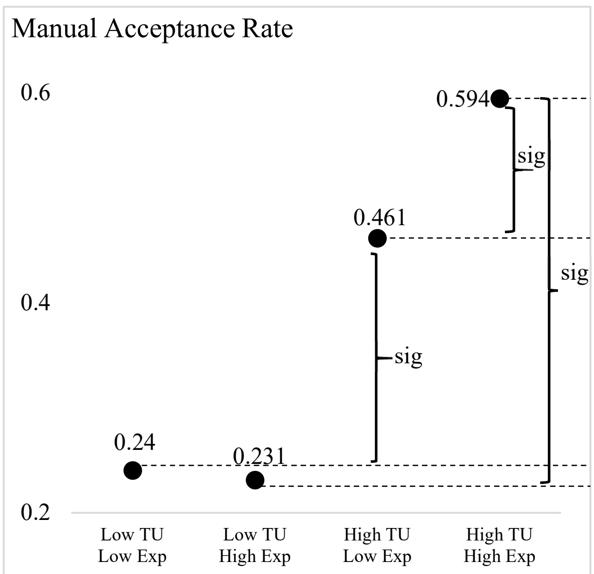
	N	Mean	SD
Manual Acceptance Rate			
Low TU Low Expertise	62	0.24	0.285
Low TU High Expertise	58	0.231	0.303
High TU Low Expertise	44	0.461	0.339
High TU High Expertise	49	0.594	0.374
Automation Acceptance Rate			
Low TU Low Expertise	62	0.492	0.343
Low TU High Expertise	58	0.389	0.346
High TU Low Expertise	44	0.273	0.280
High TU High Expertise	49	0.310	0.340

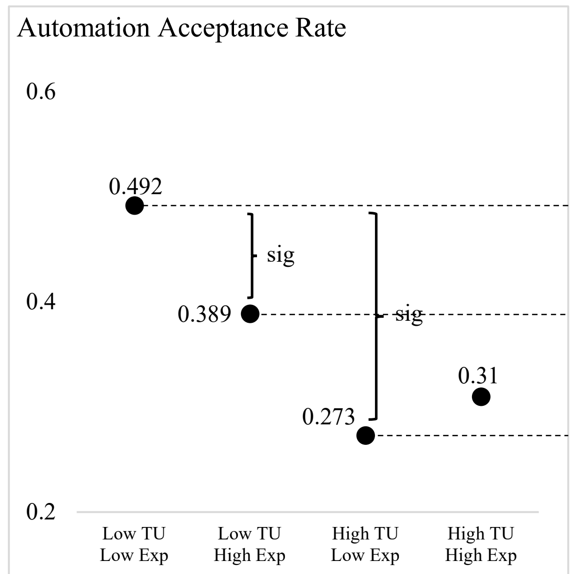
Table 5.23 and Figure 5.15a show the ANOVA results of the manual acceptance rate. In the condition of low task uncertainty, both novices ($p = 0.001$) and experts ($p < 0.001$) accepted more “switch to manual mode” advice prompts than under the condition of high task uncertainty. The results of Section 5.3.1 suggest that both novices and experts sought more control under high task uncertainty than under low task uncertainty. Therefore, the results of Figure 5.15a are consistent with the results of Section 5.3.2. It is reasonable to infer that, under high task uncertainty, the “switch to manual mode” advice prompts of IntelSys fit with the participants’ desire to seek more control, and therefore are more likely to be accepted. Furthermore, under the condition of high task uncertainty, experts accepted significantly more “switch to manual mode” advice prompts than novices ($p = 0.049$).

Table 5.23 Results of manual and automation acceptance rate in the IntelSys-Advises approach

	Manual acceptance rate		Automation acceptance rate	
Comparison (LSD)	**MD**	**p-value**	**MD**	**p-value**
Low TU Low Expertise vs. Low TU High Expertise	0.009	0.875	0.103	0.089+
Low TU Low Expertise vs. High TU Low Expertise	-0.221	0.001**	0.219	0.001**
High TU High Expertise vs. Low TU High Expertise	0.363	0.000***	-0.079	0.221
High TU High Expertise vs. High TU Low Expertise	0.133	0.049*	0.037	0.589
Model	df = 3, F= 15.890, p = 0.000***		df = 3, F= 4.644, p = 0.004**	

*: p<0.05 (two-tailed); **: p<0.01 (two-tailed); ***: p<0.001 (two-tailed); +: p<0.1 (two-tailed)





(TU = Task Uncertainty; Exp = Expertise; sig = Significant Difference)

Figure 5.15a Manual acceptance rate Figure 5.15b Automation acceptance rate

Figure 5.15 Participants' advice acceptance rate

Recall that in the scenario of high task uncertainty and high expertise, participants in the IntelSys-Advises approach used less automation than participants in the Human-Decides approach. The difference may be attributed to the fact that experts prefer to accept more "switch to manual mode" advice prompts in the IntelSys-Advises approach. Figure 5.15a suggests that under the condition of high task uncertainty, participants with high expertise accepted the highest level of "switch to manual mode" advice prompts (0.594). Figure 5.15b suggests that under the condition of high task uncertainty, almost all participants with high expertise accepted the lowest level of "switch to automation mode" advice prompts (0.310). It appears that under the condition of high task uncertainty, the advice of IntelSys enhances expert participants' desire for control. Therefore, under the condition of high task uncertainty and high expertise, relying less on automation significantly lowered the team performance of IntelSys-Advises with respect to the team performance of Human-Decides.

Finally, Table 5.23 and Figure 5.15b show the ANOVA results of the automation acceptance rate. The novices accepted more “switch to automation mode” advice prompts under the condition of low task uncertainty than under that of high task uncertainty ($p = 0.001$). This is consistent with the results of Section 5.3.2, which suggest that in low task uncertainty situations, novices used the automation mode more than the manual mode. Moreover, under the condition of low task uncertainty, novices accepted more “switch to automation mode” advice prompts than experts ($p = 0.089$). This is consistent with the results of Section 5.3.2, which suggests that experts sought more control than novices.

6 Discussion

6.1 Results Summary

The controversy about the decision-making approaches in human-automation interaction motivates me to conduct this research. In this study, I propose four hypotheses to compare the impacts of three DMAs on human-automation team performance, having task uncertainty and expertise as the two core boundary conditions. The results of the large-scale lab experiment provide support for H1 and H2; that is, for both novices and experts under the condition of low task uncertainty, team performance is the highest for the Human-Decides approach, followed by the IntelSys-Advises approach, and then the IntelSys-Decides approach. However, the results are quite different in the case of high task uncertainty. For novices under high task uncertainty, there are no significant differences in team performance across the IntelSys-Decides, Human-Decides, and IntelSys-Advises approaches, thus rejecting H3. Finally, H4 is partially supported. For experts under high task uncertainty, I found that the team performance of the IntelSys-Decides approach is higher than that of the Human-Decides and IntelSys-Advises approaches. Yet, contrary to H4, the team performance of Human-Decides is higher than that of IntelSys-Advises; this may be partially explained by the fact that experts in the Human-Decides approach used more automation than experts in the IntelSys-Advises approach.

I then conducted post-hoc analyses to examine how the participants and IntelSys made decisions in different situations. The results show that automation plays better than 95% of the participants in both the low and high uncertainty tasks (which is consistent with the premise of this study stated in Chapter 1); thus, it is reasonable to observe that the automation ratio positively affects team performance. In other words, the more humans delegate the task to automation (i.e., a higher automation ratio), the better the team performance will be.

Next, I compare the automation ratio of the three DMAs in each situation. The results suggest that (1) the automation ratio of IntelSys-Decides is very stable, running about 50% across all situations; (2) under low task uncertainty, (a) both experts and novices in the Human-Decides approach and (b) novices in the Intel-Advises approach exhibit a higher automation ratio than novices in the IntelSys-Decides approach; (3) under high task uncertainty, when having the decision-making authority (e.g., Human-Decides and IntelSys-Advises approaches), most participants (except experts in IntelSys-Advises) used automation approximately 50% in the IntelSys-Decides approach, which is higher than the automation ratio for experts in the IntelSys-Advises approach (34.4%). Interestingly, in the case of high task uncertainty, experts in the IntelSys-Advises approach use less automation than experts in the Human-Decides approach. I further examine the automation ratio in the IntelSys-Advises approach and find that expert participants are more likely to accept "switch to manual" advice prompts (59.4% acceptance rate).

I also analyzed the decision-making behaviors of each DMA across different situations. I find that (1) for the Human-Decides and IntelSys-Advises approaches, task uncertainty negatively impacts the automation ratio; (2) for the Human-Decides approach, expertise negatively affects the automation ratio in the low task uncertainty situation; and (3) for the IntelSys-Advises approach, expertise negatively impacts the automation ratio in the high task uncertainty situation. Interestingly, I find that that participants with low expertise (i.e., the novice) under the condition of high task uncertainty do not go on to exhibit learned helplessness, but instead seek more control.

Finally, I conducted a moderated mediation analysis. The results of this analysis suggest that for the Human-Decides approach, (1) the automation ratio mediates the relationship between task uncertainty and team performance for both novices and experts; and (2) the automation ratio mediates the effect of expertise on team performance, but only for participants with low expertise. For the IntelSys-Advises approach, (1) expertise moderates the impact of task uncertainty on team performance; (2) the automation ratio mediates the relationship between task uncertainty and team performance for experts; and (3) the automation ratio mediates the effect of expertise on team performance, but only for participants with high expertise.

6.2 Theoretical Contributions

6.2.1 Theory contribution to the HAI field

The findings of this study bring a number of theoretical contributions to different research streams. For the line of HAI research, first, while prior research has compared IntelSys-Decides versus Human-Decides, or IntelSys-Decides versus IntelSys-Advises, only a few papers have compared all three DMAs (Kidwell et al. 2012; Li 2013; Ruff et al. 2018). However, the results of these studies reveal inconsistent findings on which DMA is the best. Unlike these prior empirical studies, which do not propose a theoretical lens, this thesis theoretically examines the three DMAs from the lens of team-based decision making and develops a theoretical model to compare and contrast the relative advantages and disadvantages of the IntelSys-Decides, Human-Decides, and IntelSys-Advises approaches.

The theoretical model proposed in this study also identifies the boundary conditions for the three DMAs. While some researchers indicate that the Human-Decides approach ensures better team performance than the IntelSys-Decides approach due to the fact that humans can make more flexible, customized, and informed decisions than IntelSys (Parasuraman et al. 2005; Parasuraman et al. 2007), this view ignores the possibility that human decisions in the Human-Decides and IntelSys-Advises approaches could be affected by task uncertainty. Toward this end, I find that humans' decisions are subject to the level of task uncertainty. Human decisions are more likely to follow the expertise coordination principle in low task uncertainty situations than in high task uncertainty situations. This is because both novices and experts sought more control and tended to allocate more tasks to themselves in the high task uncertainty situations than in the low task uncertainty situations, albeit to a different extent. I also find that expertise can impact humans' control behavior.

Moreover, the results of the post-hoc analyses open up the black box of decision making in human-automation interaction. Most prior HAI research has focused only on decision outcomes rather than the decision-making process. To the best of my knowledge, this study is one of the first to reveal the differences among IntelSys-Decides, Human-Decides, and IntelSys-Advises from the perspective of the decision-making process. This is made possible by looking into the details of how decision-makers (e.g., participants and IntelSys) use automation and manual modes during the entire process of the experiment. Specifically, the automation ratio (as an indicator that reflects the participants' control behavior) is used to measure how participants made decisions during the experiment. I further explore the mediating effect of the automation ratio and find that this ratio mediates the effect of task uncertainty and expertise on team performance. To obtain further insights into the decision-making black box, I also examine how participants respond to the advice generated from IntelSys in the IntelSys-Advises approach.

Last, but not least, this paper is one of the first theoretically grounded studies in the HAI field, which has been criticized as lacking in theoretical accounts. Traditional HAI papers have focused more on the final results rather than constructing a theoretical framework to explore the underlying mechanism leading to the final outcomes. This study creatively combines team-based decision-making research and the personal control model; moreover, it proposes a theoretical model to explain how DMAs achieve different levels of task performance under different conditions.

6.2.2 Theory contribution to the human-automation team

This study also makes contributions to research on human-automation teams. To begin with, this thesis approaches HAI from the team perspective and clarifies two core features of the human-automation team (Section 2.2.2.3). First, automation and humans are highly heterogeneous in terms of task expertise, but they have shared team goals in my investigative context. Second, team success can be achieved by following the expertise coordination principle, which suggests that team performance can be optimized by allocating more tasks to members with stronger competency. Therefore, for the human-automation team, when humans have the decision-making authority, factors such as heterogeneous capabilities between humans and automation (and hence, the expertise coordination principle) should be considered in order to ensure the team's success (in Section 2.2.1).

Second, this thesis enriches the team literature by studying the contextual variable of task uncertainty. Scholars have proposed the need to investigate the impact of task characteristics on team performance (Lewis and Herndon 2011; Baumann and Bonner 2017). In response to this call, in this thesis, I find that task uncertainty is a significant factor impacting humans' task-allocation decisions, which then impact team performance. In particular, the results suggest that task uncertainty influences humans' control behavior, affecting how humans allocate tasks between themselves and automation.

Last, but not least, this research introduces a research problem that has been examined in engineering, but not in the information system field. To the best of my knowledge, this thesis is one of the few that study human- vs. IntelSys-based decision making in the context of HAI. Moreover, I apply the theoretical lens of team-based decision making to address problems in engineering. The cross-disciplinary nature of this dissertation enriches knowledge in both the fields of information systems and engineering areas.

6.3 Practical Implications

This research bears valuable practical implications for industry, especially for aviation, air traffic control, autonomous vehicles, and manufacturing, among others. In particular, the findings provide practical guidelines for firms on how to incorporate IntelSys in making task-allocation decisions between human operators and automation. These guidelines should promote (as well as provoke) practitioners to reconsider existing practices. For instance, current aviation designs often delegate pilots as the decision maker to switch between the manual mode and the autopilot mode in the event of extreme turbulence (Billings 1997), which represents a typical situation when experts work in a high task uncertainty environment. However, the results of this thesis suggest that when tasks are highly uncertain, it is necessary to recognize experts' impulse to take control; in this case, IntelSys-Decides could be a desirable DMA, as IntelSys can make objective, accurate, and timely responses. In contrast, when tasks are more certain, humans should take more authority to make such task allocation decisions.

As a typical unmanned ground vehicle, the autonomous car has become increasingly popular. Tesla, for instance, is the pioneer in the autonomous car industry. However, there have been at least eight accidents involving Tesla cars since 2016. Most of these accidents occurred because human drivers place an inappropriate amount of trust in the autopilot system, and thus allocate too much of the driving task to the autopilot system, which, is not capable enough to replace human driving completely. Against this backdrop, how to allocate tasks between an automated system (e.g., the autopilot system) and a human driver is a critical problem that needs to be tackled further.

The essence of the task allocation decision involves which decision-making approach to use. To this end, the research findings of this thesis provide practical implications by clarifying the boundary conditions for three decision-making approaches. Based on the research findings, I suggest that, on the one hand, human drivers should be actively involved in the driving task so as to avoid human out-of-the-loop problems. On the other hand, for expert drivers under high task uncertainty, it is necessary to limit their impulse to take more control. Therefore, the appropriate level of involvement with respect to human drivers should be studied in the future. At the same time, the advantages of IntelSys should be considered. It is reasonable to expect that in the future, IntelSys in an autonomous car can gather human drivers' physiological measures such as eye movement, heart rate, and brain waves such that IntelSys can achieve a more precise assessment of human drivers' workload. As such, IntelSys can make accurate, objective, and timely decisions and should be considered in high task uncertainty situations.

The findings of this thesis also shed light on manufacturing. With the development of AI, the Internet of Things, and the deployment of 5G, humans and automation are now working together closely. Human-centered automation is becoming the norm in the "Industry 4.0" era (Fernández-Caramés and Fraga-Lamas 2018). Industry 4.0 advocates the application of intelligent manufacturing systems (IMS). One of the main functions of the IMS is distributed scheduling in manufacturing control (Pacaux-Lemoine et al. 2017). Human-centered IMS emphasizes that both humans and IMS (e.g., automation) should be involved in the manufacturing control process (Pacaux-Lemoine et al. 2017; Ansari et al. 2018; Fernández-Caramés and Fraga-Lamas 2018). People realize that humans' intelligence (e.g., humans' innovativeness in unforeseen situations, social accountability, expertise, and proficiency) and IMS should be integrated so that the combination can attain better performance in manufacturing control tasks (Pacaux-Lemoine et al.

2017; Longo et al. 2017; Ansari et al. 2018). However, the task allocation between humans and IMS has not been fully studied (Pacaux-Lemoine et al. 2017). As a result, this thesis provides a framework that is instrumental in understanding task allocation between humans and automation. Human expertise in manufacturing control and the uncertainty of manufacturing control tasks are two critical factors that should be further considered. The findings of this thesis imply that under the condition of high task uncertainty, to achieve high task performance, the human desire for control (i.e., seeking more control) could be limited by allocating more manufacturing control tasks to IMS.

This thesis also bears implications for sustainability in manufacturing (Nguyen and Krüger 2017). Sustainable manufacturing is defined as "*...the creation of manufactured products that use processes that minimize negative environmental impacts, conserve energy and natural resources, are safe for employees, communities, and consumers and are economically sound*" (The US Department of Commerce 2010; Faulkner and Badurdeen, 2014). As a severe threat to sustainability in manufacturing, musculoskeletal disorders (MSDs) are often attributed to "*rapid work pace and repetitive motion, forceful exertions, non-neutral body postures, and vibration*" (Punnett and Wegman 2004, p. 13). This thesis reveals another possible reason that may cause MSDs. That is, human operators allocate too many tasks to themselves rather than automation, especially in the case of high task uncertainty. In this vein, one possible solution is to reduce human operators' inappropriate high level of involvement and then allocate a greater share of the tasks to automation. By involving human operators appropriately in the task execution, factors causing MSDs could be effectively alleviated.

6.4 Limitations and Future Research

Like most empirical research, this study has its limitations. Below I discuss the limitations related to automation, IntelSys, task type, and experiment design. These issues may constrain the generalizability of the current findings, but may also shed light on further opportunities.

First, with respect to automation, this thesis only considers two modes in HAI: the total automation mode and the total manual mode. As mentioned in the literature review section, for many applications, there are different levels of automation (LOA) between the total automation mode and the total manual mode. For example, the Society of Automotive Engineers International has proposed six levels (from level 0, which is totally manual, to level 5, which is total automation) of driving automation, which is the current LOA model for the driverless car area. In this vein, the question regarding task allocation decisions becomes which of the six LOA levels that should be used in a specific situation. For example, for novice drivers under high task uncertainty, the particular LOA level of automation that should be used to decrease the driver's workload while guaranteeing driving safety is a decision to be determined. The interaction between task allocation approaches and LOA is a research topic that deserves further attention in the future. Finally, in practical applications, the design of automation does not always fall into the binary categories of full automation or fully manual. For example, decision aid automation can provide a couple of possible options to humans, who can then decide which one to choose (Verplank 1978). Hence, having more than one type of decision aid could be a possible direction for future research.

Next, this thesis does not consider the malfunction of automation. In reality, automation is not 100% stable and sometimes malfunctions. When automation malfunctions (e.g., stops working), humans with the decision-making authority (e.g., Human-Decides and IntelSys-Advises approaches) may have different perceptions regarding the capability and reliability of automation, which would have an impact on humans' switching decisions. Since automation could actually malfunction in the real world, future research should factor in malfunction as a part of the research model and research design.

Also, in terms of IntelSys, in this thesis, the triggering mechanism of IntelSys only considers the workload of the human player. As mentioned in Section 2.2.2.2, there are four methods that are often used to trigger the switching decision in the IntelSys-Decides approach: critical events, performance measurement, psychophysiological assessment (e.g., workload), and the behavior modeling strategy. The IntelSys that considers two or three methods is better and more reasonable than the IntelSys that only considers one triggering method (Parasuraman et al. 1992; Kaber and Endsley 2004). In this thesis, I only considered the psychophysiological assessment method. Combining two or more methods would lead to a more capable IntelSys that can make better switching decisions. In reality, the critical event method is often combined with other methods when designing IntelSys. For example, in autonomous cars, the collision avoidance system (CAS) keeps working to monitor the speed and distance between vehicles so as to prevent or reduce potential collisions. The Automatic Ground Collision Avoidance System (Auto GCAS) of F-16 fighters monitors the flight altitude. If the fighter is too close to the ground (too close that the pilot does not have enough reaction time), the Auto GCAS will automatically pull the nose of the fighter up so that the fighter is safe. Interested scholars could model two or more methods when designing IntelSys. Hence, future research could design a more practical IntelSys with a combination of two or more triggering methods for further investigation.

In addition, the platform in this study is an action game software that may constrain the generalizability of the results. The action game used in this study tests the players' cognitive abilities, such as spatial cognition, visual short-term memory, and coordination of visual capture and physical activities. All of these abilities can be generalized to other similar tasks, such as driving and air traffic control tasks. To advance this line of research, future studies could consider a different type of task. Different types of tasks may result in different conclusions, even with the same HAI design. The current task is a time-critical and low-risk task (i.e., the player's plane never dies). What will happen if the same task is employed, but the player's plane can be eliminated by hazards (e.g., high-risk tasks)? Following this line, other tasks such as airplane flying tasks and autonomous car driving tasks could be considered in the future.

Further, the choice of student participants may limit the generalizability of the results. The literature has expressed caution in using student subjects in experiments in terms of generalizability (Compeau et al. 2012). While acknowledging this limitation, the participants of this study are not inappropriate for the investigative context (i.e., action games), given that the undergraduates in this study are eligible representatives of action game players (Ogletree and Drake 2007; Terlecki et al. 2011). Nevertheless, I encourage recruiting different populations in future research to determine if the results of this study are replicable with different groups of individuals.

Finally, there are other factors that should be considered in future studies. For instance, team performance in the current study focuses on task performance. However, in reality, the design of human-automation interaction strategy often involves multiple objectives, such as high task performance, low human workload, high human situation awareness, and improvement of human skills. Future research can look into these dependent variables simultaneously. Also, team experience is a significant factor that may impact team performance (Littlepage et al. 1997). Team experience refers to team members' experience working with other team members (Littlepage et al. 1997). Team experience was found to have a positive impact on team performance because it allows team members to accurately recognize the expertise of other members (Littlepage et al. 1997). Therefore, interested scholars can enrich the current research model by considering how human participants' team experience affects team performance.

Citations

Arciszewski H.F. R., Greef T.E. de, and van Delft J. H. 2009. Adaptive automation in a naval combat management system. IEEE Transactions on Systems, Man, and Cybernetics—Part a: Systems and Humans, 39 (6), 1188 – 1199.

Agbejule, A., 2005. The relationship between management accounting systems and perceived environmental uncertainty on managerial performance: A research note. Accounting and Business Research, 354, 295–305.

Ashford S.J., and Black J.S., 1996. Proactivity during organizational entry: The role of desire for control. Journal of Applied Psychology, Vol. 81. No. 2, 199-214.

Ansari F., Hold P. and Sihn W. 2018. Human-centered cyber physical production system: How does Industry 4.0 impact on decision-making tasks? 2018 IEEE Technology and Engineering Management Conference (2018).

Ashford S.J., and Black J.S., 1996. Proactivity during organizational entry: The role of desire for control. Journal of Applied Psychology, Vol. 81. No. 2, 199-214.

Austin, J. R. 2003. Transactive memory in organizational groups: The effects of content, consensus, specialization, and accuracy on group performance. Journal of Applied Psychology, 88: 866-878.

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.

Baumann, M. R., and Bonner, B. L. (2017). An expectancy theory approach to group coordination: Expertise, task features, and member behavior. Journal of Behavioral Decision Making, 30, 407-419. doi:10.1002/bdm.1954.

Bazarova, N. N., and Yuan, Y. C. (2013). Expertise recognition and influence in intercultural groups: Differences between face-to-face and computer-mediated communication. Journal of Computer-Mediated Communication, 18, 437-453. doi:10.1111/jcc4.12018.

Bensaou, M., N. Venkatraman. 1995. Configurations of interorganizational relationships: A comparison between U.S. and Japanese automakers. Management Sci. 41(9) 1471–1492.

Brown, R., Hall, L., and Holtzer, R. (1997). Gender and video game performance. Sex Roles, 36(11/12), 793–812.

Burke, M. J., and Day, R. R. (1986). A cumulative study of the effectiveness of managerial training. Journal of Applied Psychology, 71, 232–245.

BEA. 2012. On the accident on 1st June 2009 to the Airbus A330-203 registered F-FGZCP operated by Air France flight AF 447 Rio de Janeiro – Paris. Cedex, France: Bureau d'Enquétes etd'Analyses pour la sécurité de l'aviation civile.

Beal, D. J., Cohen, R. R., Burke, M. J., and McLendon, C. L. 2003. Cohesion and performance in groups: A meta analytic clarification of construct relations. Journal of Applied Psychology, 88: 989-1004.

Billings, C. E. (1997). Aviation automation: The search for a human-centered approach. Mahwah, NJ: Lawrence Erlbaum Associates.

Billings, C. E. (1991). Human-centered aircraft automation; A concept and guidelines (NASA Tech. Memorandum 103885). Moffet Field, CA: NASA Ames Research Center.

Billings, C.E., and Woods, D.D. (1994). Concerns about adaptive automation in aviation systems. In M. Mouloua and R. Parasuraman (Eds.), Human Performance in Automated Systems: Current Research and Trends (pp. 264-269). Hillsdale, NJ: Erlbaum.

Bliss, J. P., Dunn, M. C. 2000. Behavioral implications of alarm mistrust as a function of task workload. Ergonomics (43)1283.

Brandt, S.L., Russell, R., Lachter, J., Shively, R.J., 2017. A human-autonomy teaming approach for a flight-following task. In: Proceedings of the 8th International Conference on Applied Human Factors and Ergonomics, this edition (2017).

Bonner, B. L. (2004). Expertise in group problem solving: Recognition, social combination, and performance. Group Dynamics: Theory, Research, and Practice, 4, 277–290.

Bonner, B. L. (2000). The effects of extroversion on influence in ambiguous group tasks. Small Group Research, 31, 225–244.

Bonner, B. L., Baumann, M. R., and Dalal, R. (2002). The effects of member expertise on group decision-making and group performance. Organizational Behavior and Human Decision Processes, 88, 719–736.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M. and Gratton, G. 2008. The effects of video game playing on attention, memory, and executive control. Acta Psychologica. (Amst.) 129, 387–398.

Bottger, P. C., and Philip W. Y. 1988. An integration of process and decision scheme explanations of group problem solving performance. Organizational Behavior and Human Decision Processes, 42: 234-249.

Braver, T. S., and Ruge, H. (2006). "Functional neuroimaging of executive functions," in Functional Neuroimaging of Cognition, (Eds.) R. Cabeza and A. Kingstone (Cambridge, MA: MIT), 307–347.

Bunderson, J. S. 2003. Recognizing and utilizing expertise in work groups: A status characteristics perspective. Administrative Science Quarterly, 48: 557-591.

Byrne, E. A., and Parasuraman, R. (1996). Psychophysiology and adaptive automation. Biological Psychology, 42, 249-268.

Cabrall, C. D. D., Janssen, N. M., and de Winter, J. C. F. 2018. Adaptive automation: Automatically (dis)engaging automation during visually distracted driving. Peer J Computer Science (4), pp. e166.

Calhoun, G.L., Ward, V.B.R., and Ruff, H.A. (2011). Performance-based adaptive automation for supervisory control. In Proceedings of the Human Factors and Ergonomics Society, pp. 2059-2063.

Chen, J.Y.C., and Barnes, M.J. 2012. Supervisory control of multiple robots: Effects of imperfect automation and individual differences. Human Factors (54:2), pp. 157-174.

Chenhall, R.H., 2003. Management control systems design within its organizational context: Findings from contingency-based research and directions for the future. Accounting, Organizations and Society, 28, 127–168.

Christoffersen, K. and Woods, D. D. (2002). How to make automated systems team players. In E. Salas (Eds.), Advances in Human Performance and Cognitive Engineering Research, Volume 2. St. Louis, MO, Elsevier Science, pp. 1–12.

Chong, V., 1996. Management accounting systems, task uncertainty and managerial performance: A research note. Accounting, Organizations, and Society 21, 415-421.

Clamann, M.P., Wright, M.C., Kaber, D.B., 2002. Comparison of performance effects of adaptive automation applied to various stages of human–machine system information processing. In: Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society. Human Factors and Ergonomics Society, Santa Monica, CA, pp. 342–346.

Compeau, D., Marcolin, B., Kelley, H., and Higgins, C. 2012. "Research commentary - Generalizability of information systems research using student subjects - A reflection on our practices and recommendations for future research," Information Systems Research (23:4), pp. 1093-1109.

Cohen, J, Cohen, P, West, SG and Aiken, LS. 2003. Applied multiple regression/correlation analysis for the behavioral sciences. Third Edition. Routledge:New York.

Cosenzo, K., Parasuraman, R., Novak, A., and Barnes, M. Army Research Laboratory, Human Research and Engineering Directorate. (2006). Implementation of automation for control of robotic systems (ARL-TR- 3808).

Cordery, J. L., Morrison, D., Wright, B. M., and Wall, T. D. (2010). The impact of autonomy and task uncertainty on team performance: A longitudinal field study. Journal of Organizational Behavior, 31, 240–258, http://dx.doi.org/10.1002/job.657.

Cuevas, H.M., Fiore, S.M., Caldwell, B.S., Strater, L., 2007. Augmenting team cognition in human–automation teams performing in complex operational environments. Aviation, Space, and Environmental Medicine 78, B63–B70.

Dekker, S. A., and Woods, D. D. (2002). MABA-MABA or abracadabra? Progress on human–automation co-ordination. Cognition, Technology and Work, 1, 240–244.

deCharms, R. 1968. Personal causation: The internal affective determinants of behavior. New York: Academic Press, 1968.

Endsley, M. R., and Kiris, E. O. 1995. The out-of-the-loop performance problem and level of control in automation. Human Factors, 37(2), 381-394. doi: 10.1518/001872095779064555.

Endsley and Kaber, 1999. Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 1999, Vol. 42, No. 3, 462 - 492.

Endsley, M. 1987, The application of human factors to the development of expert systems for advanced cockpits, Proceedings of the Human Factors Society 31st Annual Meeting (Santa Monica, CA: Human Factors Society), 1388 -1392.

Endsley, M. R. (1995a). Towards a new paradigm for automation: Designing for situation awareness. In: Proceedings of the 6th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems, pp. 421-426. Cambridge, MA: MIT.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. Human Factors, 37(1), 32-64.

Endsley, M. R. and Kaber, D. B. 1999. Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 1999, VOL. 42, NO. 3, 462 - 492.

Endsley, M.R. 2017. From here to autonomy: Lessons learned from human-automation research. Human Factors. 59, 5--27 (2017)

Endsley and Kiris, 1995. The out-of-the-loop performance problem and level of control in automation. Human Factors, 1995,37(2), pp 381-394.

Endsley and Kaber, 1999. Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 1999, VOL. 42, NO. 3, 462 - 492.

Faraj S., and Sproull L., (2000) Coordinating expertise in software development teams. Management Science 46(12):1554-1568. https://doi.org/10.1287/mnsc.46.12.1554.12072.

Faulkner and Badurdeen, 2014. Sustainable value stream mapping (Sus-VSM): methodology to visualize and assess manufacturing sustainability performance. Journal of Cleaner Production. 85:15 (2014), pp. 8-18.

Feigh K.M., Dorneich M. C and Hayes C.C. 2012. Toward a characterization of adaptive systems : A framework for researchers and system designers. Human Factors: The Journal of the Human Factors and Ergonomics Society 2012 54: 1008 originally.

Feng J, Spence I and Pratt J. 2007. Playing an action video game reduces gender differences in spatial cognition. Psychological Science, Vol. 18, No. 10, pp. 850-855.

Fernández-Caramés, T.M.; Fraga-Lamas, P. A review on human-centered IoT-connected smart labels for the Industry 4.0. IEEE Access 2018, 6, 25939–25957.

Gombolay, M., Jensen, R., Stigile, J., Son, S.H. and Shah, J., 2016. Apprenticeship scheduling: Learning to schedule from human experts. AAAI Press/International Joint Conferences on Artificial Intelligence.

Greenberger, D. B., and Strasser, S. 1986. The development and application of a model of personal control in organizations. Academy of Management Review, 11: 164-177.

Greenberger. D. B., Strasser, S., Cummings, L. L., and Dunham, R. B. (1989). The impact of personal control on performance and satisfaction. Organizational Behavior and Human Decision Processes, 43, 29-51.

Green CS and Bavelier D. 2003. Action video game modifies visual selective attention. Nature 423 (2003), 534-537, Letters to Nature.

Galbraith, J. R. 1974. Organization design: An information processing view. Interfaces, 4, 28-36.

Graham, J.R., Harvey, C.R., Huang, H., 2009. Investor competence, trading frequency, and home bias. Manage. Sci. 55 (7), 1094–1106.

Groover, M.P., 2018. Automation. Encyclopædia Britannica. Encyclopædia Britannica, inc.

Groom, V., and Nass, C. (2007). Can robots be teammates? Benchmarks in human-robot teams. Interaction Studies, 8, 483–500.

Hancock, P. A., and Scallen, S. F. 1996. The future of function allocation. Ergonomics in Design (4:4), pp. 24–29.

Harris, W.C., Hancock, P.A., and Arthur, E.J. 1993. The effect of taskload projection on automation use, performance, and workload. In Proceedings of the 7th International Symposium on Aviation Psychology. pp.178-184. Columbus, OH.

Hart S.G. 2006. "NASA-task load index (NASA-TLX); 20 years later," in Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society (HFES '06), pp. 904–908, October 2006.

Hartmann, F.G.H., 2000. The appropriateness of RAPM: Toward the further development of theory. Accounting, Organizations and Society, 25, 451–482.

Hartmann F.G.H. and Maas V.S. (2011) The effects of uncertainty on the roles of controllers and budgets: An exploratory study, Accounting and Business Research, 41:5, 439-458, DOI: 10.1080/00014788.2011.597656.

Haverstock, N.B., Ruthig, J.C., and Chipperfield, J.G. (2020). Primary and secondary control strategies and psychological well-being among familial caregivers of older adults with dementia, The Journal of Social Psychology, 160:1, 61-74, DOI: 10.1080/00224545.2019.1592095.

Hayes A. F. (2015) An index and test of linear moderated mediation, multivariate behavioral research, 50:1, 1-22, DOI: 10.1080/00273171.2014.962683.

Hayes, A. F. (2018). Introduction to mediation, moderation, and conditional process analysis. (2nd Ed.). New York: The Guilford Press.

Hirst, M. K. (1983). Reliance on accounting performance measures, task uncertainty and dysfunctional behaviour. Journal of Accounting Research, 21(2), 596–605.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision making. Science 310, 1680–1683.

Hilburn B., Jorna P. G., Byrne E. A., and Parasuraman R. 1997. The effect of adaptive air traffic control (ATC) decision aiding on controller mental workload, in Human-Automation Interaction: Research and Practice, M. Mouloua and J. Koonce, Eds. Mahwah, NJ: Erlbaum, pp. 84–91.

Jackson, P. 1986. Introduction to expert systems, Addison-Wesley Publishing Company.

Janssen C.P., Donker S.F., Brumby D.P., and Kun A.L. (2019) History and future of human–automation interaction. International Journal of Human-Computer Studies.131: 99–107.

Kaber, D. B. 1996, The effect of level of automation and adaptive automation on performance in dynamic control environments, doctoral dissertation, Texas Tech University, Lubbock, TX.

Kaber and Perry 2006. Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. International Journal of Industrial Ergonomics, V36, pp 447–462.

Kaber, D. B. and 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.

Kidwell, B., Calhoun, G.L., Ruff, H.A., Parasuraman, R., 2012. Adaptable and adaptive automation for supervisory control of multiple autonomous vehicles. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 56. SAGE Publications, Boston, pp. 428–432.

Kaber, D. B., Omal, E., and Endsley, M. (1999). Level of automation effects on telerobot performance and human operator situation awareness and subjective workload. In M. Scerbo and M. Mouloua (Eds.), Automation technology and human performance: Current research and trends (pp. 165–170). Mahwah, NJ: Erlbaum.

Kaber and Riley. 1999. Adaptive automation of a dynamic control task based on the secondary task workload measurement. International Journal of Cognitive Ergonomics (3:3). pp. 169-187.

Kaber, D. B. and 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.

Kaber, D. B., Wright, M. C., Prinzel, L. P., and Clamann, M. P. 2005. Adaptive automation of human-machine system information processing functions. Human Factors (47), pp. 730-741.

Kaber D, Perry C, Segall N, McClernon C, and Prinzel L., 2006. Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. International Journal of Industrial Ergonomics (36), pp. 447–462.

Karle JW, Watter S, and Shedden JM. 2010. Task switching in video game players: Benefits of selective attention but not resistance to proactive interference. Acta Psychologica (134:1), pp. 70-78.

Kessel, C. J., and Wickens, C. D. (1982). The transfer of failure detection skills between monitoring and controlling dynamic systems. Human Factors, 24, 49-60.

Kramer, S. H. (1999). When are two heads better than one? The role of expertise and task difficulty in individuals, statistical group, and interacting group problem solving. Dissertation Abstracts International: Section B: the Sciences and Engineering, 60, 1350.

Kringelbach M. L. (2005). The orbitofrontal cortex: Linking reward to hedonic experience. Nature

Reviews Neuroscience. 6 (9): 691–702. doi:10.1038/nrn1747.

Kozlowski, S. W. J., and Bell, B. S. 2003. Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, and R. J. Klimoski (Eds.), Handbook of Psychology: Industrial and Organizational Psychology, Vol. 12: 333-375. London: Wiley.

Lamberti D.M. and Wallace W.A., 1990. Intelligent interface design: An empirical assessment of knowledge presentation in expert systems. MIS Quarterly, Vol. 14, No. 3 (Sep., 1990), pp. 279-311.

Land, S.A., Malin, J.T., Thronesbery, C., Schreckenghost, D.L., 1995. Making intelligent systems team players: A guide to developing intelligent monitoring systems (NASA Technical Memorandum 104807). NASA Johnson Space Center, Houston, TX.

Langan-Fox, J., Canty, J.M., Sankey, M.J., 2009. Human-automation teams and adaptable control for future air traffic management. International Journal of Industrial Ergonomics 39, 894-903.

Lee, M.K., 2018. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. Big Data and Society, 5(1).

Lee, J. D., and See, K. A. 2004. Trust in automation: Designing for appropriate reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society (46:1), pp. 50–80 (doi: 10.1518/hfes.46.1.50.30392).

Lewis, K. 2004. Knowledge and performance in knowledge-worker teams: A longitudinal study of transactive memory systems. Management Science, 50: 1519-1533.

Lewis, K., and Herndon, B. (2011). Transactive memory systems: Current issues and future research directions. Organization Science, 22, 1254–1265. http://dx.doi.org/10.1287/orsc.1110.0647.

Li HY, 2013. Supporting human-automation collaboration through dynamic function allocation: The case of space teleoperation. University of Michigan. Dissertation.

Li, H., Sarter, N., Wickens, C., and Sebok, A. (2013, September). Supporting human-automation collaboration through dynamic function allocation: The case of space teleoperation. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 57, No. 1, pp. 359-363). SAGE Publications.

Libby R, Trotman K.T, Zimmer I. 1987. Member variation, recognition of expertise, and group performance. Journal of Applied Psychology. Vol 72. No 1, pp 81-87.

Littlepage G, Robison W, Reddington K. 1997. Effects of task experience and group experience on group performance, member ability, and recognition of expertise. Organizational Behavior and Human Decision Processes. 69:133–47

Littlepage GE, Mueller AL. 1997. Recognition and utilization of expertise in problem solving groups: Expert characteristics and behavior. Group Dynamic. 1(4):324–28.

Littlepage, G.W., Schmidt, E.W. Whisler, A.G. and Frost. 1995. An input-process-output analysis of influence and performance in problem-solving groups. Journal of Personality and Social Psychology. 69 877–889.

Longo, F., Nicoletti, L., and Padovano, A. (2017). Smart operators in industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context. Computers and Industrial Engineering, 113, 144–159.

Malin, J. T., and Schreckenghost, D.L. (1992). Making intelligent systems team players: Overview for designers. NASA Technical Memorandum 104751, Johnson Space Center, Houston, TX.

Mani D., Barua A., and Whinston A. 2010. An empirical analysis of the impact of information capabilities design on business process outsourcing performance. MIS Quarterly, Vol. 34, No. 1 (March 2010), pp. 39-62.

Mathieu, J. E., Maynard, M. T., Rapp, T., and Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. Journal of Management, 34, 410-476.

Miller, C. and Hannen, M. (1999). The rotorcraft pilot's associate: Design and evaluation of an intelligent user interface for cockpit information management. Knowledge Based Systems, 12, 443–456.

Miller C., Funk H., Goldman R., and Wu P. 2004. A 'playbook' for variable autonomy control of multiple, heterogeneous unmanned air vehicles. In Human Performance, Situation Awareness, and Automation, D. A. Vicenzi, M. Mouloua, and P. A. Hancock, Eds. Mahwah, NJ: Erlbaum, 2004, pp. 210–214

Miller, C.A., Funk, H., Goldman, R., Meisner, J., and Wu, P. 2005. Implications of adaptive vs. adaptable UIs on decision making: Why "automated adaptiveness" is not always the right answer. Proceedings of the 1st International Conference on Augmented Cognition, Las Vegas.

Miller and Parasuraman, 2007. Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. Human Factors (49:1). pp. 57–75.

Miller, Pelican, and Goldman, 2000. "Tasking" interfaces for flexible interaction with automation: Keeping the operator in control." In Proc. Conf. Human Interaction with Complex Systems, Urbana-Champaign, IL, pp. 188-202.

Moray, N. (1986). Monitoring behavior and supervisory control. In K. Baff (Ed.), Handbook of Perception and Human Performance (pp. 40/1-40/51). New York: Wiley.

Moray, N., T. Inagaki, and M. Itoh. 2000. Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. Journal of Experimental Psychology: Applied 6: 44–58.

Mushtaq F, Bland AR, Schaefer A. 2011. Uncertainty and cognitive control. Frontiers in Psychology.2:249.

National Transportation Safety Board. (1990). Aircraft accidents report: USAIR, Inc., Boeing 737-400, LaGuardia Airport, Flushing New York, September 20, 1989 (NTSB/AAR-90-03). Washington, DC: Author.

Neale, M. A., and Northcraft, G. B. (1986). Experts, amateurs, and refrigerators: Comparing expert and amateur negotiators in a novel task. Organizational Behavior and Human Decision Processes, 38, 305–317.

Nguyen, T. D. and Krüger , J. 2017. Human-centered automation to simplify the path to social and economic sustainability. In R. S. Stark, Sustainable Manufacturing: Challenges, Solutions and Implementation Perspectives (pp. 85-95). Cham, Switzerland: Springer Open.

Ntuen, C A. and Park, E. H. (1988). Human factors issues in teleoperated systems. In: Ergonomics of Hybrid Automated Systems I (W. Karwowski, H. R. Parsaei and M. R. Wilhelm, Eds.),pp. 230-210. Amsterdam: Elsevier.

Ogletree, S. M., and Drake, R. 2007. College students' video game participation and perceptions: Gender differences and implications. Sex Roles (56) pp. 537–542.

Olson, W. A., and Sarter, N. B. 2000, Automation management strategies: Pilot preferences and operational experiences. International Journal of Aviation Psychology (10:4), pp. 327–341.

O'Doherty, J., Critchley, H., Deichmann, R., and Dolan, R. J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J. Neurosci. 23, 7931–7939.

Pacaux-Lemoine, M., Trentesaux, D., Zambrano Rey, G., and Millot, P. (2017). Designing intelligent manufacturing systems through human-machine cooperation principles: A human-centered approach. Computers and Industrial Engineering, 111, 581–595. http://dx.doi.org/10.1016/j.cie.2017.05.014. ISSN 0360-8352.

Parasuraman, R., Barnes, M., and Cosenzo, K. (2007). Adaptive automation for human-robot teaming in future command and control systems. International Journal of Command and Control, 1(2), 43–68.

Parasuraman, R. (1993). Effects of adaptive function allocation on human performance. In D. J. Garland and J. A. Wise (eds.), Human Factors and Advanced Aviation Technologies (pp. 147-158). Daytona, Beach, FL: Embry-Riddle Aeronautical University Press.

Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on Systems Man and Cybernetics Part A: Systems and Humans, 30(3), 286-297. doi: 10.1109/3468.844354.

Parasuraman and Wickens. (2008). Humans: Still vital after all these years of automation. Human Factors, Vol. 50, No. 3. pp. 511–520.

Parasuraman, R., Bahri, T., Deaton, J., Morrison, J., and Barnes, M. (1992). Theory and design of adaptive automation in aviation systems (Progress Rep. No. NAWCADWAR-92033-60). Warmister, PA: Naval Air Warfare Center, Aircraft Division.

Parasuraman, R., Mouloua, M., and Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. Human Factors, 38(4), 665-679.

Parasuraman, R., Bahri, T., Deaton, J., Morrison, J., and Barnes, M. (1992). Theory and design of adaptive automation in aviation systems (Progress Rep. No. NAWCADWAR-92033-60). Warmister, PA: Naval Air Warfare Center, Aircraft Division.

Parasuraman, R., S. Galster, P. Squire, H. Furukawa, and C.A. Miller. 2005. A flexible delegation interface enhances system performance in human supervision of multiple autonomous robots: Empirical studies with RoboFlag. IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans 35: 481–493.

Parasuraman, Molloy, and Singh. 1993. Performance consequences of automation-induced 'complacency.' The International Journal of Aviation Psychology, Vol. 3, pp. 1–23, 1993.

Pearce, J. A., and Ravlin, E. C. (1987). The design and activation of self-regulating work groups. Human Relations, 40, 751–782.

Pope, A. T., Comstock, R. J., Bartolome, D. S., Bogart, E. H., and Burdette, D. W. (1994). Biocybernetic system validates index of operator engagement in automated task. In M. Mouloua, and R. Parasuraman (Eds.), Human Performance in Automated Systems: Current Research and Trends (pp. 300-306). Hillsdale, NJ: Lawrence Erlbaum.

Power, C., Denisova, A., Papaioannou, T. and Cairns, P. (2017). Measuring uncertainty in games: Design and preliminary validation. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. (pp. 2839-2845). New York, USA: ACM Press. ISBN 9781450346566

Power C.D., Cairns P.A., Denisova A., Papaioannou T., and Gultrom R. 2018. Lost at the edge of uncertainty: Measuring player experience in digital games. International Journal of Human-Computer Interaction. pp. 1-14. ISSN 1532-7590.

Punnett L, Wegman D. 2004. Work-related musculoskeletal disorders: The epidemiologic evidence and the debate. Journal of Electromyography and Kinesiology.14:13–23.

Parasuraman, R., S. Galster, P. Squire, H. Furukawa, and C.A. Miller. 2005. A flexible delegation interface enhances system performance in human supervision of multiple autonomous robots: Empirical

studies with RoboFlag. IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans 35: 481–493.

Quaiser-Pohl, C., Geiser, C., and Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. Personality and Individual Differences, 40(3), 609–619. doi:10.1016/ j.paid.2005.07.015.

Ren, Y., and Argote, L. (2011). Transactive memory systems 1985– 2010: An integrative framework of key dimensions, antecedents, and consequences. The Academy of Management Annals, 5, 189–229. http://dx.doi.org/10.1080/19416520.2011.590300.

Rothbaum, F. M., Weisz, J. R., and Snyder, S. S. (1982) Changing the world and changing the self: A two-process model of perceived control. Journal of Personality and Social Psychology, 42, 5-37.

Rouse W. B. 1988. Adaptive aiding for human/computer control. Human Factors, vol. 30, pp. 431–438.

Rouse, W. B., Geddes, N. D., and Curry, R. E. (1986). An architecture for intelligent interfaces: Outline of an approach to supporting operators of complex systems. Human Computer Interaction, 3, 87-122.

Ruff, H., Calhoun, G., Frost, E., Behymer, K., and Bartik, J. 2018. Comparison of adaptive, adaptable, and hybrid automation for surveillance task completion in a multitask environment. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (62), pp. 155–159. SAGE Publications Sage CA: Los Angeles, CA.

Santiago-Espada, Myer, Latorella and Comstock. 2011. The Multi-Attribute Task Battery II (MATB-II) software for human performance and workload research: A user's guide. NASA/TM–2011-217164.

Sasaki, J. Y., and Kim, H. S. (2010, December 27). At the intersection of culture and religion: A cultural analysis of religion's implications for secondary control and social affiliation. Journal of Personality and Social Psychology. Advance online publication. doi: 10.1037/a0021849.

Sauer, J., and Chavaillaz A. 2017. The use of adaptable automation: Effects of extended skill lay-off and changes in system reliability. Applied Ergonomics 58: 471–481.

Sauer, Kao and Wastell. (2012). A comparison of adaptive and adaptable automation under different levels of environmental stress. Ergonomics, 55(8): 840-853, 2012.

Scerbo, M. (2008). Adaptive automation. In R. Parasuraman and M. Rizzo (Eds.), Neuroergonomics: The Brain at Work (pp. 238–252). New York: Oxford University Press.

Scerbo, M. (2001). Adaptive automation. In W. Karwowski (Ed.), International Encyclopedia of Ergonomics and Human Factors (pp. 1077–1079). London: Taylor and Francis.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman and M. Mouloua (Eds.), Human Performance in Automated Systems: Theory and Applications (pp. 37-64). Mahwah, NJ: Lawrence Erlbaum Associates.

Sheridan T.B. and Parasuraman R.P., 2006. Human-automation interaction. Reviews of human factors and ergonomics 1(1):89-129. DOI: 10.1518/155723405783703082.

Sheridan, T. B., and Verplank, W. L. (1978). Human and computer control of undersea teleoperators. Man-Machine Systems Laboratory Report. Cambridge, MA: MIT.

Shiff, B. 1983. An experimental study of the human-computer interface in process control. Unpublished thesis, University of Toronto, Toronto.

Shively, R.J., Lachter, J., Brandt, S.L., Matessa, M., Battiste, V., Johnson, W. 2017. Why human-autonomy teaming? In: Proceedings of the 8th International Conference on Applied Human Factors and Ergonomics, this edition.

Squire, P. N., and Parasuraman, R. (2010). Effects of automation and task load on task switching during human supervision of multiple semiautonomous robots in a dynamic environment. Ergonomics, 53(8), 951-961.

Stasser, G., Stewart D. D., and Wittenbaum G. M. 1995. Expert roles and information exchange during discussion: The importance of knowing who knows what. Journal of Experimental Social Psychology, 31:244-265.

Stewart, D. D., and G. Stasser 1995 Expert role assignment and information sampling during collective recall and decision making. Journal of Personality and Social Psychology, 69: 619-628.

Steinhauser, N.B., Hancock P.A. 2009. Design principles for adaptive automation and aiding. Ergonomics in Design. pp. 6-10.

Terlecki, M. S., Brown, J., Harner-Steciw, L., Irvin-Hannum, J., Marchetto-Ryan, N., Ruhl, L., and Wiggins, J. 2011. Sex differences and similarities in video game experience, preferences, and self-efficacy: Implications for the gaming industry. Current Psychology, (30:1), pp. 22-33.

The US Department of Commerce, 2010. The International Trade Administration and The U.S. Department of Commerce's definition for sustainable manufacturing.

Thompson, S. C. (1981) Will it hurt less if I can control it? A complex answer to a simple question. Psychological Bulletin, 90, 89-101.

Ujma, M., & Ingram, T. (2019). Perception of ability-motivation-opportunity oriented HRM practices and organizational commitment: The role of task uncertainty. Journal of Entrepreneurship, Management and Innovation, 15(1), 139-162.

Van de Ven, A. H. and Delbecq. A. L. 1974. A task contingent model of work-unit structure, Administrative Science Quarterly (June 1974) pp. 183-197.

Vagia, M., Transeth, A.A., Fjerdingen, S.A., 2016. A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? Applied Ergonomics. 53, 190–202.

Visser E. De, Shaw T., Mohamed-Ameen, A., and Parasuraman, R. 2010. Modeling human-automation team performance in networked systems: Individual differences in working memory count. Proceedings of the Human Factors and Ergonomics Society, San Francisco, pp 1087-1091.

Wall, T. D., Cordery, J. L., and Clegg, C. W. 2002. Empowerment, performance, and operational uncertainty. Applied Psychology: An International Review, 51: 146-169.

Warm, J.S., Parasuraman, R., Matthews, G., 2008. Vigilance requires hard mental work and is stressful. Human Factors 50 (3), 433–441.

Wickens, C. D. (2002). Multiple resources and performance prediction. Theoretical Issues in Ergonomics Science, 3(2), 159–177.

Wickens, C. D. (2008). Multiple resources and mental workload. Human Factors, 50, 449–455.

Wiener, E. L. (1988). Cockpit automation. In: Human Factors in Aviation (E. L. Wiener and D. C Nagel, Eds.), pp. 433-459. San Diego, CA: Academic Press.

Wickens, C. D. and Kessel, C (1981). Failure detection in dynamic systems. In: Human Detection and Diagnosis of System Failures (J. Rasmussen and W. B. Rouse, Eds.). NewYork: Plenum.

Wickens, C. D. (1994). Designing for situation awareness and trust in automation. In Proceedings of the International Federation of Automatic Control (IFAC) Conference (pp. 174–179). Baden- Baden, Germany: IFAC.

Wittenbaum, G. M., Vaughan, S. I., and Stasser, G. 1998. Coordination in task-performing groups. In R. S. Tindale and L. Heath (Eds.), Theory and Research on Small Groups: 177-204. New York: Plenum.

Wijngaards N., Kempen M., Smit A., and Nieuwenhuis K. 2006. Towards sustained team effectiveness, in coordination, organizations, institutions, and norms in multi-agent systems. In O. Boissier, J. Padget, V. Dignum, G. Lindemann, E. Matson, S. Ossowski, J. S. Sichman, and J. Vázquez-Salceda, (Eds.) Springer Berlin Heidelberg, 2006, pp. 35–47.