A Study on Human Trust in Machine under Supervisory Control

September　２０２０

Lee Jieun

A Study on Human Trust in Machine under Supervisory Control

Graduate School of Systems and Information Engineering
University of Tsukuba

September　２０２０

Lee Jieun

ABSTRACT

Human trust in machines plays a vital role in determining the use of automation in the context of modern society where automated machines are widespread, with increasing the complexity of automation. Despite of the advances, newly developed automated machines still have a requirement for human in supervisory position to ensure high performance and safety. As yet, the mechanisms underlying how users' trust in unfamiliar machines grows in the current context remain unclear. Accordingly, the objective of this study was to investigate factors influencing the development of trust in machines with regard to current contexts as reinterpret a theoretical model conceptualized by Muir (1994) which trust develops faith, then dependability, and finally predictability in a simulated supervisory control task. A series of experiments were implemented to address trust development for users in general populations with issues of initially formed trust and failures that system occurs in supervisory control situations by replicating process control tasks and applying the model to driving automation domain under supervisory control. In STUDY I, Muir and Moray's (1996) experimental tasks were fully replicated with the newly reprogrammed raw-milk pasteurization plant and 12 male students. STUDY II attempted to the framework to partial vehicle automation that demands human drivers' supervisory control regarding operators' knowledge levels in an initial stage and the types of system failures. STUDY III revisited Muir and Moray's (1996) process control tasks with only engineering major students. The results have implications in terms of trust development that (i) dependability prompted initial trust in machines for untrained users regardless of automation domains and the supervisory roles in task allocation, (ii) in general, dependability was consistently the best predictor of trust throughout human interaction with machines, disconfirming the original findings of Muir and Moray (1996), (iii) the occurrence and type of system failures can have relatively small impacts on trust in process control tasks, and users are likely to focus the occurrence of system failure relevant to critical risk rather than which type of failure they encountered. This series of experiments reveals that modern society calls dependability as the best attributor of trust in machines. With the findings, the present results confirm previous findings that human trust in machines increases through human-machine interaction, and decrease in trust due to system fault can be recovered by subsequent experience of error-free machine. This thesis research contributes approaches for modern automation designers to gain and establish users' trust when general users attempt to interact unfamiliar machines and remains implications in the context of information building appropriate levels of trust.

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisor. All my researches from my junior to PhD would not have been possible without Prof. Makoto Itoh's generous support. His excellent insight and advice have been improving my research ability, patience, and knowledge to survive as an independent researcher, further, a sense of humanity.

I greatly appreciate Dr. Yusuke Yamani at Old Dominion University to his tremendous effort on the collaboration for human-machine trust study and warm advices on future career. I am also grateful to Prof. Genya Abe at Japanese Automobile Research Institute for his practical suggestion into study on trust in vehicle automation. Thanks to Dr. Neil Millar for kind suggestion to develop academic writing and presentation skill. Thanks to Prof. Toshiyuki Inagaki for his encouragement, critical questions, and a special present.

The members of Laboratory Cognitive Systems Science should be acknowledged for their support and invaluable feedbacks on this study at lab seminars. I especially thank two great counsellors, Dr. Huiping Zhou and Yumi Ohhama, for letting me lean on their warmth privately as well as having constructive discussions on research. Thanks to Sugiyama Taisei for developing the pasteurizer in MATLAB. Thanks to my friends for encouragement with kind words. I express big thanks to my best friend M.S. Eunsoo Kwak for her tremendous support, with wishing our wonderful friendship.

Finally, I should articulate my heartfelt appreciation to my family. Especially thanks to my dear mom and sister, Jeonga Kim and Jihye Lee, for their unconditional love and enduring support throughout my studies.

Jieun Lee

Tsukuba, June 2020

CONTENTS

List of Tables

LIST OF FIGURES

LIST OF APPENDICES

Chapter 1. INTRODUCTION

1.1 Background

Automation executes functions that human operators has been performed, with supporting human information processing. Automation has been developed for a wide range of systems with ultimate aims to diminish operators' workload, to improve performance, and to enhance human safety by supporting human information processing, to be specific, sensing, perceiving, decision making, and execution (Parasuraman, Sheridan, & Wickens, 2000). Usage of automated systems has been dramatically increasing for the past two decades. Nowadays, automated systems are increasingly prevalent in modern society for supporting human life across a variety of domains, including the internet of things devices (Lopatovska et al., 2019), surface transportation (Society of Automotive Engineers, 2018) and healthcare (Mitzner et al., 2018). As an example of highly automated systems for supporting daily life, the automotive industry, over 14 companies, including car manufacturers and technology companies has accelerated the realization of autonomous and semi-autonomous vehicles in the market – e.g., Tesla Model S and Volvo series 90. Public in modern societies is, or soon will be, surrounded by highly automated systems that can have significant consequences to their lives including surgical robots (Gerber et al., 2020), peacekeeping robots (Bliss et al., 2020), autonomous vehicles (Yamani & Horrey, 2018), and unmanned aerial vehicles such as Uber Elevate to name a few.

Automation has been mainly utilized for designated tasks, appearing in work environments. Manned autoflight systems are capable of yielding high levels of performance and surveilling remote locations with less operation by commercial pilots (Billings, 1997) also automated systems fully run nuclear power plants and public transit systems (Skjerve & Skraaning, 2004). The widespread of automation in the workplace leads human-machine collaboration to accomplish same tasks. In this situation, as a supervisory controller, human should make decisions to best interact with automated machines in information-rich environments. Supervisory control is defined as "the human activity involved in initiating, monitoring, and adjusting processes in systems that are otherwise automatically controlled (Sheridan & Hennessy, 1984)." Supervisory controllers have to interact with the system through different levels of manual and automatic control. That is, human operators need adaptive function allocation between themselves and automatic controllers of systems. As supervisory control situations require human operators to detect and analyse failures in systems and intervene the process in unfamiliar situations, adaptive supervision based on deep understanding of the process and control capability of system is essential for the supervisory controller. Therefore, operators in professional domains, such as aviation or maritime transportation, undertake an extensive mandatory training to accomplish tasks in supervisory control situations.

The supervisory controller should adequately understand systems for better human-machine interaction. If deep understanding is not prepared, it is impossible to predict upcoming situations in work environments, resulting in poor decision making. Expert operators may, for example, form an accurate mental model of the automated systems, achieved through extensive training under supervisory control, thus allowing them to achieve high levels of situation awareness (e.g. Endsley, 1995). The high levels of situation awareness are important to comprehend the meaning and environment of situation and carry out appropriate actions in operating automated machines. That is, accumulating knowledges about systems through professional training may improve operators' ability to judge encountering circumstance and to apply an appropriate decision making strategy for performance incline (Damos & Wickens, 1980; Gopher et al., 1989; Kramer, Larish, & Strayer, 1995).

Recent technological advancements provided human-automation interaction schemes in which the human performs supervisory roles (e.g., Gao & Lee, 2006; Cummings & Clare, 2015; Xu & Dudek, 2015). The role of supervisory controller has shifted from an active controller to a passive monitor of highly automated systems (e.g., Sheridan, 1970; Parasuraman et al., 1996; Metzger & Parasuraman, 2001; Hoogendoom et al., 2014). Despite the shift, humans should play a fundamental role in supervisory controllers, such as detection of automation failures and intervention in control if necessary. Given that increasing availability and decreasing cost of automated systems will allow general users to begin interacting with these highly developed systems with little to no training, the general users may not figure out how to interact with systems under supervisory control, leading to inappropriate usage of automated systems (Parasuraman & Riley, 1997). To address this concern, it is critical to consider what factors influence the successful human-automation interaction under supervisory control.

One factor shown to enormously influence human-automation interaction in the context of supervisory control is trust (Sheridan, 1992; Halpin et al., 1993). Operators' trust in automation is playing a vital role in guiding their allocation strategy. In a supervisory control situation, several variables, such as risk or mental workload, may have relatively large impacts on the operator's choice between automatic and manual control in comparison with trust (Riley, 1994). However, trust has been addressed as a prominent factor to mediate the relationship between the supervisory controller and the subordinate machines (Lee & Moray, 1992; Muir & Moray, 1996; Parasuraman & Riley, 1997).

1.2 Human Trust in Machines

Most existing researches on human trust in machines have been expanded based on theories of general human trust in psychological domains (e.g., Barber, 1983; Rotter, 1967; Rempel et al., 1985). In general, the term 'trust' has been used to refer to the personality characteristics that make a person 'trusting' and 'trustworthy' (e.g., Deutsch, 1960). Historically, diverse paradigms have been employed to define trust because it cannot be elaborated by only one concept. For instance, Deutsch (1960) outlined trust as a behavioural result or state of vulnerability or risk based on intention and ability, and Mayer et al. (1995) regarded trust as a willingness to take vulnerability in a similar way with Deutsch (1960). However, Barber (1983) viewed trust as a belief or attitude toward others. These assorted attributions of general human trust have been applied to examine human trust in machines. Reeves and Nass (1996) stated that human reaction to technology and computers resembles reaction to human collaborators. Many studies have attempted to establish concepts of trust between human and automation following frameworks of general human trust and explored trust is a key contributor to establish effective relationship between human and human as well as human and machines even though the formation of trust between those is not completely identical (e.g., Madhavan & Wiegmann, 2007).

Trust in automation is defined in diverse ways from several human factors researchers. Riley (1994) defined trust as the operators' subjective estimate of the probability that the automation will be correct in the next action or decision it makes. Bentley et al. (1995) considered trust with familiarity as a process that is partially a function of technology. After the publication of Lee and See (2004), approximately all studies on trust in automation have followed their definition which is "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (Lee & See 2004, p. 54). A variety of automation characteristics shape human operators' trust (Hoff & Bashir, 2015). Specifically, operators' trust towards automation can vary due to factors such as instruction about an automated system and reliability of the system (Bliss et al., 1995; Chancey et al., 2017; Dixon & Wickens, 2006; Dzindolet et al., 2002; Körber et al., 2018a; Meyer, 2001; Sato et al., 2019). For example, in a study examining the impact of introductory trust information when using an automated driving system, researchers found that the manipulation led to different levels of driver reliance on automated driving systems, resulting in less attention of drivers in trust-prompted group compared to those of trust-unprompted group (Körber et al., 2018a). Additionally, reliability has proven to be an important basis for trust in several tasks, such as a pasteurization monitoring task (Wiegmann et al., 2001), a visual search task (Chancey et al., 2017), and a flight simulator (Sato et al., 2019) with cognitive support from various automated systems.

Trust is a multidimensional concept and dynamically changed by situation or operator experience (e.g., Sanchez, 2006; Manzey et al., 2012; Hergeth et al., 2017; Körber et al., 2018b; Hartwich et al., 2019). Similar to experience with automation, several factors may impact trust, such as operators' self-confidence (Lee & Moray, 1994; Moray et al., 2000) or system faults (Moray et al., 1995). Also, when considering general users who do not undergo an extensive training for operating automated systems, an interaction with automation in an early stage may be a critical factor shaping future trust and use of automated aids (Hoff & Bashir, 2015). Occurrence timing of automation error as well as degrees of difficulty in dealing with error during human-machine interaction are critical to trust as trust evolves and adapts over time along users' accumulating knowledge about systems (Lee & See, 2004; Parasuraman & Miller, 2004). In this sense, initial interaction with automated machines is likely to foster trust in automation compared to prior to automation experience (Hartwich et al., 2019). For example, automation error during human-machine interaction in an early stage leads decreased trust and decreased use of automated aids (Manzey et al., 2012). Therefore, examining the factors influencing initial trust is important for understanding and ultimately designing systems to support the development of appropriate human-machine trust in general users.

Operators' expectations of automation before interacting with automation are considered to be trust in automation, determining the use of automation. Operators' trust may be initiated if machines provide outcomes that operators anticipated (Muir, 1994; Cahour & Forzy, 2009). In related vein, imperfect part of automation can lead to disuse of automation when human expects perfect automation in advanced of interaction with machines (Dzindolet et al., 2002). Even when automation performs reliably, discrepancy between operators' expectation and the actual behaviour of machine can influence automation use (Rasmussen et al., 1994). This indicates that users who never interact with automation may not develop an adequate mental model of the automation and may over- or undertrust the system (Lee & See, 2004; Parasuraman & Riley, 1997; Wiener & Curry, 1988). Thus, understanding how trust evolves over time is particularly important for designers to support successful human-automation interaction especially for users who are relatively inexperienced with the automated system. Despite a myriad of literatures that investigated users' initial trust in automated machines of various domains, the definition of initial trust with respect to time point varies depending on what the study aimed to observe. That is, diverse views exist whether users' initial interaction with automation has increased or decreased degrees of human-machine trust, resulting in overtrust, mistrust, or distrust (e.g., Wiegmann et al., 2001).

Initial trust in machines can be explained with the three-layers of human-automation trust proposed by Hoff and Bashir (2015) in detail. The work by Hoff and Bashir (2015) summarized human-automation trust studies which published between 2002 and 2013 and provided three-layers in order to regard the development of trust (see Figure 1.1). First, dispositional trust refers an individual's enduring overall tendency to trust automation. Gender (Hillesheim et al., 2017), cultural background (e.g., nationality; Yamagishi & Yamagishi, 1994), age (Ho et al., 2005; Sanchez et al., 2004), and personality (Merritt & Ilgen, 2008) represent factors shaping dispositional trust. It seems to correspond to the concept of initial trust regarding that initial levels of trust are determined by individual and organizational contexts (Lee & See, 2004). Second, situational trust is framed as trust develops depending on encountering situations and contexts. Here, external (e.g., workload and perceived risk) and internal variabilities (e.g., self-confidence and mood) are components influencing situational trust. Hoff and Bashir (2015) regarded that context-dependent components of automation trust foster situational trust. Lastly, learned trust concept divided trust into before and after experience of automation. Pre-existing knowledge in advance of human-machine interaction forms initial learned trust, and system performance and design feature contribute to dynamic learned trust. To be specific, pre-existing knowledge refers users' knowledge relevant to the system, such as reputation of system or brand and prior exposure to similar technology. Design features (i.e., transparency or appearance) guide system performance components, including reliability (Madsen & Gregor, 2000), dependability (Muir & Moray, 1996; Merritt & Ilgen, 2008), or usefulness of system (Igbaria & Iivari, 1995), then the components construct dynamic learned trust.

As shown in Figure 1.1, dispositional trust, situational trust, and initial learned trust formed trust prior to interaction with machines (Hoff & Bashir, 2015). According to this theoretical model, trust development throughout human-machine interaction can be determined by factors consisting of dynamic learned trust. However, factors of system performance guiding dynamic learned trust also are embodied by other factors consisting of initial trust. For instance, dispositional elements are main factors shaping dependability of system (Muir, 1994).

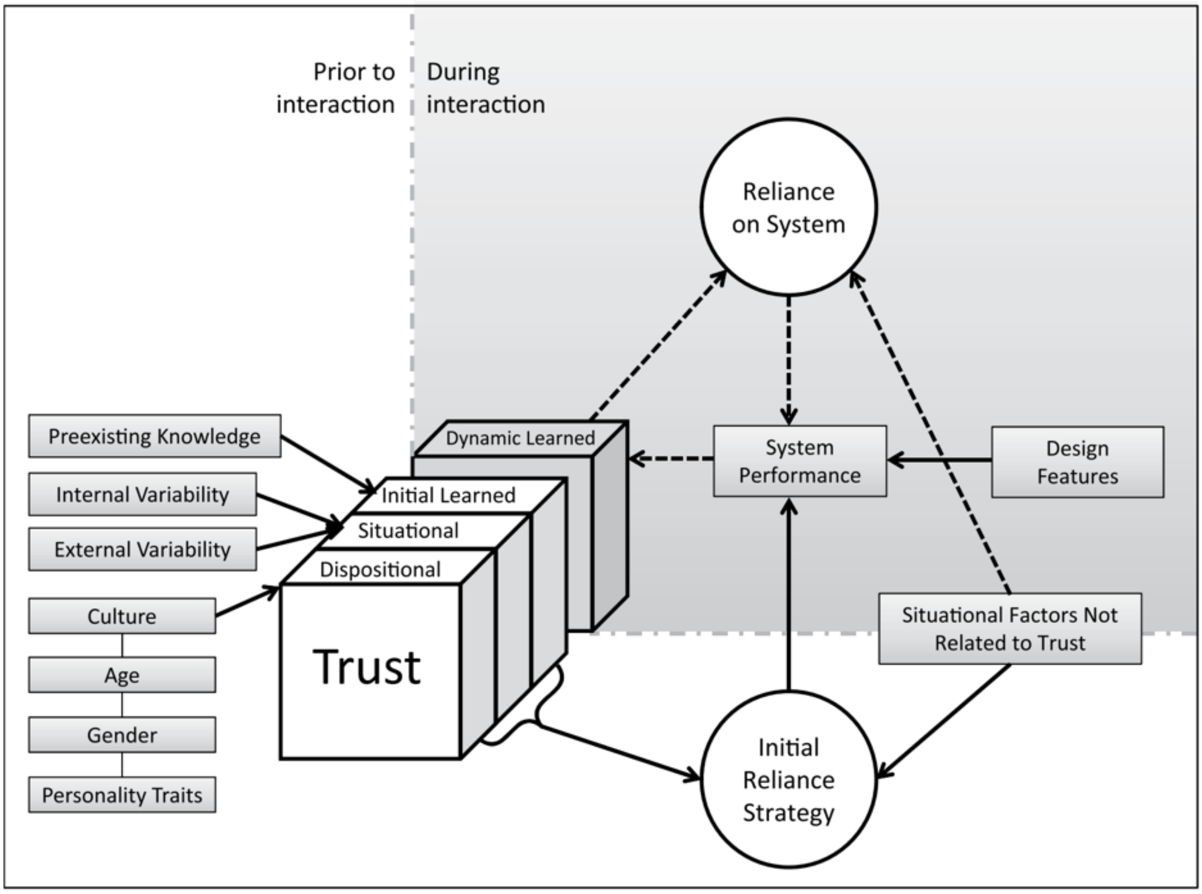


Figure 1.1: Full model of factors that influence trust in automation from Hoff & Bashir (2015). The dotted arrows represent factors that can change within the course of a single interaction.

In summary (see, Figure 1.2), dispositional quality, pre-existing knowledge, specific knowledge about system with intention to use machines, and first impression of system after human-machine interaction are considered to be factors guiding trust in automation until and including the first interaction with automated machines. As trust is a history-dependent variable depending on the prior behaviour of the trusted person and the information that is shared (Deutsch, 1958), the accumulation of knowledge in advance of interaction with the machine incorporates with the internal variability defined by Hoff and Bashir (2015), and it may have a great impact on shaping the first impression of machine. Looking into trust from initial interaction with automation until being skilled operators seems possible to provide a clue to a question: how changes of trends between past and present function involve human-machine trust. Given that identifying factor which accounts for trust is crucial for saving cost to obtain trust, automation designer should observe not only initial trust but also how trust evolves throughout the course of interaction between human and machine. After the first interaction, the formed first impression and system characteristics, such as design features and system fault, determine trust in machine under supervisory control. After interacting with machines, users are possible to be aware of system characteristics or which benefits are derived by automation over time. Various issues remain to have better understanding of designing automation, such as how operators respond system failure or which error leads disuse of automation.

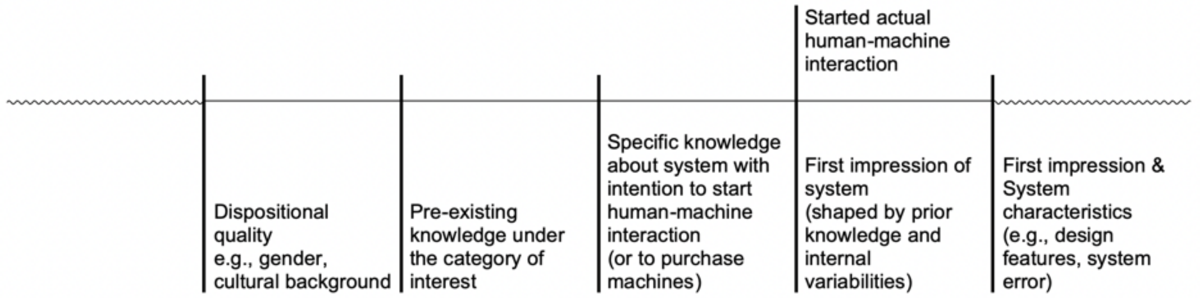


Figure 1.2: Factors which govern users' initial trust regarding time points.

As trust is resilient, one of automated system properties which are likely to influence trust development is system fault (e.g., Moray et al., 1995; Kelly et al., 2001). The occurrence of system fault negatively impacts human-machine trust in general. If once trust is broken due to system fault, rebuilding trust is slow (Moray et al., 1995). The process control experiment by Lee and Moray (1992) found that system fault first leads the lowest trust in the automation, but the trust can be rebuilt if the system fault continues in the course of interaction with the system. Further, Lee and Moray (1992) indicated that trust is less resilient because trust does not recover to the previous levels when the fault occurs. On the other hand, there are empirical evidences that trust fully recovers with subsequent exposure to system as well as continuous experience of systems despite system fault (e.g., Hergeth et al., 2017; Körber et al., 2018b). Kraus et al. (2019) examined changes of driver trust during interacting with conditionally automated vehicle that human drivers are responsible for not monitoring both external and internal environments but resuming vehicle controls when automated driving systems issue a request to intervene with vehicle control due to system failures. Driver trust continued to increase with experiencing error-free vehicle automation even though rare occurrence of the request decreased trust only right after the error experience. Trust recovery depends on magnitude and variable of system fault (Lee & Moray, 1992; Muir & Moray, 1996), where and when fault occurs in the system (Muir & Moray, 1996; Madhaven et al., 2006), and the amount of information about systems fault (Riley, 1996). These empirical findings have mainly addressed trust recovery with regard to the degree of trust which subjectively rated by users. So far, however, there has been little discussion about trust mechanism, such as which factor yields decrease in trust when users encounter fault. That is, empirical studies on what forms trust and how trust develops are relatively limited. The work of Muir and Moray (1996) is one of the most cited studies in the human-machine trust literature regarding this topic. Accordingly, the impact of a three-dimensional model of trust proposed by Muir and Moray (1996) has been particularly large. When considering the aforementioned issues, adapting the model of Muir and Moray (1996) may provide a valuable exploration to suggest new insight regarding different social contexts between three decades ago and the present.

1.3 Investigation of Muir's (1994) Framework

Muir's two-part work (Muir, 1994; Muir & Moray, 1996) has a symbolic meaning in automation trust studies because Muir (1994) first conceptualized that human-machine trust evolves similarly to interpersonal trust (i.e., Barber, 1983; Rempel et al., 1985; see Table 1.1). As aforementioned, current theories of human-machine trust depend on those of interpersonal trust. Part of them has been identically replicated findings by interpersonal trust theory, however, most researchers explore distinctions between human-machine and interpersonal trust. One famous study showing the distinction with an empirical finding is Muir and Moray's (1996) work.

Meaning of trust in machines. To examine the broader context of the meaning of trust, Muir (1987, 1994) outlined a framework of trust in automation based on Barber's (1983) definition. Barber (1983) viewed trust as the concept of expectation in society and described that human trust contains at least three different kinds of expectation in social relationships: persistence, competence, and persistence fiduciary responsibility. The persistence of the natural physical order, the natural biological order, and the moral social order is the most general expectation in order to generalize the expectation concept. Next is specific expectation of technically competent role performance which comes from the social relationship and belonging systems. The expectation of competent performance, such as patients' expectation of doctors' good operation performance, involves expert knowledge, technical facility, or everyday routine performance. Last, specific expectation that partners in an interaction will carry out their fiduciary obligations and responsibilities, that is, their duty in certain situations to place others' interests before their own. Power that relatively varies in social relationships and systems contribute the distribution of trust as fiduciary responsibility.

Muir (1987, 1994) attempted to capture the nature of trust in automation by incorporating previous social psychology research with engineering perspectives. She reinterpreted his concepts to apply it to human trust in machines. According to the Muir's definition (1987, 1994), persistence means the constancy of the physical, biological and the moral social orders. Competence as defined by Barber (1983) was integrated with Rasmussen's (1983) definition of human behaviour into knowledge-, rule-, and skill-based behaviour. Responsibility means the expectation that human motives are reliable.

Dynamic of trust in machines. To examine the dynamics of trust in machines, Muir (1994) proposed a theoretical model of human-machine trust involving three distinct elements of trust, predictability, dependability, and faith, also based on the work in terms of interpersonal trust by Rempel et al. (1985). They generalized trust as expectation with consideration of four points to model three factors for the development of interpersonal trust. Trust evolves with (1) prior interaction and experience for mature relationship, and (2) dispositional credits, such as how other people are reliable and dependable considering expectations of rewards from the people, and (3) a willingness to take risk to build a relationship based on Deutsch's (1973) definition of trust which indicates "confidence that one will find what is desired from another, rather than what is feared", and lastly (4) confidence and security in the caring responses of the partner and the strength of the relationship.

Rempel et al. (1985) modelled three dimensions of interpersonal trust for prolonged relationships and assumed that predictability initiates human-human trust which governs the effective relationship, then dependability forms it, lastly faith dominates trust. Predictability of partner's behaviour refers the consistency of recurrent behaviour and the stability of the social environment. Predictability leads human to anticipate partner's future behaviour and action under uncertain situations based on previous consistency of response and personal understanding of the partner (Rotter, 1980). Dependability is originated from personal dispositional attributes of trustworthiness. As human relation develops, people know partners' personal traits and involve risk, uncertainty, and vulnerability. Here, dependability plays a vital role in trust formation. Trust is derived by not partners' specific actions but the person him/herself. Faith, last element of trust, seems fundamentally different from predictability and dependability and an aspect that is difficult to describe in a short word. Predictability and dependability are weighted by experiences and evidences through previous interaction with partners. Given that there is no consistently effective relationship, the relationship encounters new, unexpected, situations. Rempel et al. (1985) stated that faith based on predictability and dependability contributes to the determination of trust in situations when successful relationships are not guaranteed. They described that "faith reflects an emotional security on the part of individuals, which enables them to go beyond the available evidence and feel, with assurance, that their partner will be responsive and caring despite the vicissitudes of an uncertain future." (p. 97).

Table 1.1 An integrated model of trust in human-machine relationships, created by crossing Barber's (1983) model of the meaning of trust (rows) and Rempel et al.'s (1985) model of the dynamics of trust (columns). Statements in the cells exemplify the nature of a person's expectations of a referent (j) at different levels of experience in a relationship from Muir (1994).

	Basis of expectation at different levels of experience		
	Predictability	Dependability	Faith
Expectation	(of acts)	(of dispositions)	(in motives)
Persistence			
Natural physical	Events conform to natural laws	Nature is lawful	Natural laws are constant
Natural biological	Human life has survived	Human survival is lawful	Human life will survive
Moral social	Humans and computers act 'decently'	Humans and computers are 'good' and 'decent' by nature	Humans and computers will continue to be 'good' and 'decent' in the future
Technical competence	*j*'s behaviour is predictable	*j* has a dependable nature	*j* will continue to be dependable in the future
Fiduciary responsibility	*j*'s behaviour is consistently responsible	*j* has a responsible nature	*j* will continue to be responsible in the future

Muir (1994) extended this model to human trust in machines as follows. She then hypothesized that user trust in unfamiliar automation grows from predictability, to dependability, and lastly to faith following the suggestion of Rempel et al. (1985).

- *Predictability of acts* describes the perceived consistency of actions of a machine in a given situation.
- *Dependability of dispositions* refers to the extent to which operators can rely on not specific components of the machine but the capability of the entire machine.

- *Faith in motives* means an expectation that the machine performs in future situations beyond the behavioural evidence generated by the machine in terms of predictability and dependency.

Muir and Moray (1996) examined the relationship the dimensions of human-machine trust with a prediction that competence would be more applicable than responsibility with a simple linear regression as well as which dimensions and in what order they predict human-machine trust over time the best based on the three-factor theoretical model proposed in Muir (1994), further, aimed to understand how to optimize and predict operators' task allocation behaviour based on trust or automation property that impacts trust under supervisory control situations. Participants were required to control a simulated semi-automated pasteurization plant named 'PASTEURIZER' (Figure 1.3). After each session, participants completed questionnaires about trust and aforementioned factors. The result showed the expectation of competence best captured the meaning of trust as they expected. Contrary to their hypothesis, surprisingly, faith initially predicted overall trust towards the automated system, then dependability, and finally predictability.

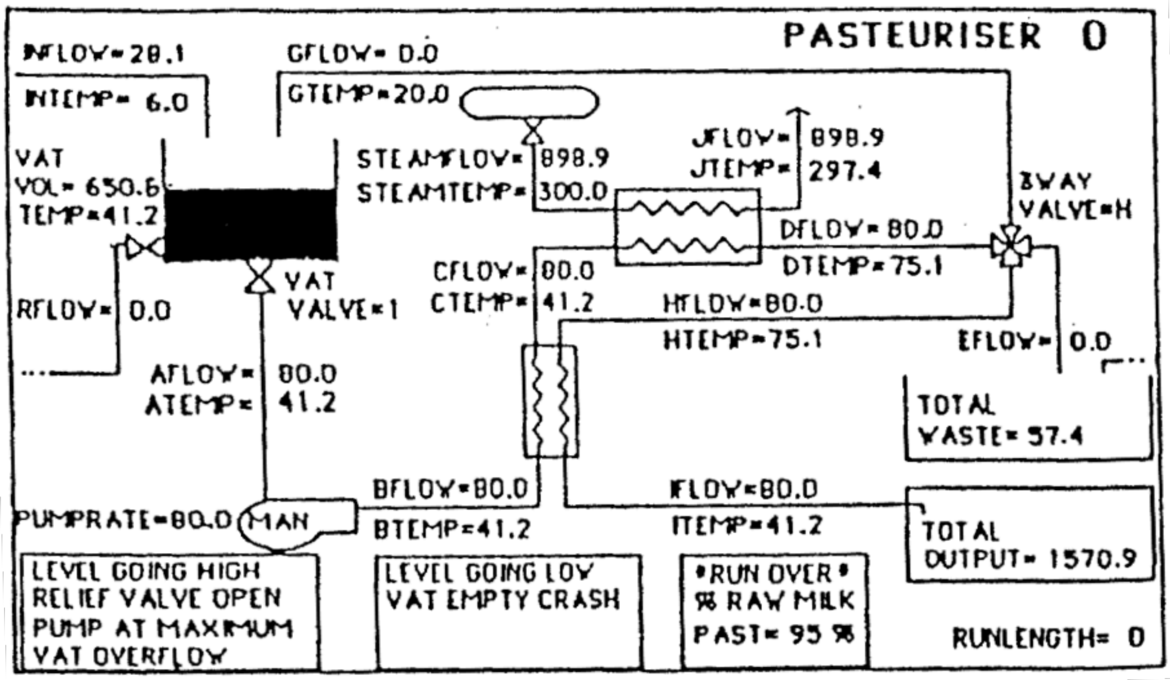


Figure 1.3: Pasteurization plant from Muir and Moray (1996).

This unexpected finding and three-factor model have been widely cited in the field of human-machine interaction and provides the foundation for a more modern theory of human-machine trust (e.g., Merritt & Ilgen, 2008; Bliss, 2009). This three-factors framework particularly affected the establishment of Lee and Moray's (1992) three bases of automation trust: purpose (faith), process (dependability), and performance (predictability) which published in advance of Muir (1994) and Muir and Moray (1996) in spite of late conduction. Lee and Moray (1992) also tested the relationship between dependability, predictability, and faith by using Muir and Moray's (1996) questionnaire of trust in automation for conceptualizing the three bases, and Lee and See (2004) reestablished three bases of trust based on an extensive literature survey. Further, empirical findings that how automation error in the course of the interaction affect levels of trust and reliance in novel systems also provided invaluable insights into human-machine trust as well as automation design (e.g. Lewandowsky et al., 2000; Körber et al., 2018a).

1.4 Research Questions and Thesis Overview

Muir (1994) opened her work with "Toffler's (1980) 'third wave' is upon us." Now, humans are facing the Fourth Industrial Revolution which fundamentally transforms human life-style with the development of technology across several domains, such as autonomous cars, the Internet of things, and artificial intelligence. How to interact with highly developed automated machines keeping appropriate levels of trust in automation is crucial in this context, thus knowing the mechanism of human-machine trust remains one of most important issues for human factors researchers in order to further design trustworthy and trustable automation (Ghazizadeh et al., 2012).

Sheridan, through his works, hypothesized the concept of operators' trust in supervisory control paradigms frequently over the years (e.g., Sheridan & Verplank, 1978; Sheridan et al., 1983; Sheridan & Hennessy, 1984). Interestingly, whilst several frameworks in terms of human-machine trust have been suggested with respect to adaptive function allocation scenarios as well as an understanding of which component contributes to shaping human-machine trust, there has been little investigation of trust development satisfying both theoretical issues and empirical results throughout human interaction with machines for last three decades since Muir and Moray's work (1996). Several studies doubted Muir's (1994) model and stated problems, however, there has been no direct and explicit challenges to their view. Inconsistent with Sheridan's expectation, still, Muir's (1994) work is considerably viewed as a seminal model to capture the development of trust for untrained individuals when automation includes several types of error and to identify factors determining human trust in unfamiliar machines.

This thesis research aims to have comprehensive understanding of automation design to consider which factor guides trust looking into how trust evolves when untrained individuals interact with automated machines in supervisory control situations. Trust calibrated by only question: "To what extent do you trust machine?" cannot account for trust development. The factor relating to trust closely should be considered to address the issue of human trust toward automation. In addition, empirical findings should be provided to have intuitive understanding of trust and automation design. The replication of previous study is a quite challenging work with reference difficulty in carrying out psychological experiment, ensuring reproducibility, and extracting very new and novel findings which can be distinguished from the original findings. In this sense, replicating the work of Muir and Moray (1996) is very challenging. However, it should be crystal cleared to discuss trust in automation for current general users because whether the old framework of trust development is still available for seizing human trust in machines in the modern society has been less explored with empirical findings. This study facilitates Muir and Moray's (1996) three-factor model to identify the key predictor which best embodies trust during human-machine interaction.

The chart as shown in Figure 1.4 illustrates the flow of a series of three experiments which explored for this thesis study. The first chapter (this chapter) describes fundamental theoretical and experimental studies which helped in the design of this thesis research. The second, third, and fourth chapters contain a series of three experiments to extend the investigation of human-machine trust in the course of human interaction with automated systems of process control systems and vehicle automation. Chapter 2 and 4 describes the replication of Muir and Moray's (1996) experiment, and Chapter 3 describes a driving simulator study. To address following issue, a total of three studies were conducted:

- Among Muir's (1994) three trust dimensions, which factor best captures initial trust in machines for untrained users in supervisory control situations
- How trust in unfamiliar machines for general users develops as experience of and exposure to automation
- Impacts of system failure as well as the type of system failure on users' trust in automated machines
- Relationship between Muir's three trust dimensions and other attributes influencing human-machine trust, such self-confidence and understanding of system mechanisms

STUDY I challenges the long-accepted view of Muir and Moray (1996) to figure out clear distinction of trust in automation between 1980s and now. It was hypothesized that trust develops from faith initially, then dependability, lastly predictability following the result of Muir and Moray (1996). That is, STUDY I aims to confirm whether the psychological structure of Muir and Moray (1996) can be identically replicated as well as show particular impacts on trust development. Full replication of Muir and Moray's (1996) experiment was carried out with the enlarged number of Japanese male students by reprogramming the PASTEURIZER. Following statistical analyses methods which facilitated in Muir and Moray (1996), difference of trust formation caused by changes of contexts discussed with quantitative comparisons between the original study and this study. Further, STUDY I discussed the difference caused by cultural contexts.

STUDY II reports on a study that sought to investigate key determinants of trust development for drivers in automated vehicles depending on different levels of knowledge in terms of vehicle automation as well as different types of system failure. STUDY II attempts to expand findings of STUDY I which dependability is a key contributor shaping human-machine trust to different automation domain. Subjective ratings of trust were collected to examine the impact of two factors: knowledge level (Detailed vs. Less) and type of system failure (by Limitation vs. Malfunction) in a driving simulator study in which drivers experienced a partially automated vehicle. It was hypothesized that trust may be governed by different trust dimension depending on different knowledge levels which presented to participants in advance of vehicle automation experience. The experience as well as type of system failure may lead decrease in levels of trust. Whilst the decreased trust by system limitation can be reestablished by subsequent experience of flawless automated driving, the decreased trust due to system malfunction cannot be recovered. Three-factor model was utilized to look into trust development, and changes of trust ratings over time was observed.

STUDY III presents partial replication of Muir and Moray (1996) with different cohort that only Japanese undergraduate students majored engineering field. Based on previous findings from STUDY I and II, trust in automation may be derived by dispositional qualities, such as academic major, gender, and occupation. In the STUDY III, gender was balanced to consider how dispositional attributes differently affect feelings of trust attributes and trust formation throughout human interaction with the PASTEURIZER. Statistical analyses – i.e., overall trust ratings of pump system and the development of trust in machines showed distinction from Muir and Moray (1996) confirming findings of STUDY I. Dependability is a main contributor initiating trust in machines and best predicts trust at most courses of interaction with machines. As faith dominates trust when participants understood the mechanism of all subsystems during training sessions, the result of STUDY III indicates that feelings of understanding system mechanism may call faith as the best predictor of trust. Further, STUDY III explores effects of the first impression on pump systems in the pasteurizer on trust and automation usage as investigates changes of self-confidence and time length that participants spent in the automatic pump mode. The results of STUDY III with participants who have relatively more affinity for technology confirmed that the increased use of automatic controller reflects difference between 1980s and the present.

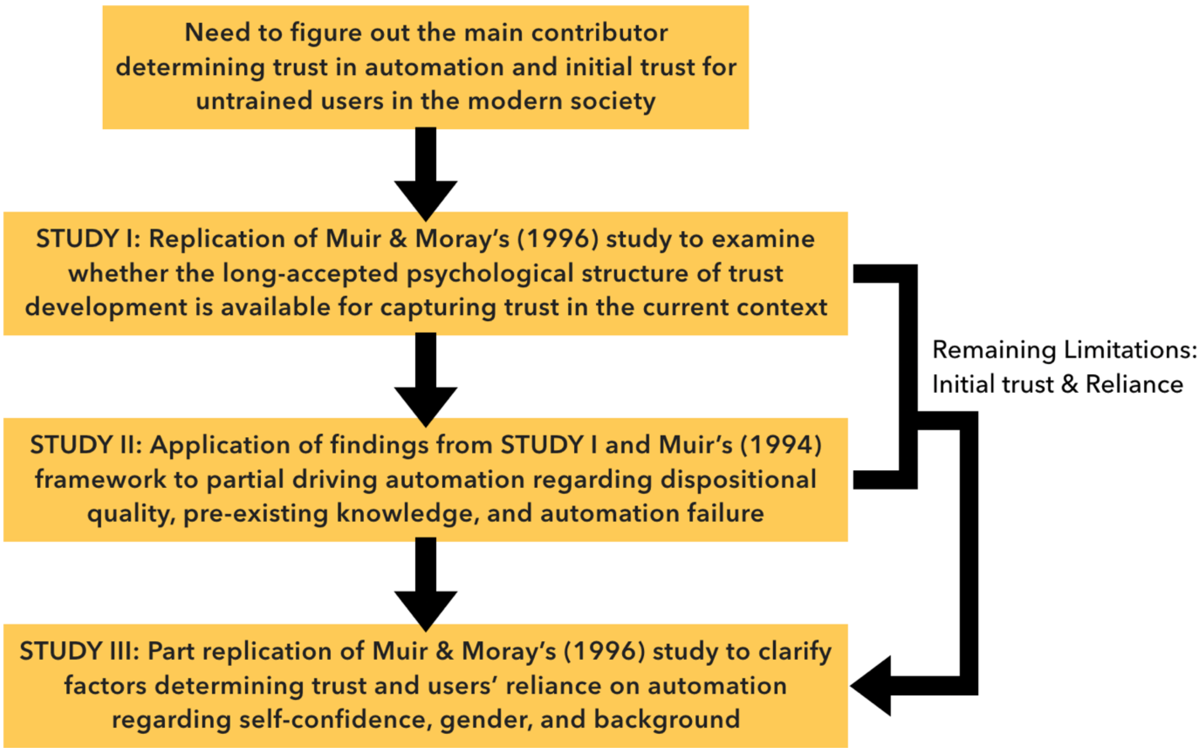


Figure 1.4: The research flow of a series of three experiments.

Chapter 2. REPLICATION OF MUIR AND MORAY (1996) WITH A NEWLY PROGRAMMED PASTEURIZER

2.1 Purpose of Replication

Chapter 1 describes why trust in automation is a critical issue from perspective of human factors researcher. As mentioned, Muir and Moray's (1996) work is particularly important with respect to theoretical as well as practical considerations. Muir (1994)'s framework is distinguished by current frameworks of human-machine trust. The three-bases or -layers of trust viewed it as integrations of several attributes of trust in automation (Lee & See, 2004; Hoff & Bashir, 2014), however, Muir (1994)'s model is grounded in the interaction among three-elements of trust during the course of human-machine interaction regardless of time point and degrees of trust. Operators' degree of trust in automation at a particular point in time may not depend solely on the state of the automation at that time but rather on the behaviour of automation in previous times, including whether automation has experienced partial or complete failures, whether the failures have occurred rarely or frequently, how long failures have lasted, and other time-based factors.

The wide and fast spread of automated machines for the last decades has greatly changed general users' perception of automation. With respect to changes of contexts between 1980s and the present, STUDY I attempted to replicate the findings of Muir and Moray (1996). This study focus on the overall trust and how human-automation trust develops. Three decades after the initial experiment, automation has become ubiquitous, with new technological advances constantly available both in the workplace and in the home. Though models of human-automation trust have evolved at least partly based on Muir and Moray (1996), it is unknown whether current general users of automation form faith initially to guide their trust development as suggested by Muir and Moray (1996). Process control experiment where participants performed the pasteurizer task described by Muir and Moray (1996; Figure 1.3) was reprogrammed. Thus, hypotheses of STUDY I mirrored Muir and Moray's (1996) findings. That is, it was hypothesized that trust in machines is initially best predicted by faith, then dependability, and finally predictability as interaction time increases. In addition, it was also hypothesized that automation failures decrease operators' subjective ratings of trust in automation, and recovery of trust is depending on type of failures. Thus, following research hypotheses were established:

- H1: Trust in machines is initially best predicted by faith, then dependability, and finally predictability as interaction time increases. That is, trust develops from faith, dependability, then predictability.
- H2: Continuous experience of machines for untrained users leads an increase in trust.
- H3: Automation failures decrease operators' subjective ratings of trust in automation.
- H4: The recovery of trust is depending on the type of failures. To be specific, variable system error which operators could not recognize where and which error occurs leads a large decrease in operators' trust in comparison with constant system error which they could recognize its symptom and expected future situations.
- H5: Competence can account for operators' meaning of trust in machines rather than responsibility.

2.2 Methodology

This research complied with the University of Tsukuba's ethics code and was approved by the ethical review board in the Faculty of Engineering, Information and Systems at the University of Tsukuba. Informed consent was obtained from each participant.

2.2.1 Participants

Twelve male undergraduate/graduate students (6 graduate students; M_{age} = 22 years, SD_{age} = 1.95 years) from the University of Tsukuba were recruited and participated in the study. This study conducted a power analysis with the pwr package in the R (R Core Team, 2018) to determine the number of participants in the current study. A sample size of N = 10 in the current experiment will result in power > .8 to detect an association between faith and overall trust early in the interaction of the original study. Each session lasted approximately 3 hours, and they were paid 2,460 JPY for each session.

2.2.2 Apparatus

A Fujitsu LIFEBOOK A577/RX laptop with a LED display (1366 x 768) was used to conduct simulation that programmed in MATLAB 2018a (Mathworks Inc., Natick, MA).

2.2.3 Pasteurization plant

"Pasteurizer", a simulated raw milk pasteurization plant task used by Muir and Moray (1996) was reprogrammed in MATLAB 2018a (Mathworks Inc., Natick, MA). Figure 2.1 presents a sample display of the pasteurizer plant simulation. From the upper left pipeline, raw milk entered a main vat. The raw milk travelled to an active heater from the main vat through pipelines, and a three-way valve determined whether the milk is appropriately pasteurized or not. If the temperature of milk was between 70 and 85 degrees Celsius, the milk entered the output vat. If the milk was hotter than 85 degrees Celsius, it was burned and went to the waste vat. If the temperature was lower than 70 degrees Celsius, the milk was not pasteurized and was transported to the main vat again. The operators' task was to maximise the amount of pasteurized milk while operating two subsystems: pump subsystem and heating subsystem.

The pump subsystem was used to adjust the flow of raw milk from the main vat. The system was operated in semi-automated, and participants were able to switch between manual and automatic mode. In the manual mode, participants were required to set a value for new pump target by typing commands, but in the automatic mode, the system automatically adjusted the value matching the milk flow rate into the system. Participants could check which mode was operating by looking at the "Pump System Mode" at top-right corner.

The heating subsystem was used to adjust the temperature of raw milk. Participants were required to adjust both the flow and temperature of steam in manually. The steam and raw milk stream's outlet temperatures are calculated based on the parallel flow heat exchanger. The values of overall heat transfer coefficient and heat transfer surface area are obtained by referring Muir and Moray (1996).

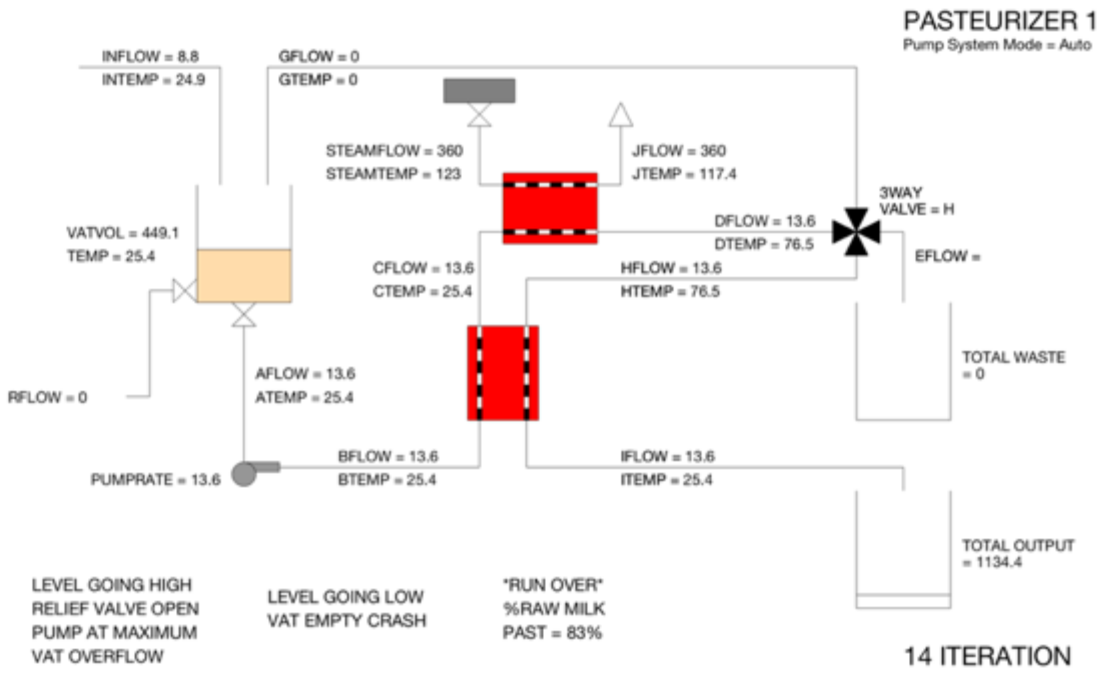


Figure 2.1: Screenshot of milk pasteurization plant simulation in this experiment. The orange box denotes where the milk enters the system. The black valve denotes the three-way valve outlined above.

2.2.4 Experimental Design

Both errors in the display and control properties were manipulated to examine the operators' trust in the pasteurization system, replicating Muir & Moray (1996). This study also manipulated control and display properties to be exact or honest (accurate), to have constant error, or to have variable error. The three control and three display manipulations were completely crossed, producing nine experimental conditions (Table 1). A within-subjects, repeated-measures design was used, and all participants performed nine plant sessions. Each session included four controlling simulation runs, with 80 iterations in each run.

The control manipulation. The pump's response to a new target pump was manipulated across three types: the exact response as requested (Exact), a 10% higher response than requested (Constant error), or a random response that is sampled from a Gaussian distribution with the mean of the requested pump target and a standard deviation of 2% of the pump target value (Variable error).

The display manipulation. The display conditions were manipulated across three types: an honest display of the pump rate (Honest), a rate 10% higher than manipulated pump rate (Constant error), or a rate with a variable Gaussian error (Variable error).

Table 2.1 Experimental condition.

	Display Manipulation		
Control Manipulation	Honest	Constant Error	Variable Error
Exact	C1	C2	C3
Constant Error	C4	C5	C6
Variable Error	C7	C8	C9

2.2.5 Dependent Variables

After every run, participants were required to complete the set of subjective rating scales for the pump system. This study translated the questionnaire of Muir and Moray (1996) from English to Japanese. The questionnaire with 100 mm-scale and the poles labelled “none at all” or “not at all” on the left to “extremely high” on the right (Table 2.2).

In Muir and Moray (1996), there was no significant effect in the three-way valve. Thus, the current study asked participants to take questionnaires with respect to the pump system. This study examined the relationship between trust and operators’ performance on the plant: scores, pump control actions, proportion of time in automatic pump mode. The score was calculated by dividing the total amount of successfully pasteurized milk by the total amount of flowed milk from the main vat. The score was displayed at the bottom-centre of the plant simulation. The pump control actions indicate how many times the operator adjusted the pump rate. The proportion of time in automatic pump mode means time length that the operators used the automatic controller during the trial.

Table 2.2 Constructs examined in subjective trust questionnaire (adapted from Muir & Moray, 1996)

Construct	Relevant questions
Competence	To what extent does the pump perform its function properly? To what extent does it produce the requested flow rates?
Predictability	To what extent can the pump's behavior be predicted from moment to moment?
Dependability	To what extent can you count on the pump to do its job?
Responsibility	To what extent does the pump perform the task it was designed to do in the system? To what extent does it maintain the system volume?
Reliability over time	To what extent does the pump respond similarly to similar circumstances at different points in time?
Faith	To what extent will the pump be able to cope with other system states in the future?
Trust in pump	To what extent do you trust the pump to respond accurately?
Trust in pump's display	To what extent do you trust the accuracy of the pump's display?
Overall trust	To what extent do you trust the pump?

2.2.6 Procedure

The procedure in the current study followed that of the original study by Muir and Moray (1996). The experiment included a training program and an experimental program. To become skilled at operating the pasteurizer, participants were required to carry out the training program at least eight sessions. There were five stages while the training: (1) to be introduced to the pasteurization plant with a detailed manual, (2) to practice the manual pump control mode, (3) to learn and practice how to switch between manual and automatic pump control mode, (4) to practice the automatic pump control mode, (5) to be acclimated with free choice of manual or automatic modes. In the stages of (1) and (3), the experimenters individually supported participants' first operation of plants. Further, when the participant's performance was kept over 80 % during last four runs, they could move on the experimental program. The mean number of training sessions required was 9.7 sessions with a range between 8 and 17 sessions.

After training program, participants proceeded to the experimental program. Their task was to maximise performance keeping the volume of main vat at around 500 litres. Participants were instructed that, in contrast to the training session, if the main vat volume reaches empty, the plant would crash and automatically shut down (Figure 2.1). The experimenter did not provide specific information about each control and display setting, but participants were told that the setting may vary from trial to trial due to equipment problems in some plants, making manual or automatic control less effective. Each participant completed the nine experimental conditions. After all trials, participants completed a brief interview asking about their impression on the pasteurizer, such as the first interaction with the pasteurizer, and whether they detect the differences. Finally, participants were debriefed, paid, and dismissed.

2.2.7 Statistical Analyses

Data were analysed by following statistical methods applied in Muir and Moray (1996). Subjective ratings of trust were collected via the questionnaires. There were three experimental factors: control properties of the pump (exact, constant error, variable error), display properties of the pump (honest, constant error, variable error), runs within sessions (run 1, 2, 3, 4). To measure operators' performance on pasteurization plants, this study used three performance measures: scores, the number of pump control actions, and the proportion of time in automatic pump mode. The three experimental factors were also analysed on these performance measures. Three-way (3 x 3 x 4) repeated-measures ANOVAs were conducted to analyse subjective ratings of trust (overall trust, trust in the pump, and trust in the pump's display) and the three performance measures across the three experimental factors. For all post-hoc tests, this study utilized the Bonferroni correction against alpha inflation due to multiple comparisons. This study assessed correlations between overall trust and three performance measures (performance score, manual control pump actions, proportion of time in automatic pump mode). Independent-samples t-test was used to compare differences between these three measures produced by the original study (Muir and Moray 1996) and the current study. To examine the meaning of trust in machines, simple linear regression and stepwise regression were conducted. To look into differences between operators who controlled the plant with only automatic pump modes and with only manual operation, independent-samples t-test and linear regression were performed. Further, stepwise regression was used to assess hypothesized models in terms of trust development. Variance inflation factors (VIF) were calculated to check multicollinearity in all regressions. If the VIF is greater than 10, it indicates a multicollinearity problem (see, Kabacoff, 2020). Thus, if the multicollinearity problem is detected, separate stepwise regression was performed. Bayesian analyses were employed to against conventional null-hypothesis significance tests. Bayes factors (*BF*) which represent the degree to which the observed data favour one statistical model over another model quantify whether the data are more compatible with a null model or an alternative (Rouder et al., 2009; Schönbrodt et al., 2015). For the current study, *BF*s obtain evidence for a null hypothesis as it can distinguish between uninformative results and results supporting the null hypothesis. Three participants failed to keep the main vat volume, resulting in system crashes. Thus, the data of failed participants in the experimental program were eliminated. All statistical analyses were conducted in R (R Core Team, 2018).

2.3 Results

2.3.1 Overall Trust in Pump System

Operators' overall trust was significantly affected by the display manipulation, $F(2, 16) = 3.87$, $p = .04$, $\eta_G^2 = .03$, $BF_{10} = 6.23 \times 103$. Post-hoc t-tests revealed that the operators reported higher levels of trust towards automatic pumps with the honest display more than the variable display error, paired-samples t-test $t(8) = 2.40$, $p = .04$, $BF_{10} = 1.15 \times 104$, $M = 82.99$ vs. 74.93. The remaining effects were not significant, all ps $> .13$. Table 2.3 describes the average of subjective ratings of trust and performance variables with standard deviation, and Figure 2.2 presents the mean ratings of overall trust in the pump system toward control and display properties. This indicates that variable display error in the pump system leads decreased trust compared to when using the honestly displayed pump system.

Table 2.3 Mean and standard deviation of subjective ratings of trust and performance.

	Subjective Ratings of Trust			Performance		
Experimental Condition	Overall Trust	Trust in Control	Trust in Display	Score	Manual Control Action	Proportion of Time in Automatic Pump Mode (%)
C1	83.81	84.08	82.39	96.02	24.83	64.40
	(17.30)	(15.25)	(17.85)	(3.54)	(41.73)	(46.37)
C2	82.53	79.72	81.56	95.94	22.39	64.61
	(17.08)	(24.14)	(16.82)	(3.94)	(31.96)	(46.56)
C3	74.66	76.41	77.75	93.29	19.00	63.37
	(20.18)	(24.85)	(15.94)	(5.85)	(32.95)	(46.08)
C4	80.24	82.40	80.65	88.29	38.44	67.76
	(12.63)	(13.12)	(15.51)	(14.33)	(73.25)	(42.77)
C5	75.86	76.52	78.10	92.63	22.75	59.57
	(20.20)	(18.54)	(22.22)	(6.87)	(37.42)	(48.38)
C6	72.78	75.10	76.70	91.43	30.64	62.89
	(18.66)	(13.56)	(19.40)	(8.57)	(49.12)	(44.87)
C7	84.91	86.03	82.24	96.28	14.22	63.99
	(16.97)	(13.70)	(19.04)	(3.81)	(30.64)	(46.47)
C8	78.56	78.89	79.03	94.74	29.69	64.20
	(18.90)	(14.43)	(17.30)	(5.28)	(47.91)	(46.52)
C9	77.38	79.91	79.09	93.64	38.89	64.03
	(19.20)	(13.33)	(19.54)	(7.79)	(63.11)	(46.28)
Overall	80.16	80.11	78.56	94.42	19.15	71.40
	(18.17)	(18.84)	(18.98)	(4.95)	(36.32)	(42.43)

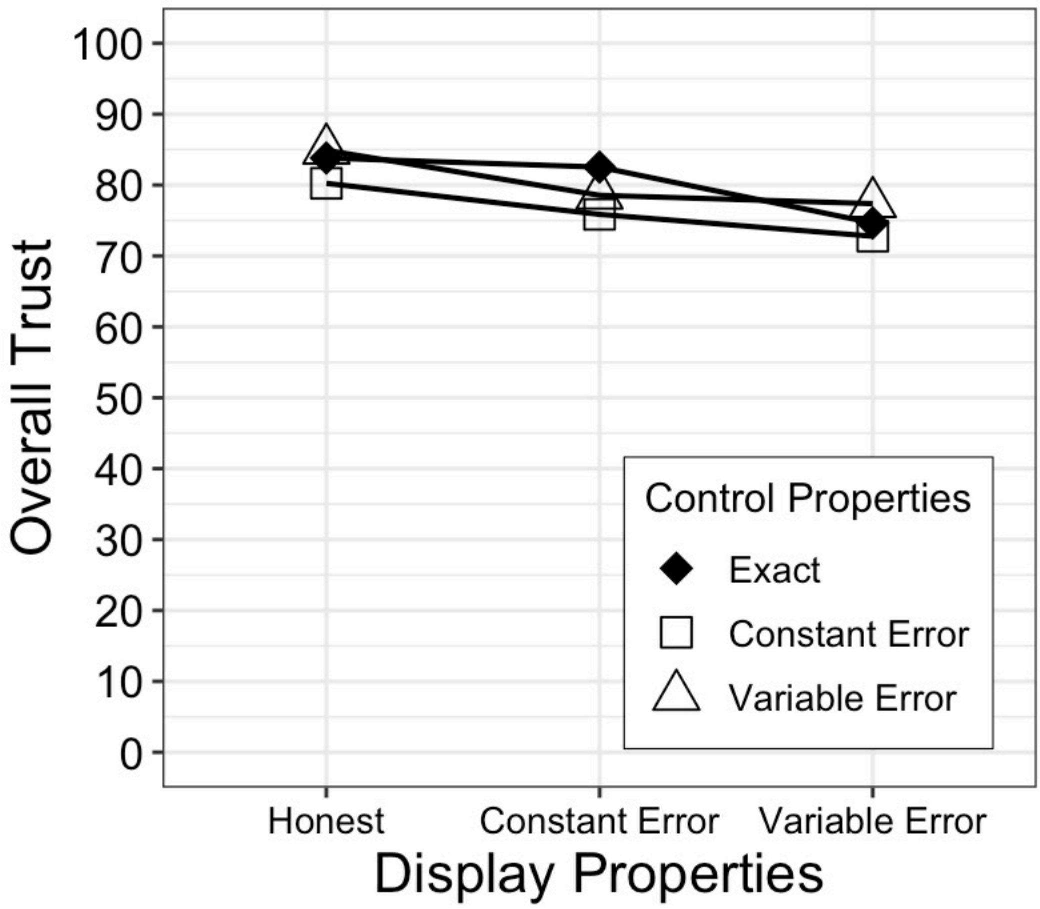


Figure 2.2 Mean ratings of overall trust in the pumps, pooled over operators and runs, as a function of the control and display properties of the pumps.

Figure 2.3 illustrates mean ratings of overall trust in the pump system, assembling operators and runs, as a function of the control and display properties of the pumps. Figure 3 shows mean ratings of overall trust in the pumps, assembling run, as a function of the control property of the pumps and operators' experience respectively.

A significant two-way interaction between control property and run was observed, $F(6, 48) = 2.42, p = .04, \eta_G^2 = .004, BF_{10} = .01$. One-way ANOVAs for each of control property revealed that main effects of the run factor on levels of trust towards the exactly controlled pump, $F(3, 24) = 4.3, p = .015, \eta_G^2 = .01, BF_{10} = .25$, and the pump with the variable error , $F(3, 24) = 3.61, p = .027, \eta_G^2 = .01, BF_{10} = .09$. For the exact control property, post-hoc t-tests observed that operators' levels of trust at the first trial was lower than those at the second, paired-samples t-test $t(8) = -2.4, p = .045, BF_{10} = 1.97, M = 77.54$ vs. 81.76, and at the third, paired-samples t-test $t(8) = -2.5, p = .037, BF_{10} = 2.27, M = 77.54$ vs. 80.33, and lastly at the fourth trial, paired-samples t-test $t(8) =$ -

4.4, $p = .002$, $BF_{10} = 21.83$, $M = 77.54$ vs. 81.71. For the variable control property, post-hoc t-test revealed that operators' levels of trust at the first trial was lower than the fourth, paired-samples t-test $t(8) = -2.4$, $p = .046$, $BF_{10} = 1.94$, $M = 76.92$ vs. 80.39. The ratings of operator's trust at the 2nd trial was lower than those at the third, paired-samples t-test $t(8) = -2.3$, $p = .048$, $BF_{10} = 1.87$, $M = 80.38$ vs. 82.42. These results indicate that operators' levels of trust in the pump system increase with the experience of machines which exactly controlled or contained random error. The remaining effects were not significant, all ps > .07.

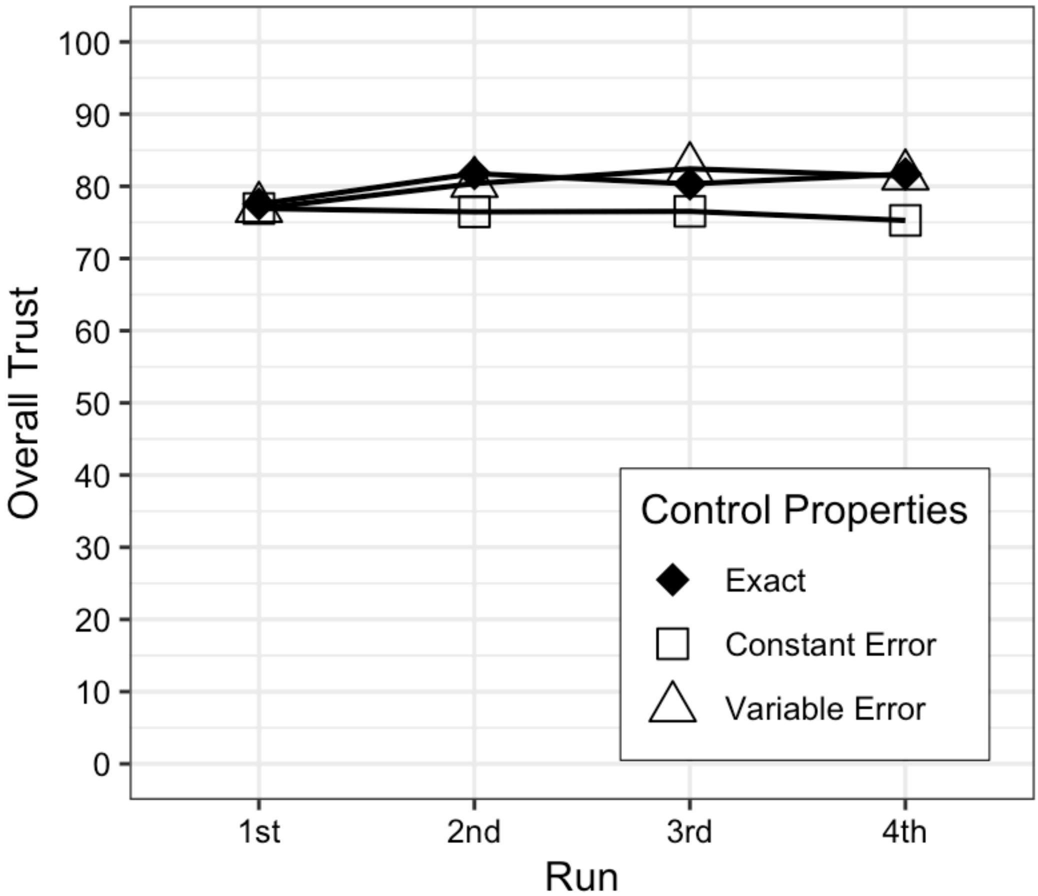


Figure 2.3 Mean ratings of overall trust in the pumps, assembling the run factor, as a function of the control property of the pumps and operators' experience.

2.3.2 Trust in Pump Control and Display Properties

Figure 2.4-(a) and –(b) illustrates mean ratings of trust in the control and display properties of the pumps respectively. Trust which independently rated for pump's control and display properties did not show significant difference, all *p*s > .07.

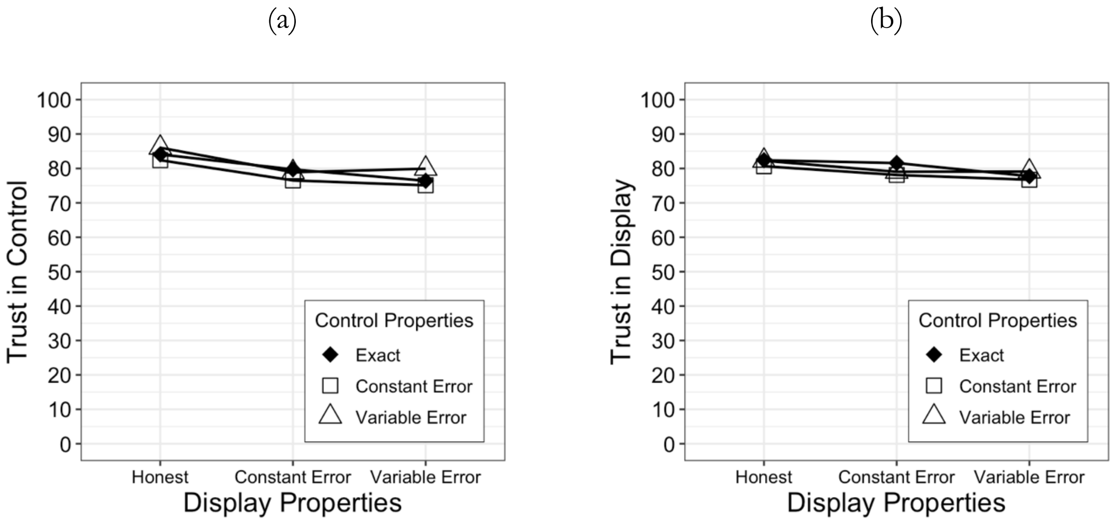


Figure 2.4 Mean ratings of trust in the control (a) and display properties (b) of the pumps.

2.3.3 Performance

Figure 2.5 presents the mean performance score, manual pump control actions, and proportion of time in automatic pump mode of Muir and Moray (1996) and the current study separately.

Performance score. The main effect of control properties on performance scores was statistically significant, $F(2, 16) = 5.34, p = .02, \eta_G^2 = .09, BF_{10} = 4.61 \times 10$. Post-hoc t-tests revealed that performance scores in constant control error were lower than those of variable control error pump, paired-samples t-test $t(8) = -2.5, p = .037, BF_{10} = 22.65$ $M = 90.78$ vs. 95.08. The remaining effects were not statistically significant, all *p*s > .06.

Independent-samples t-test revealed that scores of the current study were significantly lower than those of Muir and Moray (1996), $t(16) = 3.88, p < .001, BF_{10} = 23.89$, Cohen's $d = 1.83$. These imply that participants in both the original study (Muir and Moray, 1996) and the current study were able to figure out how to achieve high scores.

Manual pump control actions. No effect was statistically significant, all *p*s > .19. Unlike Muir and Moray (1996), mean manual pump control actions in this experiment were significantly lower, independent-samples $t(16) = 1.17$, $p = .26$, $BF_{10} = .66$, Cohen's $d = .55$.

Proportion of time in automatic pump mode. The three-way interaction was statistically significant, $F(12, 96) = 1.995$, $p = .03$, $\eta_G^2 = .07$, $0 < BF_{10} < .1$. The proportion during the third run under the C5 condition (constant control error-constant display error) was the lowest, 42.94%. The second lowest was 61.32% under the C6 condition (constant control error-variable display error) during the first run. It indicates that operators spent less time in the automatic pump mode in the C5-third run compared to the all runs in the all conditions, all *p*s < .01.

The operators in the present study spent more time with the automatic pump than Muir and Moray (1996), independent-samples t-test $t(16) = 50.20$, $p < .001$, $BF_{10} = 2.51$ x 1015, Cohen's $d = 23.67$.

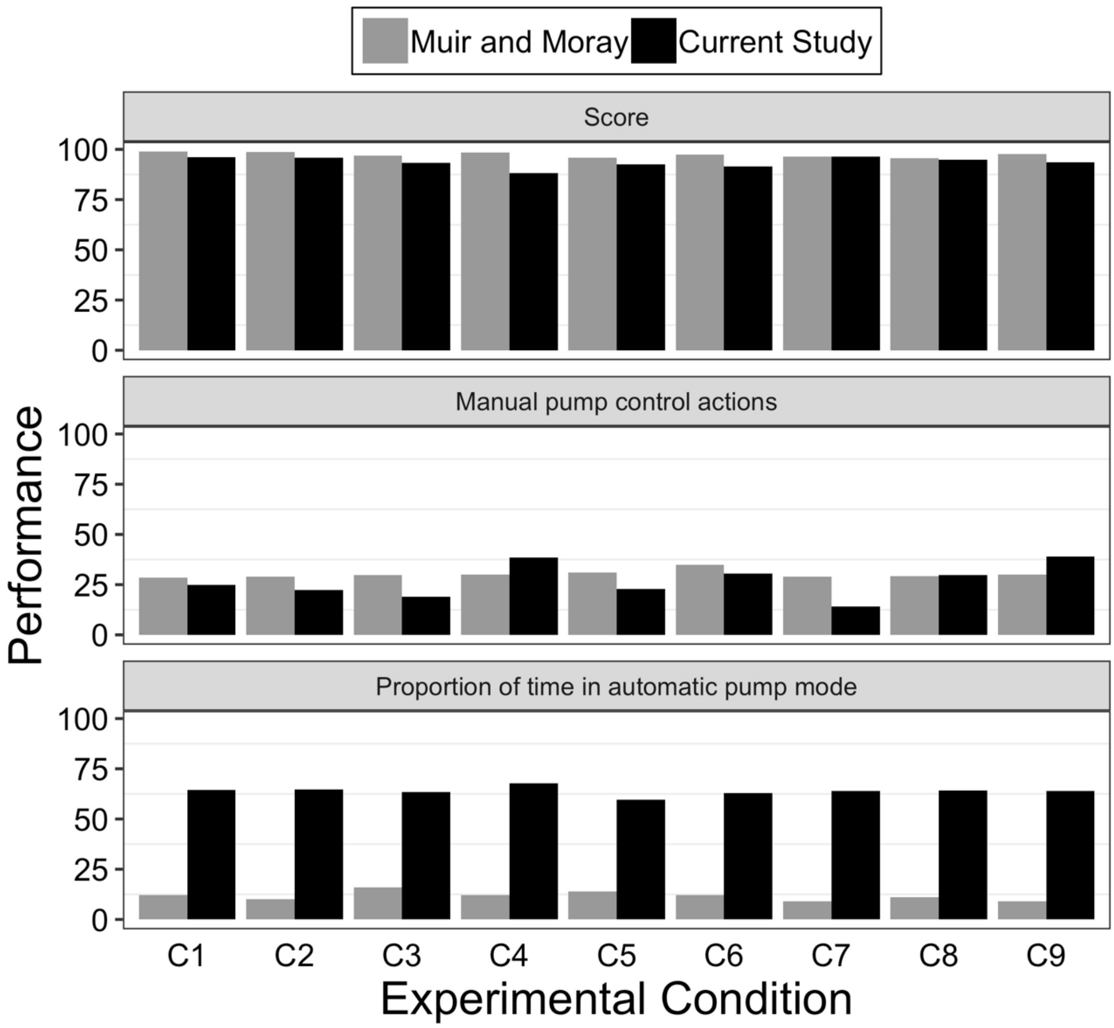


Figure 2.5 Mean values of performance score, the number of manual control operations, and the proportion of time in automatic pump mode.

2.3.4 The Meaning of Trust in Machines

Muir and Moray (1996) applied both simple linear regression and stepwise regression analyses to examine whether competence and/or responsibility predict overall trust in machines. Muir and Moray (1996) found that competence is a better predictor of trust accounting for the meaning of trust than responsibility. The current data showed that responsibility was the most predictive in all the experimental conditions except for the C6 condition (constant error-variable error) as shown in Table 2.4.

Table 2.4 Summary of analyses of the meaning of trust. *C* = competence, *Rs* = responsibility, T = overall trust.

Experimental Condition	R^2 *C*	BF_{10} *C*	R^2 *Rs*	BF_{10} *Rs*	R^2 *C*, *Rs*	Stepwise Best Predictor	*F* (Best Model)	*p*
Total	0.62	1.32 x 1065	0.76	8.62 x 1098	0.58	*Rs*	1042.16	< .001
C1	0.74	3.52 x 108	0.77	1.6 x 109	0.58	*Rs*	110.83	< .001
C2	0.79	2.01 x 109	0.94	5.23 x 1019	0.90	*Rs*	601.80	< .001
C3	0.59	2.58 x 105	0.88	1.79 x 1014	0.77	*Rs*	261.40	< .001
C4	0.48	3525.26	0.69	1.76 x 107	0.46	*Rs*	75.55	< .001
C5	0.76	1.42 x 109	0.83	1.92 x 1011	0.67	*Rs*	160.40	< .001
C6	0.58	9.45 x 103	0.56	8.78 x 103	0.31	*C*	44.98	< .001
C7	0.85	1.49 x 1012	0.85	2.55 x 1012	0.72	*Rs*	193.77	< .001
C8	0.64	1.4 x 105	0.77	2.1 x 108	0.58	*Rs*	113.10	< .001
C9	0.45	2179.93	0.72	1.61 x 107	0.52	*Rs*	91.66	< .001

There were three participants who controlled the plant in only the manual operation and other three participants who operated the plant using only the automatic pump mode. Mean ratings of overall trust and responsibility for the automatic pump-preferred operators were significantly higher than those of the manual control-preferred operators, U = 6 and 7, respectively, *p*s > .4. Results of a stepwise regression analysis showed that both competence and responsibility predicted overall trust in machines. Competence was more predictive than responsibility for the manual operation group whilst competence was less predictive compared to responsibility for the automatic pump mode group.

A regression analysis was performed to further investigate a relationship between operators' overall trust and responsibility with respect to the types of operators: manual- and automatic pump-preferred (Figure 2.6-(a) and (b) respectively). The ratings of responsibility significantly predicted overall trust towards the automatic pump system for operators who preferred the manual pump control, (R^2 = .76, $F(2, 106) = 342.9$, $p < .001$) and for operators who preferred the automatic pump control ($R2$ = .60, $F(1, 106) = 162$, $p < .001$).

Bivariate correlation analyses showed strong correlations between competence and responsibility (C1, $r = .88$; C2, $r = .92$; C3, $r = .64$; C4, $r = .70$; C5, $r = .97$; C6, $r = .78$; C7, $r = .95$; C8, $r = .91$; C9, $r = .78$; Pooled, $r = .79$). Additionally, correlation analyses were conducted regarding two operator types and showed strong correlations between two factors, r = .76 and .89 for the automatic pump- and manual control-preferred operators respectively. These results confirm that two factors are highly intercorrelated, but variance inflation factor analyses did not reveal the problems of multicollinearity except for C5 (VIF = 14.38) and C7 (VIF = 10.81) conditions.

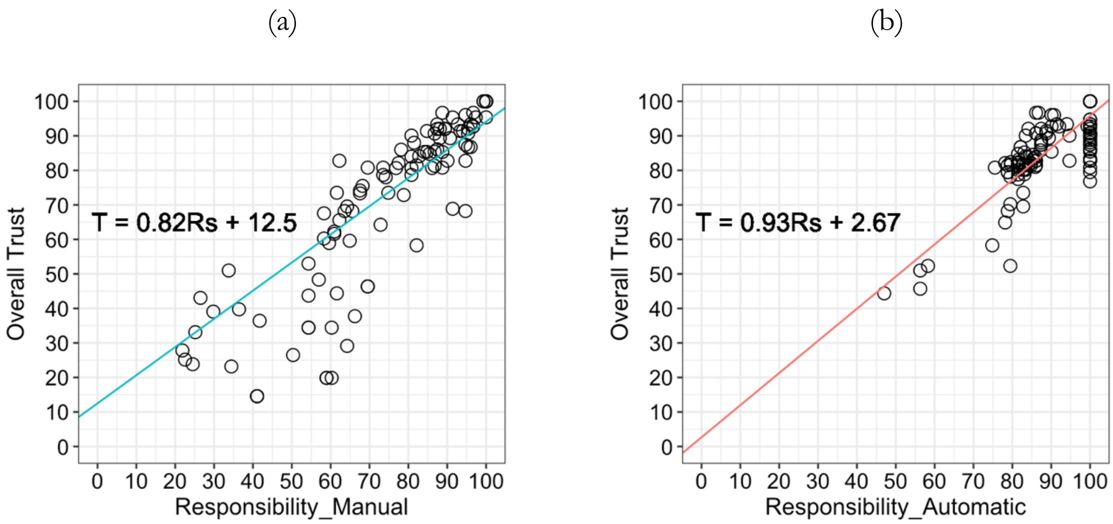


Figure 2.6 Scatter plots of ratings of responsibility with overall trust in the pump for operators who preferred the manual pump control (a) and the automatic pump mode (b) with the fitted regression line and the regression equation for each.

2.3.5 The Development of Trust in Machines

Following Muir and Moray (1996), the data were extracted with respect to three points during operators' experience: the first training session, the last training session, and the last run of the experimental session in C1. Further, this study selected the first questionnaire after the first experience of automatic pump control system in the current study. Contrary to the findings of Muir and Moray (1996), the results showed that dependability was the best predictor of overall trust in the automated pump system throughout all time points (Table 2.5). Bayesian regression analyses also showed that dependability was the best predictor of trust in the pump system after the first interaction with the automatic pump controller (predictability, BF_{10} = 1.77; dependability, BF_{10} = 182.05; faith, BF_{10} = 1.88), the first training (predictability, BF_{10} = 18.5; dependability, BF_{10} = 32.67; faith, BF_{10} = 2.92), in the last training (predictability, BF_{10} = 1.23; dependability, BF_{10} = 13.65; faith, BF_{10} = .63), and in the C1 (exact control-honest display) experimental condition (predictability, BF_{10} = 7.98; dependability, BF_{10} = 168.78; faith, BF_{10} = 2.54).

Table 2.5 Summary of analyses of the development of trust.

Session	R^2 Predictability	R^2 Dependability	R^2 Faith	Stepwise Best Predictors	*F* (Best Model)	p
Automatic Pump	0.42	0.90	0.45	*D*	65.43	< .001
First Training	0.83	0.92	0.69	*D*	85.66	< .001
Last Training	0.42	0.86	0.26	*D*	46.24	< .001
Plant (C1)	0.69	0.90	0.52	*D*	63.55	< .001

Figure 2.7 illustrates bivariate correlation between three dimensions of trust in automation at each point. VIFs at all points showed that there are multicollinearity problems in predictability and dependability at two points: the first training and the experimental session in the C1 plant. Separate stepwise regressions regarding predictability and faith, and dependability and faith showed both predictability and faith captured trust in automation after experiencing the first training, $F(2, 6) = 21.41$, $p = .002$, $R^2 = .84$, and that dependability only best predicted trust, $F(1, 7) = 32.01$, $p < .001$, $R^2 = .8$. Similar results were observed for the point after the C1 plant. Both predictability and faith predicted trust in automation, $F(2, 6) = 13.7$, $p = .006$, $R^2 = .76$, and dependability only

accounted for trust, $F(1, 7) = 63.55$, $p < .001$, $R^2 = .89$. All results indicated that dependability predicted overall trust in the pump system for operators.

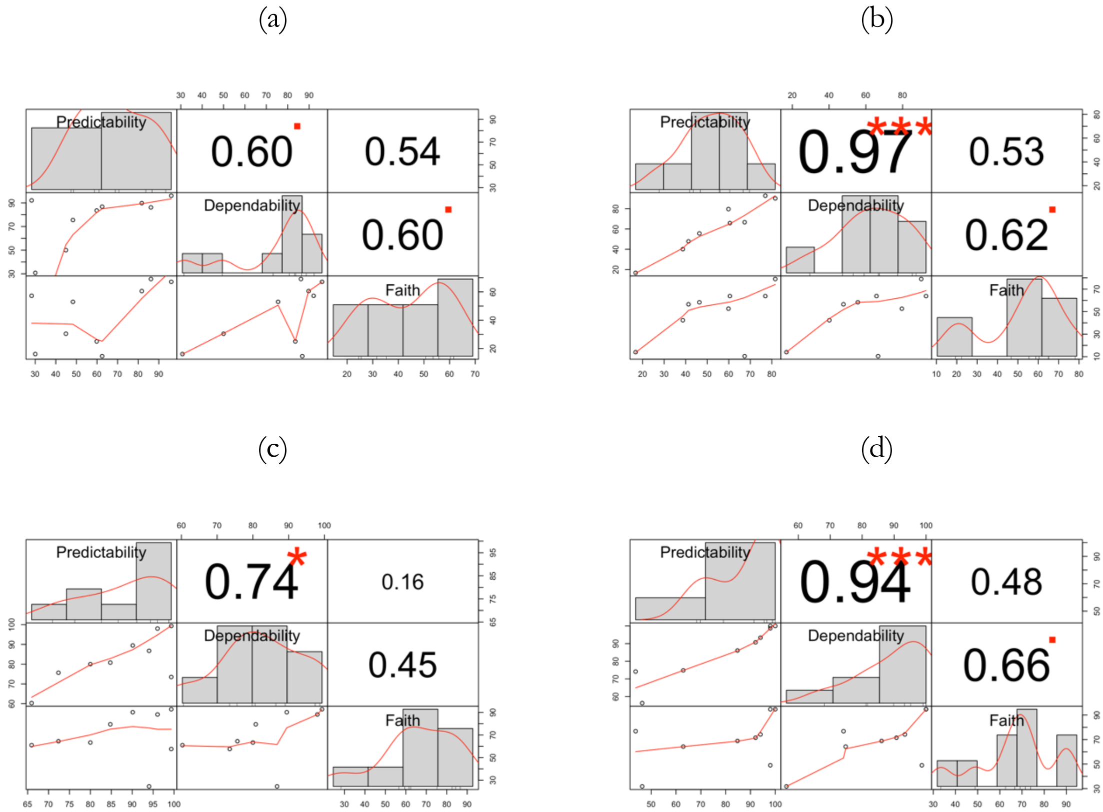


Figure 2.7 Bivariate correlation between all dimension of trust at all time points: (a) first interaction with automatic pump mode, (b) first training, (c) last training, and (d) experimental session of C1 plant.

2.3.6 Correlation between performance and trust

The correlation between overall trust in the pumps and performance scores was statistically significant, $r(322) = .18$, $p = .002$. The negative correlation between overall trust and the number of manual pump operation was significant, $r(322) = -0.48$, $p < .001$. Lastly, operators' trust was significantly and positively correlated with use of the automatic controller, $r(322) = .31$, $p < .001$. Figure 8 illustrates scatter plots for each correlation.

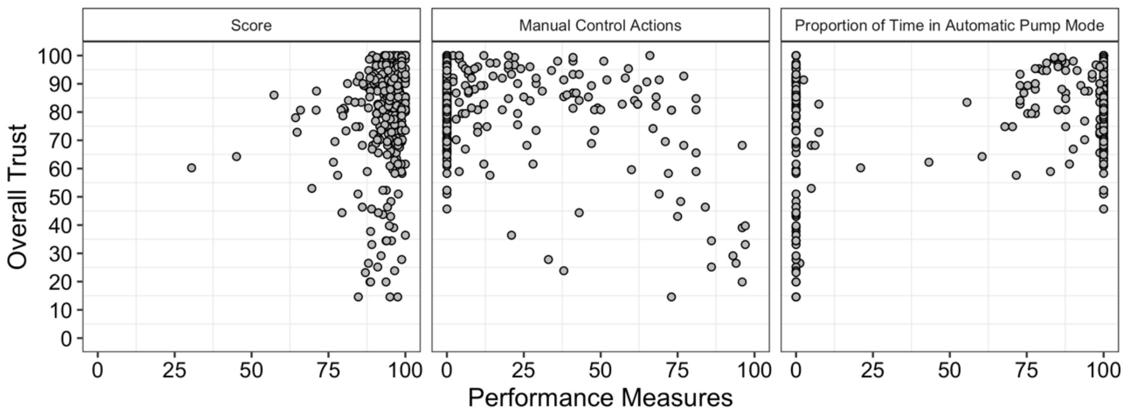


Figure 2.8 Scatter plots of correlations between ratings of overall trust in the pump and three performance measures.

2.4 Discussions

In this study, the current study aimed to replicate Muir and Moray (1996) by asking twelve participants to operate a simulated pasteurizer plant task between manual and automatic pump control modes.

2.4.1 Subjective Ratings of Trust on Pump System

Muir and Moray (1996) showed that the three experimental factors, control and display properties and run factor, affected operators' trust. They found the main effects of both control and display properties on overall trust towards the automatic pumps and an interaction effect between the run factor and the control property. Thus, this replication study also expected to find same results which automation failures have different impacts on the locus and type of failure (H3, H4) as well as the increase in trust over time (H2). In our study, the display properties significantly impacted operator's trust, with the variable error condition leading to decreased overall trust scores. However, the analysis of trust in the pump control and trust in display exhibited no interaction between control and display manipulations. Thus, the finding of overall trust analysis may partly support that information display plays an important role influencing operators' trust (e.g. Muir & Moray, 1996; Lee & See, 2004).

Muir and Moray (1996) reported trust is likely to grow with continuous human-machine interaction when either non-error existed or the constant error occurred except for the occurrence of the variable error. The current study found a similar tendency when the operator experienced the exactly controlled pump and the pump with the variable control error inconsistent with the finding of the original study. This unexpected result is indicative of the relationship between the performance score and trust because the mean scores in the experimental conditions including the variable control errors were significantly higher than those of conditions with the constant control error. As operators' performance is one of determinants affecting operators' trust (Rempel et al., 1985) and post-experiment interviews revealed participants were highly sensitive to maximising the performance score, it was expected that the score impacts operators' trust formation. The scores in the constant control error conditions were significantly lower than those of the variable control error and exact control conditions. Furthermore, the current results support that performance score is positively associated with overall trust, indicating that those who achieved higher levels of the pasteurizer task also displayed higher levels of trust. Muir and Moray (1996) also concluded that the score perhaps influences the determination of operators' trust with a small correlation between trust and score. Further corroborating this view, the current results also indicate that trust is negatively associated with the total number of manual pump operation and positively with the use of the automatic pump. Other empirical studies also describe that when automation fault occurred, both operators' subjective level of trust performance scores decreased (Lee & Moray, 1992) and the relationship between trust and performance may vary depending on types of automation or automation fault (e.g. Lewandowsky et al., 2000). Future research should explore this finding with respect to the relationship among trust, system property, and system performance using the pasteurizer task to replicate and other more advanced automated systems to generalize the current findings.

2.4.2 The Meaning of Trust in Machines

Responsibility and competence have been reported as important factors influencing human-machine trust (e.g., Sheridan, 1988; Merritt & Ilgen, 2008). In this study, competence is related with production of the requested pump flow rates, and responsibility is related with maintenance of the volume of the main vat. Muir and Moray (1996) reported that competence was the dominant factor accounting for the meaning of human-machine trust (H5). However, our study exhibited results inconsistent with Muir and Moray (1996). Responsibility was the dominant factor accounting for the meaning of trust in the current study.

For each participant, it was important to understand the mechanism of the pasteurizer to maximise the performance score. In other words, being competent in operating the pump system should be closely associated with maximising the score. The proportion of time in automatic pump mode showed that about a half of the operators used automatic mode for adjusting the pump rate rather than the manual operation. Most of the operators (8 of 12) also selected automatic mode from the beginning of the training program to the end of the experimental program. The participants were highly sensitive to producing wastes and the performance score, and it is possible to increased use of the automatic mode to effectively maximising the output. Therefore, it may be that how the pump is integrated in the plant is perceived more important for them than the direct comparison between their actual control behaviour and the operation of system. This finding implies that operators' use of the automatic mode may be an important factor determining the meaning of trust in automated systems. It is possible that the best predictor of meaning of trust perhaps differs by the proportion of using the automated aids. This is further supported by differential relationships between trust, competence and responsibility for those who used the manual pump mode and those who used the automatic pump mode. For operators who used the manual operation mode rather than the automatic pump mode, competence was the best predictor of overall trust. On the contrary, for operators who used the automatic pump mode more rather than the manual operation mode, responsibility as the best predictor of overall trust. This contrasting finding might have an implication for the impacts of users' choice between manual operation and automatic systems on the role of two attributes differentially affecting automation trust: competence and responsibility. However, this result should be carefully considered with the multicollinearity problems due to high correlations between the two contributors. Indeed, Muir and Moray (1996) concluded that both competence and responsibility significantly contribute to capturing the meaning of trust in machines, and the two factors are highly intercorrelated (Lee & Moray, 1992, Merritt & Illgen, 2008).

2.4.3 The Development of Trust in Machines

Many of studies on human-machine trust employ the framework Muir and Moray (1996; e.g., Lee & See, 2004; Jian et al., 2000). This replication study hypothesized that trust develops from faith initially, then dependability, finally predictability following the result of Muir and Moray (1996; H1). Our result showed that dependability was a consistently dominant dimension among three dimensions of human-machine trust in all conditions across different time points during their interaction with the pasteurizer (see, Table 2.5). Separate stepwise regression was conducted considering the multicollinearity problem, and the results of regression and Bayesian analysis demonstrated that dependability is highly likely to account for trust in automation at all time points. This result might suggest a new cohort perspective on the human-machine trust between 1989 and 2018. Automation has become pervasive and has been frequently introduced for the general public in 2018 compared to the 1980s (e.g., Autor, 2015), and this result may reflect a factor that initiates the development of automation trust that is different from users in the 1980s. One clear difference between 1989 and 2018 is the frequency of use of the automatic mode. Although there was no significant difference of how many times participants toggled between manual and automation modes between 1989 and 2018 cohorts (*M* = 30.13 vs. 26.8), more participants in 2018 spent more time with the automatic pump in the experimental program than those of 1996 (*M* = 63.9% vs. 11.67%). This substantially higher percentage of the automatic pump use by the current participants perhaps mirrors operators' reliance on the ability of the automatic pump system. Additionally, the higher percentage of the automatic pump use and the high performance score provide evidence that adequate human-machine interaction was emerged.

Another interpretation of the present result is that the task difficulty of using manual pump control mode in the first practice session might undermine operator self-confidence, leading dependence upon the automatic control mode. Operator self-confidence in managing the system is an important factor influencing use of automation (Moray et al., 2000). Operators with low self-confidence greatly rely on the automatic pump mode, and most operators depended on the automatic control rather than operating the pump rate manually (Lee & Moray, 1994; Moray et al., 2000). de Vries et al. (2003) stated that when operators' trust is high, and self-confidence is low, operators are more likely to employ automated systems. Furthermore, when operators utilize manual mode, their self-confidence grows over time (Chavaillaz et al., 2016). In this study, a higher proportion of time in the automatic mode was observed than that of Muir and Moray (1996). The participant interviews conducted after the experiment support the idea that operators chose the automatic mode under their expectation that the automatic system ensures better scores than the manual operation because they found the task challenging. This may reflect that low self-confidence deprives operators of the confidence to use the system, resulting in a higher use of the automatic control mode. Low self-confidence may relate to the current finding that overall trust is best predicted dependability, which means the extent to which operators can rely on the machine. Additionally, this also may be seen as a possible explanation why responsibility could be a critical predictor of overall trust.

Our results question predictability being a determinant of human-machine trust at the beginning of users' interaction with a novel automated system. Based on the framework of human-

human trust proposed by Rempel et al. (1985), Muir and Moray (1996) hypothesised that predictability might be a predictor of initial trust in machines. However, Muir and Moray (1996) exhibited faith as the best predictor, and our present study found that dependability can account for initial trust. One possible interpretation is a discrepancy between laboratory study and real-world study. For instance, the study of Balfe et al. (2018) which examined professional rail operator's trust in the real-world rail automation system found predictability is one of important factors pertaining to trust. Since operators in this laboratory setting used a simulation, they were aware that safety is ensured, and no penalty or bonus will be imposed depending on their performance. However, even though most operators in the real-world frequently use automated aids, the supervisory control system still requires human operators to execute complex information-processing performance. Thus, the operator in the real-world should be skilled at controlling all information processes, and they cannot ignore any minor errors contrary to participants in the current study (Hoffman et al., 2013). The present result might suggest that the dominant dimension of human-machine trust development depends on perceived risk (e.g. Sato et al., 2019).

2.4.4 Limitations

Findings of the current study should be interpreted with the following limitations. First, cultural difference may have an impact on the development of trust. When considering the cultural context associated with trust and acceptance of automation (see Zuboff, 1988; Baba et al., 1996), our finding is limited to generalizing human-machine trust in Japanese male university students. This may also account for some differences between the current study and Muir and Moray (1996). The cross-cultural research of Yamagishi and Yamagishi (1994) described that trust displayed by Japanese depends on the networks of mutually committed relationships and is relatively lower than those of American. Their low level of trust is based on tendency that keeps social stability and reduces uncertainty about transactions (Doney et al., 1998). This perhaps takes account into the presence of a gap of trust Japanese and American. Indeed, this cross-cultural research on interpersonal trust is a gap in the literature, and more research should consider individual characteristics such as cultural background and organizational context as factors that potentially impact human-machine trust (see, Bliss et al., 2019).

Secondly, the present experiment was conducted with limited information of the previous study (i.e. default values in the pasteurization plant). This might account for some of the discrepancy between the prior and present studies. Muir and Moray (1996) described that operators could found the presence of errors as well as discrepancies by control or display manipulations, but according to the post-experiment participant interview in the current study, some participants were not able to identify the errors. Whilst nine of twelve operators reported that they were aware of the presence of errors, they could not characterise how the errors differed between the control and display errors. This study placed the default pump rate value as 20 litre/s, referring to the sample values of Muir and Moray (1996, Table 1, p. 434). However, the pump and heating subsystem values shown in the mimic display of the pasteurization plant simulation was 80 litre/s (Figure 1, p. 430). The larger values in the previous study might have led operators to better understand the plant mechanism and manipulations. This means that if the default values of pump and heating subsystems were higher, the study may yield different results. Lastly, as Muir and Moray (1996) also described, the trust data obtained by the questionnaire encompasses a multicollinearity issue. The current study also observed multicollinearity problems, and additional separate analyses regarding trust development were conducted to avoid this issue. Future study could investigate the relation of Muir and Moray (1996)'s three-factors with other trust development structures (e.g. Lee & See, 2004; Hoff & Bashir, 2015) with respect to such statistical problem.

Chapter 3. Impacts of Prior Instruction and Automation Failure on Driver Trust in Partially Automated Vehicles

3.1 Introduction

With the successful introduction of advanced driver assistance systems (ADAS), such as adaptive cruise control (ACC) and lane keeping assistance system, driving automation has been expected to transform the role of human drivers with promising outcomes, for example, reduction of traffic accidents caused by human error (Choi & Ji, 2015) and improving traffic efficiency (Payre et al., 2014). The taxonomy of driving automation defined by the Society of Automotive Engineers (SAE International, 2018) has been widely used to account for different levels of automation and specific features of automated vehicles. Vehicles currently on the market have already adopted lower levels of driving automation systems and sophisticated ADAS (e.g., Tesla Model S and Volvo S90). Following the taxonomy, lower levels of automation require driver monitoring of roadway situations and of the status of the automated systems. In particular, a partial driving automation (SAE level 2) demands a supervisory intervention of the driver to resume vehicle control when necessary, for example, when the system encounters functional constraints, such as failure of traffic object detection. Therefore, drivers need to understand automation functioning and its failures to use it appropriately (Rajaonah et al., 2006; Seppelt & Lee, 2007). That is, insufficient understanding of interaction with the system or passive monitoring task leads to poor performance in decision making (Louw et al., 2017) and inappropriate use of automation (Sarter et al., 1997). Here, trust is an important component in the successful use of automation (Lee & See, 2004; Parasuraman & Riley, 1997).

In the context of the use of automation, trust is one of the fundamental determinants of relying on a complex automated system instead of human beings to perform tasks (Sheridan &

Ferrell, 1974). People are likely to use automation that they trust, and disuse it when they do not trust it (Muir & Moray, 1996; Lewandowsky et al., 2000; Choi & Ji, 2015). Many existing studies have followed the definition of trust in automation conceptualised by Lee and See (2004) as 'the attitude that an agent will help achieve an individual's goal in a situation characterized by uncertainty and vulnerability' (p. 51). Trust is associated with not only automation use but also other aspects, such as reducing workload (Wickens & Holland, 2000), acceptance of automation (Ghazizadeh et al., 2012), and traffic safety (Payre et al., 2016). This indicates that establishing trust is a key element for the successful acceptance of driving automation technology (Lee & Kolodge, 2019). User interaction with automation leads to an increase in human–machine trust levels (Lee & Moray, 1992; Parasuraman & Riley, 1997). Likewise, continuously experiencing automated driving systems increases driver trust in the driving automation (e.g., Hergeth et al., 2016; Körber et al., 2018a; Kraus et al., 2019). Given that today's drivers are not familiar with the functions of driving automation and resuming manual vehicle control when the failures of automation occur, the loss of driver trust is likely to occur due to the lack of driver understanding of driving automation.

Providing information to operators is a means to obtaining a better understanding of automated systems leading to appropriate levels of trust. Continuous feedback from automated systems led to the appropriate use and reliance on the system even in imperfect ACC automation, and it kept drivers informed (Seppelt & Lee, 2019). Du et al. (2019) investigated the importance of prior information and concluded that providing information regarding automation prior to a failure led to higher levels of driver trust than providing the information after the failure. An explanation of the automation failure is not likely to have considerable impacts on prompting trust compared with providing a preliminary description (e.g., Körber et al., 2018b). In this sense, it seems that the more drivers know the behaviour of the system beforehand, the more they trust it. However, deep and comprehensive knowledge does not always ensure the establishment of appropriate trust: contradictory views exist in the literature in terms of prior information about driving automation. Victor et al. (2018) reported that a detailed description of automated driving systems led to drivers' excessive trust in automation. In contrast, the work of Hergeth et al. (2016) found that drivers who received prior explanation about system limitations showed low levels of trust compared to drivers who did not receive the information in the early stage. In addition, Beggiato and Krems (2013) reported that a preliminary introduction including potential critical situations lowered the levels of trust formation in the initial stage.

Besides human-machine interaction and operators' level of knowledge, automation failure resulting in critical situations has been considered a crucial factor affecting users' trust. Because system performance has a large impact on trust (Lee & Moray, 1992; Hancock et al., 2011), automation should be designed to promote safety and trust (Ghazizadeh et al., 2012). The imperfection of vehicle automation can impose the risk of traffic accidents on drivers, leading to decreased levels of trust in automation. Given that automation failure is a fault that betrays user expectations of automation performance, two types of failures can be considered in the partial driving automation: system limitation and malfunction. System limitation is regarded as a part of the system design and predicted by prior informed knowledge and experience of the automation, for example, a weak ability to detect lane marks in bad weather conditions (Dikmen & Burns, 2016). However, system malfunction represents errors in mechanics or software and an unforeseen part of the automation; identifying failures caused by malfunction is difficult because of the various ways malfunction is manifested (McClellan, 1994). Many existing studies in different automation domains have described that automation failure leads to a decrease in trust, but it can be re-established by subsequent experience of automation (e.g., Dzindolet et al., 2003; Lacson et al., 2005). However, driver knowledge of the automation failures prior to automation experience does not ensure high levels of trust, and experiencing system malfunction can result in different forms of trust toward automation. Therefore, there is a need to examine levels of trust as well as the elements governing driver trust when information about automation failures is provided.

Trust is a multidimensional concept and evolves with shifts of its bases (Hoff & Bashir, 2015), for example, reliability (Madsen & Gregor, 2000) or robustness (Sheridan, 1988). As described in the Introduction, Muir (1994) first proposed a framework of trust in automation, including three central dimensions of trust in automated machines: predictability, dependability, and faith. The dimensions were defined based on the study of interpersonal relationships conceptualised by Rempel et al. (1985), with the premise that trust in automation can be formed by interaction among the three dimensions, similarly to the development of trust between humans. A subsequent study (Muir & Moray, 1996) that experimented with human–machine interaction under supervisory control settings reported that the trust of general users in unfamiliar machines evolves from faith, then dependability, and finally predictability. The feature of Muir and Moray's (1996) study has been widely applied to classify trust attributions in the literature, such as the three trust bases of trust proposed by Lee and See (2004): purpose (faith), process (dependability), and performance (predictability). Furthermore, it is expected to be an approach to examine the shift of factors governing the dynamics of trust for untrained individuals who will begin to interact with automated cars as well as to facilitate vehicle design (Walker et al., 2016). However, relatively few studies have corroborated this foundational work with empirical findings in the context of driving automation. To clarify a key factor predicting trust and to create a better vehicle design leading to appropriate levels of trust and automation use, the concept of Muir (1994) is useful to offer comprehensive insights into driver trust in automated vehicles.

Given examining driver trust toward vehicle automation, demographic factors, such as age, gender, or nationality, should be considered. Older drivers' general attitude toward self-driving cars is relatively negative compared to that of younger drivers (Hulse et al., 2018). Surveys in terms of driver age such as that conducted by Lee et al. (2017) have revealed that younger drivers showed a more positive attitude towards fully automated vehicles than older drivers. Payre and Cestac (2013) showed a contradictory view that older drivers were likely to adopt automated vehicles. Ro ̈del et al. (2014) reported that drivers' intention to use automated vehicles increased with increasing drivers' age. With respect to gender, female drivers are less likely to adopt automated vehicles than male drivers, indicating that women were uncertain about the use of autonomous cars rather than men (Ipsos MORI, 2014). However, KMPG (2013) found that females were more interested in self-driving cars owing to their benefits than males. Hulse and his colleagues (2018) found no differences in general attitudes toward self-driving cars between respondents, driver, and non-drivers. In addition, the trip characteristics influence driver acceptance of automated vehicles (w.g., Bansal & Kockelman, 2016; Krueger et al., 2016). Recent studies on driver acceptance in terms of demographic traits suggested recommendations for policymakers, but the exploration of driver acceptance from various perspectives is still needed.

Other factors besides demographic traits influence drivers' intentions and willingness to adopt new technology. Human-machine trust in automation is playing an important role in shaping driver acceptance (Muir, 1987; Lee & See, 2004). Indeed, Hillary et al. (2017) investigated consumers' trust, finding a willingness to adopt autonomous vehicles as alternatives to their vehicle. To describe the basis of trust with respect to goal-oriented information that needed to support appropriate trust, Lee, and Moray (1992) first proposed a three-dimensional (3P) model: purpose, process, and performance which revised by Lee and See (2004). Purpose refers to why the automation was designed, the process represents how the automation functions, and performance refers to how the automation is operating. Therefore, automation designers are expected to consider these three categories to determine an operator's trust in automation. Given that myriad studies are investigating general users' acceptance of and trust in automated vehicles (König & Neumayr, 2017; Kaur & Rampersad, 2018), various insights have been suggested with respect to the high levels of driving automation. To the best of our knowledge, however, there has been little discussion about relationships between demographics and the trust dimension in contemporary automated vehicles that are designed following the SAE Level 2. In addition, while the 3D model by Lee and Moray (1992) has been applied in many studies about human-machine trust (e.g. Chien et al., 2014), the information that should be provided to drivers during partially automated driving has not been closely examined with empirical data.

The primary aims of STUDY II is to examine the application of previous finding obtained from the STUDY I which the main contributor of trust toward unfamiliar machines is the feeling of dependability. The objective of this study was to identify the key dimensions that influence driver trust development the most, with the premise that driver trust may be predicted by different dimensions and formed in different ways depending on the levels of knowledge that drivers have in advance of vehicle automation experience and the type of automation failure. This driving simulator study adopted the framework of trust development proposed by Muir (1994) to determine what creates the shift between trust and distrust, when the shift occurs, and the point where levels of trust significantly change throughout driver interaction with partial driving automation. Further, the current study attempted to explore the relationship between demographic factors and driver trust in driving automation. Drivers' subjective ratings of trust were interpreted regarding the 3P model (Lee & Moray, 1992; Lee & See, 2004) to determine the information needs for appropriate automation design. Additionally, the attribute of distrust, which is an opposite concept of trust and refers to a negative expectation of consequences by automation, was measured to explore comprehensive understandings of trust. Following research hypotheses were built for examining driver trust:

- H1: Having not only generic information about driving automation but also detailed prior information of the reason why automation failure occurs leads less decrease of trust after the intervention task in comparison with having only the generic information.
- H2: Both system-limitation or -malfunction lead decreased trust. Decreased trust due to the system-limitation can be rebuilt by the subsequent experience of flawless automation, whilst the decrease by the system-malfunction cannot be recovered.
- H3: Shaping trust is depending on the levels of driver knowledge in terms of driving automation and the automation failure type.
- H4: Dependability is a main contributor determining driver trust toward vehicle automation.
- H5: Dispositional qualities leads different forms of trust. To be specific, daily drivers, non-student drivers, and female drivers are likely to show relatively less levels of trust compared to non-daily drivers, student drivers, and male drivers.

- H6: As dependability was found as a significant attributor of trust formation, driver trust toward the dimension of process is likely to be higher than the dimensions of purpose and performance.

3.2 Methodology

3.2.1 Participants

Fifty-six drivers (28 males) who aged from 19 to 75 (M_{age} =31.7 years, SD_{age} = 14.7) were recruited for the current experiment. They had held a valid driver's license and no experience of any activities in terms of driving automation – e.g., driving simulation and on-road driving experiment. Drivers who were comfortable speaking and reading Japanese were recruited because of the appropriate understanding of descriptions and questionnaires. They received reimbursements for their participation (JPY 1660).

3.2.2 Apparatus

The current experiment was conducted by using a driving simulator comprising of a driving sheet, a steering wheel by MOOG Inc., an accelerator and a brake pedal at the University of Tsukuba (Figure 3.1). The driving scenarios generated by the D3Sim by MITSUBISHI PRECISION Co. LTD. were projected five displays by SHARP Corporation. An LCD monitor was used to present the speedometer and human-machine interface and audio was paired with the driving environment and played through two speaker systems. A video camera recorded the driver's side-backward view with the driving simulation scene and the recorded video.

Figure 3.1 Fixed-base Driving simulator.

3.2.3 System Description

Driving automation systems in the current study performed both longitudinal and lateral vehicle controls. The longitudinal controller was designed like Adaptive Cruise Control. The controller maintained vehicle speed at 80 km/h which was a maximum speed. For instance, when a merging vehicle appeared in front of host vehicle in a merging lane, the system automatically reduced speed for avoiding rear-end collision. The lateral controller was designed like Lane Keeping Assistance system, and the automation was able to adjust vehicle position on the centre of driving lane. Further, the automated driving system carried out lane change when there were slow leading vehicles rather than 80 km/h. The drivers could activate the automation by pressing a button on the steering. When the input of steering wheel was greater than 5 Nm as well as the input of accelerator or brake pedal was greater than 5%, the automation was deactivated. That is, the driver could disengage the automation. When entering a simulated drive, drivers were encouraged to engage the available automation system, thus no participants made an attempt to disengage from the automation.

Human-machine interface (HMI) consisted of a visual display in the auditory output, instrument clusters and display on the dashboard (Figure 3.2). When the automated driving system activated, visual cluster was projected on dashboard. The presented visual cluster and auditory output were changed depending on situations – i.e., detecting a leading vehicle and a merging vehicle, and system's request to intervene driving task in automation failure. As shown in Figure 3.2, four visual clusters were prepared to provide information about system status with the drivers.

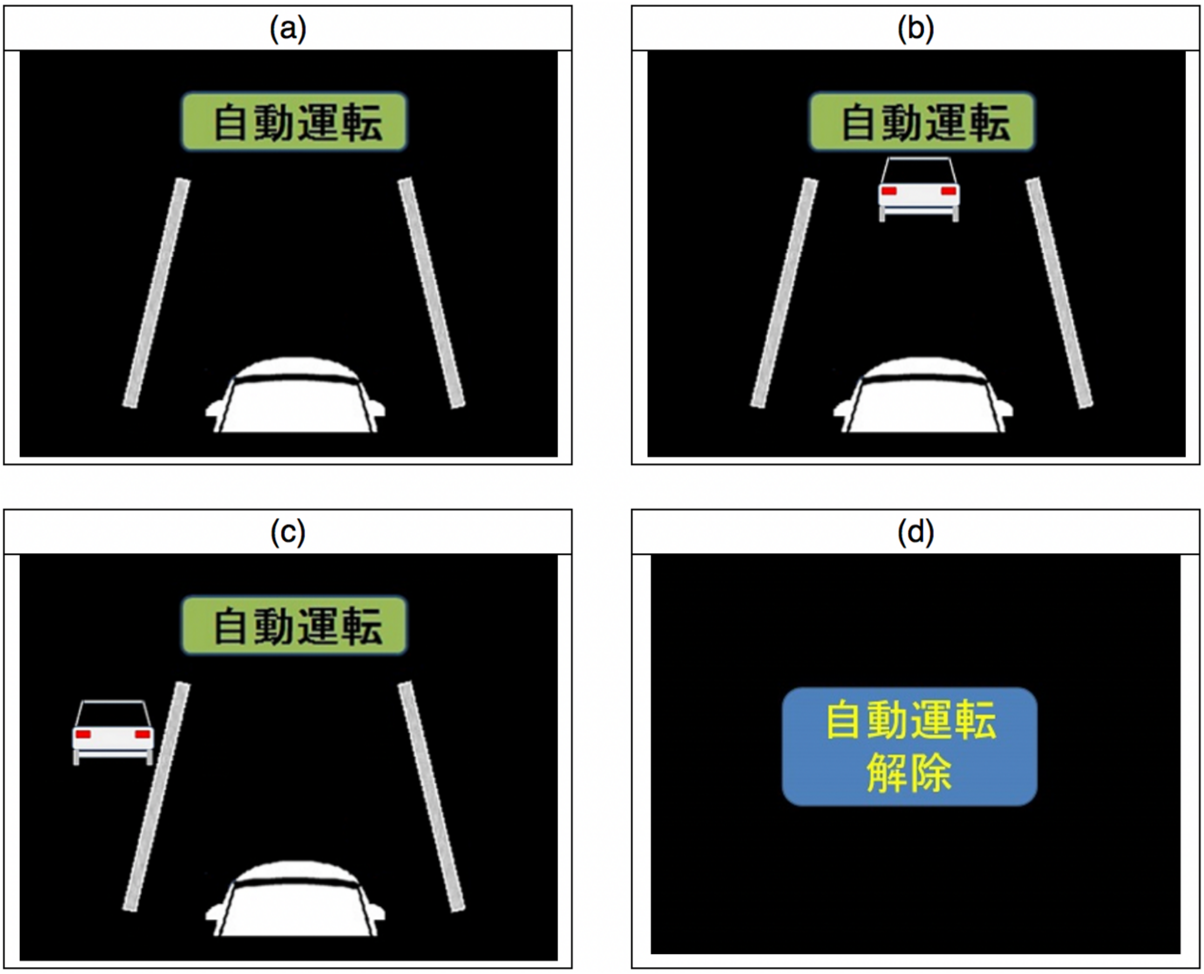


Figure 3.2 HMI screens: (a) driving automation activation, (b) detection of a leading vehicle, (c) detection of a merging vehicle, and (d) driving automation disengaged.

3.2.4 Driving Scenarios

Each driver was instructed about how to disengage the execution of driving control – e.g., grasping a steering wheel or pressing pedals. The Japanese highway road with two lanes was simulated by the driving simulator, and four trials were presented to all participants. Each trial encompassed several events those the automated driving system was able to handle.

1st trial included three driving events. First, the driver encountered three leading vehicles those drove in 40 km/h. When time-headway between the host vehicle and the leading vehicle was equal to 4 sec, the automated driving system started to change driving lane from left to right for 5 sec, then came back to the left lane after taking over the three vehicles. Secondly, a fog was presented on the straight road scene for approximately 1 min. Here, the fog reduced the driver's visibility, but the system was able to drive coping with this situation. Thirdly, a merging vehicle was appeared in a merging lane. In this situation, the system gave way to the merging vehicle by reducing vehicle speed from 80 to 70 km/h because the host vehicle could not change the lane due to four vehicles in the next lane. After the successful merging, the host vehicle increased speed to 80 km/h.

2nd trial contained two driving events. The merging car event occurred again. Contrary to the merging vehicle event in the first trial, here, the system changed driving lane from left to right after detecting the merging vehicle. After certain time, the system execute lane changing for coming back to the initial lane. Secondly, a vehicle on the right attempted to change driving lane to the left. The system decreased vehicle speed to 70 km/h, then recovered speed to 80 km/h after the successful lane changing.

In 3rd trial, the intervention task was presented to drivers. Here, two types of automation failures were prepared: limitation and malfunction. The driver entered a curve road with 800R curvature in both limitation and malfunction events. In the malfunction scenario, the system issued the short notice when faced the curve road. In the limitation scenario, the fog was suddenly appeared to the driver, and the notice was issued. In both scenarios, the notice was issued for 5 sec, and the system was disengaged. After drivers' successful regaining driving control, the third trial was finished.

4th trial consisted of two events. The drivers experienced the slow leading vehicle event and merging vehicle event which were identical to those of the first trial.

3.2.5 Experimental Design

A 2 × 2 × 6 repeated measures mixed design with between-subject factors: the levels of knowledge (relatively detailed and less) and the type of automation failure (system-limitation and -malfunction) and with a within-subject factor: the point of measurement (Inst, Practice, 1st trial, 2nd trial, 3rd trial, and 4th trial) was used for the current experiment. The sequence of whole experiment was identical regardless of two factors, however, the participants experienced different automation failure in third trial according to the factor of the type of automation failure. Therefore, there were four experimental designs: relatively detailed and system-limitation (Detailed-Lim), relatively detailed and system-malfunction (Detailed-Mal), relatively less and system-limitation (Less-Lim), and relatively less and system-malfunction (Less-Mal), and 14 participants were allocated for each condition.

All participants were instructed about the description of automation, situations where the system was capable of coping with, the presence of intervention task and what they will be required in the situation (description A), and a trust questionnaire. Further information about intervention task (description B, C) was presented to the drivers in the relatively detailed knowledge group. Information of Fig. 2-(d) was presented to only participants in the Detailed group. Full description of instruction is as follows:

A. The driving automation system stops its driving control in certain situations. In this case, the participant will be required to drive (take over driving tasks).
B. HMI instruction when the system issues an intervention task to the driver: There is some situations where the system cannot maintain safe driving according to the traffic situations or weather. The visual instruction of "Disengagement of automated driving" will be presented on the dashboard with the three times beep sound. In this case, you should takeover driving operation immediately.
C. Specific situations where the system issues the intervention task: (1) sensor error due to environmental factor – i.e., heavy fog or heavy rain. (2) malfunction of automated driving system when supplement – i.e., sensor is mechanically broken down.

3.2.6 Procedure

Figure 3.3 describes the experiment procedure. Upon arrival, the participants were provided with an overview of the current experiment: the driving automation will control dynamic driving tasks on behalf of the participant, and they should resume controls when the system issues the short notice. The description was different depending on the knowledge groups. After the instructions, they filled out the first trust questionnaire. The participants moved to a practice drive, which lasted approximately 5–10 min depending on their sense of familiarity with the automation. After the practice, the second trust questionnaire was given to the participants. The four trials were presented to the participants, and they experienced system limitation or malfunction in the 3rd trial. As mentioned above, all trials contained the merging vehicle event, except for the 3rd trial. All participants were asked to answer a questionnaire about driver trust after each trial. Altogether, the whole experiment lasted approximately 40 min, and participants completed six trust questionnaires. There was a 10-minute break between the 2nd and 3rd trials.

Figure 3.3 Procedure of experiment in STUDY II.

3.2.7 Dependent Variables

The participants completed six trust questionnaires in the course of the experiment (after the instruction, practice, 1st, 2nd, 3rd, and 4th trial). The trust questionnaire proposed by Muir and Moray (1996) was used with a 100 mm-scale and the poles labelled 'none at all (0)' or 'not at all (0)' on the left to 'extremely high (100)' on the right (e.g., Domeyer et al., 2018; Du et al., 2018). The trust questionnaire included four items about feelings of predictability, dependability, faith, and trust in the driving automation (Table 1). For each measurement point, we asked the participants to answer the question: 'Do you trust the automated vehicle?' with the options: 'I trust the automated vehicle (Yes)' and 'I do not trust the automated vehicle (No)'.

Table 3.1 Constructs examined in subjective trust questionnaire (adapted from Muir & Moray, 1996).

Construct	Relevant questions
Predictability	To what extent can the behaviour of driving automation be predicted from moment to moment?
Dependability	To what extent can you count on the driving automation to do its job?
Faith	To what extent can you count on the driving automation to do its job?
Trust	To what extent do you trust the driving automation?

This study used the questionnaire of trust dimension proposed by Chien et al. (2014). The participants were asked to rate their likelihood using a seven-point rating scale (where 1 = "Disagree completely", 2 = "Disagree moderately", 3 = "Disagree somewhat", 4 = "Neither agree nor disagree"; 5 = "Agree somewhat", 6 = "Agree moderately", 7 = "Agree completely"). The questions required the participants to rate the automation on the following aspects: distrust, purpose, process, and performance. Chien et al. (2014) developed this questionnaire by adopting existing questionnaires in terms of technology acceptance (Davis, 1989) as well as human-automation trust (Jian et al., 2010). They performed the validations of the questionnaire to examine the relationship between cultural factors and the three trust dimensions (Chien et al., 2014; Chien et al., 2015). Table 3.2 describes the constructs and statements of the trust questionnaire.

Table 3.2 Statement summing up four attitudes toward the automated vehicle

Construct	Statement
Purpose	I can rely on automation to ensure my performance
	I am confident in automation
	Automation does not fail me
Process	It is easy to follow what automation does
	Automation is friendly to use
	Automation uses appropriate methods to reach decisions
	I understand how automation works
Performance	Automation improves my performance
	Using automation increases my productivity
	Using automation enables me to accomplish tasks more quickly
Distrust	Automation may result in unpredictable situations
	I believe automation could make errors
	I am wary of automation
	I am suspicious of automation's intent
	Automation is deceptive
	Automation behaves in an underhanded manner

3.2.8 Statistical Analyses

A 2 × 2 × 6 analysis of variance (ANOVA) was performed on the subjectively rated driver trust in the automated vehicle, and post-hoc comparisons were conducted with a Bonferroni alpha correction. There were three factors in this analysis: the levels of knowledge (detailed, less) and the type of automation failure (limitation, malfunction) as between-subject factors and the point of measurement (inst, practice, 1st trial, 2nd trial, 3rd trial, and 4th trial) as a within-subject factor. To further explore the effect of the subsequent experience of flawless automation on the recovery of decreased trust due to the automation failure, a 2 × 2 × 2 ANOVA was conducted with the two between-subject factors and two measurement points (3rd trial and 4th trial) as a within-subject factor. Stepwise regression was conducted to identify the best predictor of trust at each measurement point regarding the level of knowledge. Additional stepwise regressions were implemented to investigate the best predictor from the 3rd to the 4th trial. The issue of multicollinearity was handled by calculating variance inflation factors (VIFs), and if a predictor with VIF larger than 10 was observed, additional analyses were conducted (Kabacoff, 2020). To examine the difference in responses about trust in driving automation, Chi-square tests were performed with regard to the knowledge level and each measurement point. The difference between the Pre (2nd) and Post (3rd), and the Subsequent (4th) and Post (based on the trust ratings in Post) was calculated to examine whether which has larger impacts on driver trust and tested by t-tests. When the difference of trust ratings between the 4th and 3rd trial is equal to or larger than those between the 2nd and 3rd trials, this study considered that driver trust is recovered. In other words, when the difference between the 4th-3rd trials is significantly lower than those between the 2nd-3rd trials, it was considered that driver trust is not repaired (see Table 3.3). The reliability for each questionnaire item was calculated by Cronbach's alpha, and it exhibited very high internal consistencies of predictability (α = .95), dependability (α = .90), faith (α = .94), and trust (α = .91). All statistical analyses were implemented with the R (R Core Team, 2019).

Table 3.3 The meanings of differences of trust ratings between the 2nd-3rd trials, and the 4th-3rd trials.

	4-3 trials positive	4-3 trials negative
2-3 trials positive	Decreased-Recovered	Increased-Increased
2-3 trials negative	Decreased-Unrecovered	Increased-Decreased

3.3 Results

3.3.1 Driver Trust in The Course of Automation Experience

Figure 3.4. shows the changes of trust ratings by the measurement point across the levels of knowledge. The ANOVA showed that the subjective ratings on trust of drivers significantly differed across the measurement point, $F(5, 270) = 38.61$, $p < .001$, $\eta_G^2 = .13$. However, no significant effect of the knowledge levels observed, $F(1, 54) < .001$, $p = .99$, $\eta_G^2 < .001$. The two factors did not significantly interact, $F(5, 270) = .54$, $p = .75$, $\eta_G^2 = .002$.

Table 3.4 describes the results of post-hoc comparisons towards the point of measurement. The comparison observed that trust ratings after the instruction and the practice drive were significantly lower than those at later all measurement points, the 1st to 4th trials, all *p*s < .05. The trust ratings did not differ between the 1st to 3rd trials. However, the rating in the 4th trial was significantly higher than that of 3rd trial, $M = 61.54$ vs. 67.76, $p = .01$.

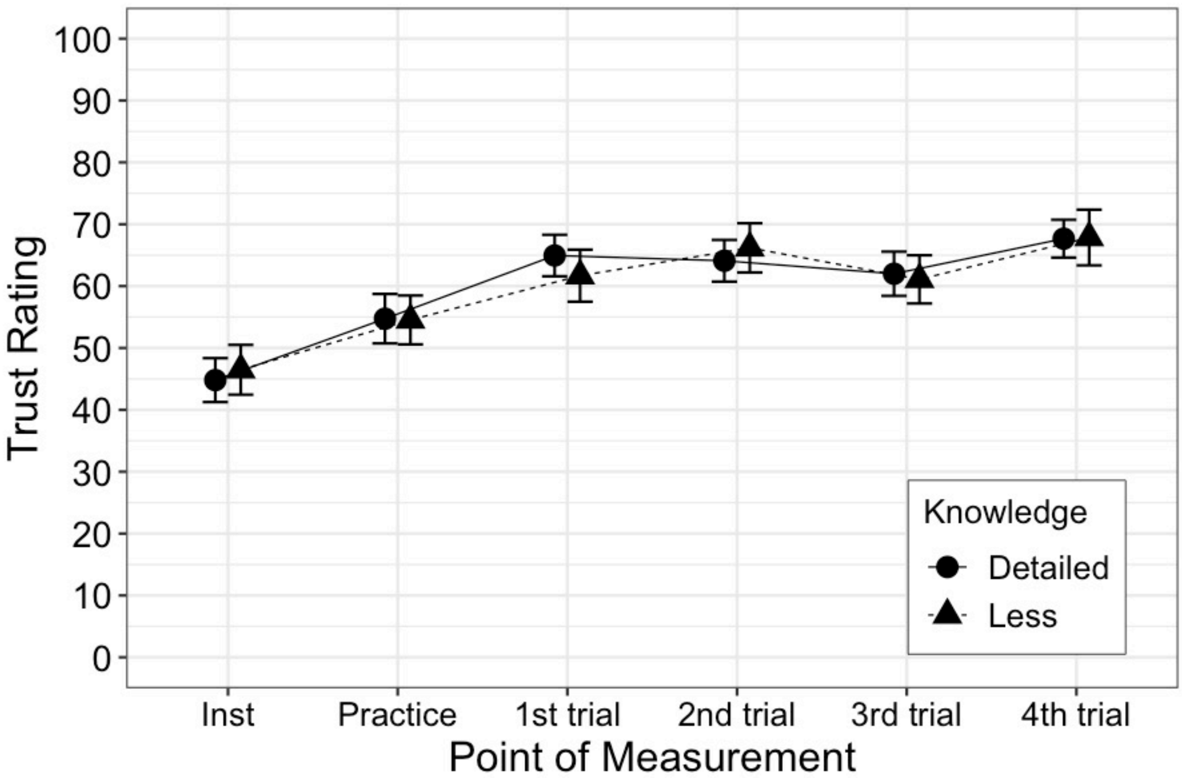


Figure 3.4 Changes of driver trust ratings in the course of the experiment across the knowledge levels.

Table 3.4 Post-hoc comparison results.

	Inst	Practice	1st trial	2nd trial	3rd trial	4th trial
Inst	-	< .001**	< .001**	< .001**	< .001**	< .001**
Practice		-	< .001**	< .001**	< .001**	< .001**
1st trial			-	n.s.	n.s.	.243
2nd trial				-	.771	n.s.
3rd trial					-	.012*
4th trial						-

note. * $p < .05$, ** $p < .01$.

n.s. means no significance was observed ($p = 1$).

Figure 3.5. illustrates changes of trust ratings across system transparency at the three measurement points (2nd, 3rd, and 4th trial) in response to the automation failure type. The 2 x 2 x 2 ANOVA revealed a significant main effect of the measurement point on driver trust, $F(1, 52) = 33.53$, $p < .001$, $\eta_G^2 = .024$, as described in the section of 3.1. The ANOVA implies a possible significant interaction between the measurement point and the automation failure type, $F(1, 52) = 2.92$, $p = .09$, $\eta_G^2 = .002$. Post-hoc comparisons revealed that the ratings after the 3rd trial were higher than those after 4th trial for the drivers in the Limitation group, $M = 65.06$ vs. 67.58, $p = .03$. This trend was also observed for the drivers in the Malfunction group, $M = 58.94$ vs. 67.75, $p < .001$. The remaining effects were not significant, all ps > .22.

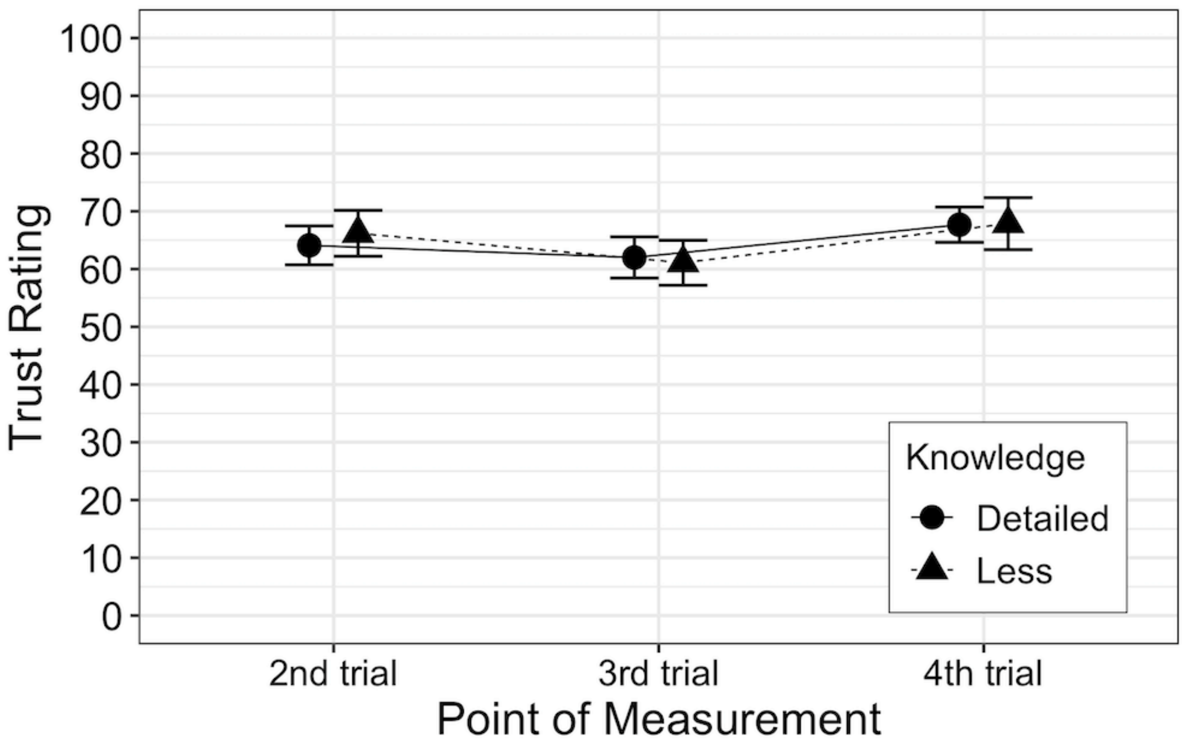


Figure 3.5 Changes of driver trust ratings between the 2nd and 4th trial across the knowledge levels.

Table 3.5 summarised the descriptive statistics of the differences across each condition and statistical values by the t-test. As shown in Table 3.5, the mean difference of trust ratings between the Pre and Post indicated that the intervention task led decreased levels of trust in automated driving system. Accordingly, the mean difference between the Subsequent and Post showed greater values than those between the Pre and Post. It meant that driver trust was successfully recovered by the subsequent experience. Also, paired samples t-tests examined the differences between the Pre and Post, and the Subsequent and Post.

The t-test resulted no significance of differences between the Pre and Post, and Subsequent and Post in the Detailed-Limitation condition. This result indicated that a subsequent experience of error-free system could recover decreased trust by the limitation experience. Accordingly, the identical trends were observed in the Less-Limitation condition and in the Less-Malfunction condition. However, a significant difference was observed in the Detailed-malfunction condition. This implies that the drivers who have the detailed information of the system rated trust more after the malfunction experience than before the experience.

Table 3.5 Descriptive values of the differences of trust ratings between the Pre and Post, and the Subsequent and Post for each condition.

		Pre-Post	Subsequent-Post		
		M (*SD*)	*M* (*SD*)	t	p
Detailed	Limitation	1.22 (7.11)	2.52 (4.87)	-.58	.572
	Malfunction	2.97 (10.83)	8.82 (9.05)	-2.34	.036*
Less	Limitation	6.04 (7.48)	6.24 (7.77)	-.06	.951
	Malfunction	4.15 (9.69)	7.29 (9.61)	-1.14	.274

Note: * means $p < .05$.

Further, this study examined whether the t-test result is convincing or not by checking the difference by the participant (Table 3.6). Table 3.3 summarized the meanings of differences of trust ratings between the 2nd-3rd trials, and the 4th-3rd trials. For example, when the difference which subtracted trust after the 3rd from those of the 2nd trial is positive, it means that the occurrence of intervention task decreased driver trust.

Table 3.6 The number of participants according to the each meaning of the difference and group.

	Detailed-Limitation	Detailed-Malfunction	Less-Limitation	Less-Malfunction	Total
Decreased-Recovered	7	8	11	6	32
Decreased-Unrecovered	2	0	1	1	4
Increased-Increased	4	5	1	6	16
Increased-Decreased	1	1	1	1	4

3.3.2 The Development of Driver Trust in Partial Vehicle Automation

Stepwise regressions were performed at each measurement point following the hypothesis proposed by Muir and Moray (1996): *Trust* = *Faith* + *Dependability* + *Predictability*. Table 3.7 illustrates the stepwise regression results with the factor of the levels of knowledge and the measurement point. Dependability consistently as well as best accounted for trust of drivers in the Detailed group throughout the course of the experiment. The similar trend was observed in the Less group, however, faith best predicted their trust after undertaking the practice.

The stepwise regression was performed at the point of the 2nd, 3rd and 4th trials across automation failure type (Table 3.7). Looking into the best determinant of trust after the 2nd trial, dependability accounted for trust in automation for the drivers in the Detailed-Lim, Detailed-Mal, and Less-Lim groups, and faith governed trust for the drivers in the Less-Mal group.

With respect to the point after the 3rd trial, faith and dependability were the best predictors of trust for the drivers in the Detailed-Lim group. However, dependability and predictability were the best predictor for the drivers in the Detailed-Mal group. For the drivers in the Less group, dependability the predicted trust when both the limitation and malfunction experiences.

Considering the point after the 4th trial, dependability governed trust of the drivers in the Detailed group regardless automation failure type. Dependability also best accounted for trust for drivers of the Less-Limit group. For the drivers of the Less-Mal group, dependability and faith best predicted their trust.

Table 3.7 Stepwise regression results. P = Predictability, D = Dependability, F = Faith.

		R^2 *P*	R^2 *D*	R^2 *F*	Stepwise Best Predictor	*F* (Best Model)	*p*
Detailed	Inst	0.41	0.76	0.67	*D, P, F*	51.08	< .001
	Practice	0.41	0.62	0.61	*D*	44.44	< .001
	1st trial	0.42	0.79	0.67	*D*	102.47	< .001
	2nd trial	0.41	0.66	0.55	*D*	48.42	< .001
	3rd trial	0.56	0.79	0.72	*D, F*	65.83	< .001
	-Lim	0.05	0.44	0.45	*F, D*	7.4	< .001
	-Mal	0.90	0.96	0.94	*D, P*	229.9	< .001
	4th trial	0.69	0.90	0.64	*D, P*	147.15	< .001
	-Lim	0.58	0.88	0.38	*D*	89.23	< .001
	-Mal	0.90	0.96	0.90	*D*	247.7	< .001
Less	Inst	0.45	0.72	0.66	*D, F*	60.09	< .001
	Practice	0.55	0.76	0.85	*F, D, P*	93.55	< .001
	1st trial	0.58	0.83	0.71	*D, F*	74.34	< .001
	2nd trial	0.56	0.77	0.64	*D, F*	54.59	< .001
	3rd trial	0.58	0.81	0.71	*D, F*	63.11	< .001
	-Lim	0.61	0.81	0.69	*D*	53.41	< .001
	-Mal	0.53	0.81	0.74	*D*	52.96	< .001
	4th trial	0.76	0.88	0.79	*D, F*	126.79	< .001
	-Lim	0.76	0.94	0.74	*D*	162.79	< .001
	-Mal	0.76	0.88	0.86	*D, F*	66.45	< .001

3.3.3 Demographic Factors

Table 3.8 shows the number of drivers and descriptive values of the drivers' ages and years of licensure across their genders, occupations, and frequencies of driving.

Table 3.8 Descriptive values of drivers' ages and driving experience across gender, occupation, and frequency of driving

			Age	Year of Licensure
		N	*M (SD)*	*M (SD)*
Gender	Female	28	33.96 (15.10)	14.58 (14.96)
	Male	28	29.79 (14.30)	10.69 (13.75)
Occupation	Student	36	22.42 (1.65)	3.43 (1.58)
	Non-Student	20	48.90 (12.15)	29.2 (12.02)
Frequency	Daily	30	40.17 (15.98)	20.82 (15.52)
	Non-Daily	26	22.31 (1.69)	3.21 (1.57)

Gender. Independent-samples t-tests revealed no significant differences of age ($t(54) = 1.06$, $p = .29$), driving experience ($t(54) = 1.01$, $p = .32$)), and driving mileage per week ($t(54) = 1.64$, $p = .11$) between female and male drivers.

The 2 × 3 repeated-measures ANOVA revealed a main effect of the three trust dimension on driver trust ($F(2, 108) = 6.55$, $p = .002$, $\eta_G^2 = .04$). Post-hoc tests showed significantly lower levels of trust for the dimension of purpose than for the dimensions of process and performance (ps = .01). There were no effects of gender, ($F(1, 54) = .01$, $p = .93$, $\eta_G^2 < .001$) and no interaction between gender and dimension factors ($F(2, 162) = 1.22$, $p = .30$, $\eta_G^2 = .01$). Figure 3.6 describes the mean ratings of distrust and trust for each dimension across the gender factor.

Female drivers rated higher levels of trust than male drivers ($t(54) = 2.04$, $p = .046$, M = 73.11 vs. 62.41). An independent-samples t-test did not reveal significant differences in distrust ratings between female and male drivers ($t(54) = .18$, $p = .86$, M = 4.58 vs. 4.55).

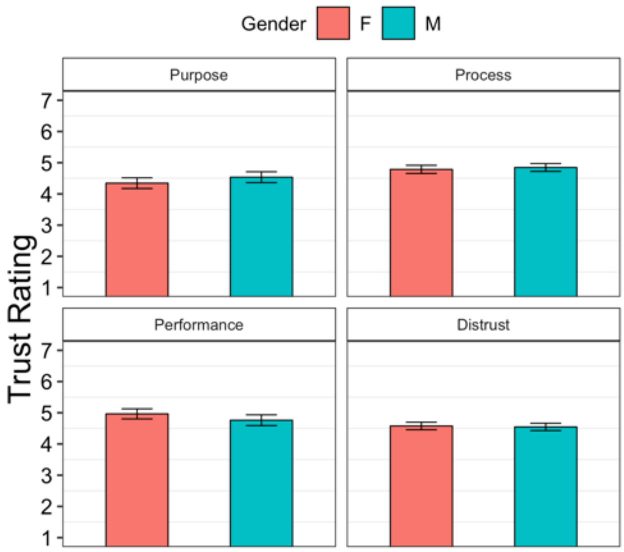


Figure 3.6 Female and male drivers' trust across the dimensions of purpose, process, and performance, and the attribute of distrust

Occupation. As shown in Table 1, non-student drivers' ages and driving experience were significantly higher than student drivers' ages ($t(54) = 12.96$, $p < .001$) and driving experience ($t(54) = 12.76$, $p < .001$). The independent-samples t-test revealed that the non-student drivers drove more frequently per week than the student drivers ($t(54) = 3.63$, $p < .001$).

Driver trust ratings differed across the dimension of trust ($F(2, 108) = 6.46$, $p = .002$, $\eta_G^2 = .04$). The post-hoc comparison also revealed that trust ratings for the dimension of purpose were significantly lower than trust ratings for the dimensions of process and performance (ps = .02). The remaining effect and interaction were not significant. Figure 3.7 illustrates the mean ratings of distrust and trust for each dimension across the occupation factor.

There was a significant difference in driver trust ($t(54) = 2.81$, $p < .01$). Non-student drivers exhibited higher levels of trust than student drivers ($M = 4.72$ vs. 4.28). Accordingly, the t-test also revealed that the attribute of distrust did differ significantly ($t(54) = 2.64$, $p = .01$). The student drivers rated higher levels of distrust toward driving automation than the non-student drivers ($M = 4.72$ vs. 4.28). These two statistical results are consistent in revealing that non-student drivers are more likely to trust automated vehicles than student drivers.

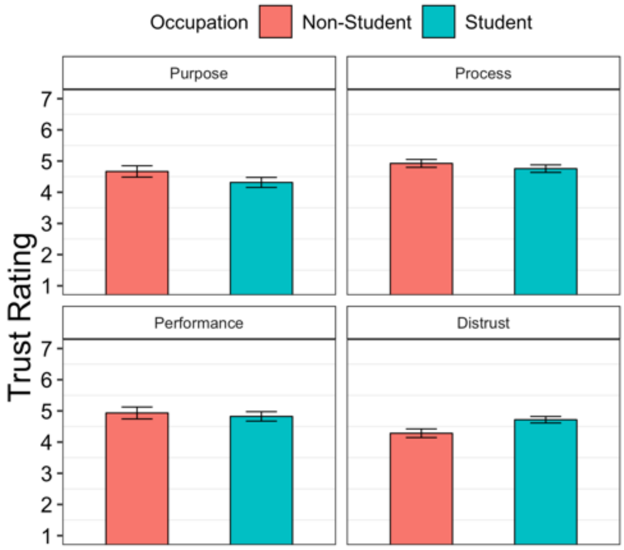


Figure 3.7 Student and non-student drivers' trust across the dimensions of purpose, process, and performance, and the attribute of distrust

Frequency of driving. The t-test revealed that the daily drivers' ages were significantly higher than those of the non-daily drivers (independent-samples $t(54) = 5.67$, $p < .001$). The year of licensure and mileage of the daily driver were longer and higher than those of the non-daily driver($t(54) = 5.76$, $p < .001$, and $t(54) = 4.61$, $p < .001$ respectively).

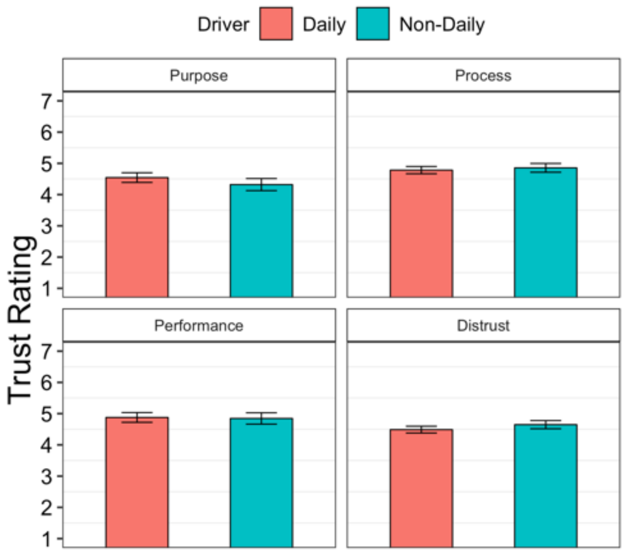


Figure 3.8 Daily and non-daily drivers' trust across the dimensions of purpose, process, and performance, and the attribute of distrust

The trust dimension has a significant effect on driver trust ($F(2, 108) = 6.49, p = .002, \eta_G^2 = .04$). The levels of trust for the dimensions of process and performance were higher than those for the dimension of purpose (ps = .01). The data did not provide substantial evidence for a significant effect of the frequency of driving ($F(1, 54) = .09, p = .76, \eta_G^2 = .001$) as well as interaction between the frequency of driving and the trust dimension ($F(2, 108) = .67, p = .51, \eta_G^2 = .01$). Figure 3.8 describes the mean ratings of distrust and trust for each dimension across the frequencies of the driving factor.

No significant difference in trust was found between daily and non-daily drivers ($t(54) = 1.26, p = .25, M = 70.9$ vs. 64.15). For the attribute of distrust, the same result was found ($t(54) = .95, p = .35, M = 4.49$ vs. 4.65).

3.3.4 Driver Willingness of Automation Trust and Usage

Table 3.9 displays the number of drivers who answered 'Yes' to the question about willingness to trust automated vehicles across each experimental group. The experience of driving automation led to an increased willingness to trust the automated vehicle. As shown in Table 3.9, half of the drivers for the detailed (14 of 28) and less (14 of 28) groups answered 'Yes' after the instructions. The number of 'Yes' responses increased after the 4th trial for both the detailed group (24 of 28) and less group (22 of 28). No change in the number of 'Yes' responses occurred after the 3rd trial for the drivers in the detailed-lim group. However, for the detailed-mal, less-lim, and less-mal groups, the number of respondents who answered 'Yes' after the 2nd trial decreased after the 3rd trial. The subsequent experience of flawless automation led to an increased number of 'Yes' responses in the detailed-lim, detailed-mal, and less-lim groups.

Table 3.9 Overall number of 'Yes' responses to the question: 'Do you trust the automated vehicle?' across the measurement point and experimental groups.

		Inst	Practice	1st trial	2nd trial	3rd trial	4th trial
Trust	Detailed-Lim	6	8	10	11	11	12
	Detailed-Mal	8	9	12	12	11	12
	Less-Lim	10	11	13	12	11	12
	Less-Mal	4	6	8	12	10	10

There was no significant difference in the response number between the detailed and less groups, all $\chi2 < .49$, all ps > .49. Chi-square difference tests with regard to the measurement point found differences in the response number between time points of the inst and all trials, respectively, as shown in Table 3.10. A significant difference was observed between the practice and the 2nd and 4th trials.

Table 3.10 Chi-square statistics (all *df*s = 1).

	Inst	Practice	1st trial	2nd trial	3rd trial	4th trial
Inst	-	1.3	8.56**	14.57**	8.56**	12.91**
Practice		-	3.37	7.54**	3.37	6.3*
1st trial			-	.91	0	.49
2nd trial				-	.91	.06
3rd trial					-	.49
4th trial						-

3.4 Discussions

The current experiment was conducted to identify key dimensions of driver trust in automated vehicles as extends findings from STUDY I which dependability initiates as well as governs human trust in machines. The current study manipulated prior information about driving automation (Detailed, Less) and the type of automation failure (Limitation, Malfunction). The partial driving automation was simulated to clarify how driver trust, such as subjective ratings of trust and dimensions of trust, changes over the experience of automation, and the occurrence of automation failure was handled by drivers' resumption of vehicle control. The levels of knowledge in terms of driving automation had no effect on driver trust in the automated vehicle, and the automation failure type also had no effect on their trust. The presence of failure negatively affected levels of trust, but the decreased trust was recovered by experiencing error-free automation after the intervention task.

3.4.1 Subjective Ratings of Driver Trust on Vehicle Automation

Results of subjective ratings for drivers' trust in automated vehicles show a clear distinction before and after the experience of driving simulation. The drivers were likely to rate higher levels of trust throughout the experience of driving automation consistent with previous findings that the consistent exposure to automation develops users' trust (Lee & Moray, 1992; Parasuraman & Riley, 1997). It was also expected that the levels of trust ratings would be shaped in different ways depending on the levels of knowledge and the failure type (H1, H2). However, trust ratings grew steadily in the identical way regardless of the levels of knowledge as well as automation failure type which they experienced. Our result did not observe a significantly negative effect of the intervention task on driver trust between the 2nd and 4th trial. In spite of no significance, the intervention task in the 3rd trial led lower trust ratings than those of the 2nd trial. The subsequent experience of error-free automation in the 4th trial significantly rebuilt decreased trust due to automation failure in the 3rd trial, and the levels of recovered trust was greater than those of the 2nd trial. This finding implies that drivers may trust the automation more if they are aware of the presence of automation failure by experiencing the intervention task (see Kraus et al., 2019). The results of 'Yes' responses to the question of whether or not drivers trust the automation support the implication. Drivers in the Detailed-Lim group showed no changes by the intervention task, whilst there were decreased numbers of the responses for the other group (Walker et al., 2018). Interestingly, the subsequent error-free automation led the increase of 'Yes' responses for the Detailed-Lim, Detailed-Mal, and Less-Lim groups except for the Less-Mal group. Consistent with findings of Beggiato & Krems (2013), the automation failure does not lead substantial loss of trust if they have detailed information about automation beforehand. Additionally, the result in terms of the 'Yes' response for drivers in the Detailed-Lim and Less-Mal groups indicates impacts of not only the knowledge levels but also automation failure type. The driving simulator study of Kraus et al. (2019) found that malfunction occurrence for drivers who did not have a deep knowledge in terms of driving automation showed larger decrease in trust compared to limitation occurrence for drivers who have a detailed knowledge. Even though automation failure is not likely to have a long-term effect on corrupting driver trust, it can lead distrust of automation. Malfunctioned automation has a larger negative impact on operators' trust than automation limitation (e.g., Muir & Moray, 1996). Therefore, our result implies that malfunction by driving automation is the most critical for drivers who have relatively less knowledge. In order to mediate distrust and a large decrease in trust in such case, a comprehensive description should be provided for drivers.

Our result of trust ratings has implication that the initial experience of system, experience of automation failure, and experience of subsequent flawless automation influence both increase and decrease of drivers' trust ratings. In particular, the trust rating significantly increased from when the description was presented until the practice drive, whilst no increase was observed between the 1st and 3rd trial. This implies that initial and early experience with a system is a key factor influencing operators' trust (Moray & Inagaki, 1999). This study confirmed successful initial human-machine interaction led increased levels of trust in automation.

3.4.2 The Development of Driver Trust

The driving simulator study was designed to observe how the dimension of trust interacted according to the levels of knowledge presented to the drivers in the very initial stage. The current study had expected that the key dimension of driver trust would be changed by all factors prepared in the current study: the levels of knowledge, automation failure type, and continuous experience of automation (the measurement point), and. Driver trust was shaped by only the knowledge levels until the 3rd trial where automation failure occurred. Thus, it was expected that initial trust can be shaped by different dimensions according to the levels of knowledge (H3). The stepwise regression regarding the point after receiving instructions (TQ1) resulted that all three traits initiated trust for drivers who received further information of automation, and dependability and faith were administered as best predictors of trust for drivers who received the general description. Then, only dependability generally kept affecting trust for the drivers in the Detailed group, whilst both dependability and faith kept accounting for trust for the drivers in the Less group. It implies that the detailed instruction attributed predictability to emerging as one of determinants shaping driver initial trust. That is, the levels of knowledge contributed the formation of driver trust in different ways.

Dependability was the most predictive indicator accounting for driver trust for both groups at every stage (H4). The replication study of Muir and Moray (1996) conducted in STUDY I showed the similar finding with statistical relationship between operators' trust and dependability. The different result between Muir and Moray (1996) and STUDY I may reflect effects of current social contexts. It can be interpreted that the changed social contexts influence the current study in spite of the different automation domain. Feelings for dependability dominantly governed their trust of drivers in the Detailed group from the practice and 2nd trial, then faith was added after the intervention task in the 3rd trial. Lastly, dependability and predictability dominated trust as the best predictors after experiencing the error-free automation in the 4th trial. For the drivers in the Less group, both dependability and faith were consistently best predictors of trust from the 1st to 4th trial. Whilst there was no difference in trust ratings between the Detailed and Less group, the best predictive features were changed after the exposure to the flawless automation between two groups. This also confirm that driver trust can be differently evolved by the levels of knowledge. Driver trust can be shaped based on dependability of automation, then predictability and faith interact depending on their knowledge levels and automation experience.

One interesting finding is an implication for the relationship between predictability and faith of automation. Predictability determined trust of drivers in the Detailed group after experiencing the error-free automation in the 4th trial, however, faith consistently governed trust of drivers in the Less group. The result described in Muir and Moray (1996), predictability could not be the best predictor of trust and showed the less correlation with operators' trust after the first training of a process control that operators should supervise raw milk pasteurization simulation. However, continuous exposure to the process control degraded the faith of pump system and led predictability to be the determinant of trust. That is, predictability may be associated with faith in an inverse way. Consistent with Muir and Moray (1996)'s finding, the present study observed that continuous human-machine interaction and experiencing automation failure led high levels of predictability and low levels of faith in the Detailed group. Additionally, our result for both the Detailed and Less groups did not partially follow Muir and Moray (1996)'s result as well as the original assumption suggested by Rempel et al. (1985). It also mirrors changes of social contexts between 1980s and 2020 in which automated machines are pervasive and familiar with for not only professional operators but also general users. For a better understanding of automation design, the relationship between predictability and faith could be explored with respect to not only different automation domain but also specific features of driving automation or traffic situations.

It is worth noting the changes of the best predictors between the 3rd and 4th trials for the drivers in the Detailed-Lim and Less-Mal groups. The automation failure changed the best predictor from both faith and dependability to dependability for the Detailed-Lim group as well as changed it from dependability to both faith and dependability for the Less-Mal group. This result may be interpreted with the result of response to the question of whether the driver actually trust the automation or not (Table 3.9). The respondent number for the Detailed-Lim group did not change, but there were two changes from trust to distrust for the Less-Mal group. System limitation can be considered as designed failure unlike system malfunction, also the drivers in the Detailed group knew specific information about the presence of automation failure unlike drivers in the Less group. For the Detailed-Lim group, faith functioned for preventing distrust from automation failure. The respondent number in the Less-Mal group steadily increased with trust which shaped by both dependability and faith, and faith could not be a determinant of trust when malfunction occurred. Even though the faith governed their trust after the error-free automation like previous stages, the respondent number was not recovered. These results indicate that faith is differently functioned for trust formation depending on type of automation failure and the extent of knowledge that they have. Muir (1994) described three components producing faith: dependability, predictability, operators' perception of the appropriateness and flexibility of the software which defines the goals and directs the behaviour of the machine subordinate, and this may account for the different functioning of faith toward trust formation. Our present result implies that faith incorporates dependability. The R^2 value of faith was greater than that of dependability for the Detailed-Lim after the 3rd trial, and the R^2 of dependability was greater than that of faith for the Less-Mal group after the 4th trial. However, regarding that the present result is partly limited to distinguish which elements incorporate faith and personality variables have a large impact on faith development (Muir, 1994), separate questions in terms of faith need to observe the evolution of faith in further study (e.g., Sanbonmatsu et al., 2018; Balfe et al., 2018). Additionally, in order to gain a comprehensive insight, this issue should extensively take into consideration other trust attributes, such as reliability or perceived risk (e.g., Wickens & Dixon, 2007; Beller et al., 2013; Chancey et al., 2017).

3.4.3 Demographic Factors

The results of statistical analyses revealed substantial differences in trust ratings across the three demographic factors, with supporting the H5 partly. This result indicates that driver gender and occupation influence driver trust in driving automation. As mentioned previously, recent studies found that higher trust in automated vehicles is observed in males than in females (e.g., Hulse et al., 2018; Lee et al., 2017); however, opposite results have also been found. In Feldhütter et al. (2016), driver trust varied with gender, and male drivers showed higher levels of trust of and intention to use highly automated vehicles than female drivers. Furthermore, females found highly automated driving more uncomfortable compared to males. However, our data revealed that females are more likely to trust driving automation than males. Moreover, the frequency of driving did not affect their trust. This result agrees with the results of Hulse et al. (2018), who found that drivers and non-drivers showed similar attitudes toward an autonomous car. Most non-daily drivers in our present study had not driven since they obtained their driver licenses. The effect of daily driving on driver trust may therefore be minimal. Future studies can expand on the sample sizes and the current experimental design used in the current study.

For the distrust attribute, the student drivers showed higher levels of distrust of the driving automation than non-student drivers. Furthermore, the analysis of trust ratings also revealed that non-student drivers were more likely to give higher trust ratings than student drivers, which does not agree with the result that student drivers are likely to trust and have an affinity for vehicle technology (Chien et al., 2014). Interestingly, although most non-daily drivers were student drivers, driving-related differences in both trust and distrust were not observed. The result in terms of the occupation may differ according to drivers' affiliation. The students at the University of Tsukuba perhaps strictly rated their trust compared to the non-student drivers. If more students from the engineering departments participated in the next study, the data may reveal higher levels of trust in automated cars for the student drivers than for the non-student drivers. For example, an engineering student driver reported that his previous work on developing sensors led to a distrust of automation; however, a non-student driver who has used advanced driver assistance systems for many years expected rapid commercialization of automated vehicles. A future study should consider drivers with a wide range of occupations. Additionally, detailed participant interviews need to clarify why drivers trust or distrust automation in the next study.

Further, this study addressed what automation designer should focus on information influencing driver trust with respect to general trust bases of Lee and Moray (1996): purpose, process, and performance. As shown in Figures. 3.6, 3.7, and 3.8, trust across the three dimensions was moderate, with most drivers rating trust with 4 or 5 on the seven-point scale ("Neither agree nor disagree" and "Agree somewhat"). Therefore, this result implies that drivers' attitudes toward driving automation are slightly favourable. The ratings of trust were collected after all trials had been completed in the driving simulator. All the drivers completed four trials with the simulated driving automation, and in one of the trials, the drivers were asked to intervene by the automation. Because the driver was informed about the failure of the automated vehicle, it may be considered that the driver's trust rating was not quite high. In particular, according to the responses given by the drivers in the short interview conducted after completion of the questionnaire, several drivers were concerned about driving safety when unexpected system failures occurred in the urban road because the driving scenario in the present study is less hazardous than the real-life driving situation on an urban road with a high car density. This suggests that automation designers should be concerned about how to ensure driving safety for drivers.

This study hypothesized that the dimension of process liken dependability may show the highest ratings among the three-dimension (H6), however, differences in the levels of trust between each dimension were found in this study. The driver ratings for the purpose dimension were significantly lower than those of the process and performance dimensions across all demographic factors. Information about the three dimensions could be used to determine appropriate levels of trust. This result may be interpreted to mean that information relevant to what the automation is supposed to do needs to be provided to drivers in comparison with information relevant to the actual performance of automation and how the automation works. The finding suggests that recommendations for automation designers as well as is helpful for maintaining the vigilance of drivers who already purchased contemporary automated vehicles. However, this result should be interpreted cautiously because the number of items varied across each dimension. Therefore, each questionnaire item should be considered to discuss the trust dimension with the basis of trust, such as reliability or understanding. In line with the discussion about the occupation result, further studies require specific questionnaire items to clarify the dimension that most affects driver trust.

3.4.4 Driver Willingness to Trust Vehicle Automation

Many studies on trust have been conducted without discussion of drivers' actual trust or distrust of automated vehicles. The question was presented to the drivers during this experiment, and the change in response number over time is quite similar to the result of the trust ratings. Even with different instructions, the same number of drivers in both groups answered 'Yes' to the question after reading the instructions. Then, the number steadily increased with the automation experience in the practice drive, and subsequent human-machine interactions changed 19 drivers' minds from distrust to trust, as measured by the last questionnaire. This corresponds to findings that driver trust increases by experiencing automation (Hergeth et al., 2017). Even though there was no clear distinction by the knowledge level until the 2nd trial, the changes of response number confirmed that both knowledge levels and automation failure type influence driver trust in different ways as aforementioned. Chi-square tests (see Table 3.10) support this tendency that the participant developed trust in the automation with the experience. Even though the data did not show a clear distinction by knowledge level, the changes in the response number reflected that both the knowledge level and automation failure type influence driver trust in different ways. This result observed a clear finding that continuous exposure to automation can lead driver trust and use of automation, and potential for developing training methods for drivers.

3.4.5 Limitations

The current study should be interpreted with several limitations. First, this driving simulator study included an intervention task. Parasuraman and Manzey (2010) described that the repeated exposures to the automation imperfection led calibrations of operators' trust reliance to the system. Future study needs the greater number of automation failure for more accurate trust calibration. Second, in line with the first limitation, only highway environment was presented to the drivers. As driving automation is expected to be used for various traffic situations, future study should be designed with respect to road environments (Deb et al., 2017) and features of driving automation (Abraham et al., 2017). Lastly, the findings described in this paper may be not necessarily generalized without considerations about individual differences. As shown in Table 3.9, the participants' responses were considerably varied even though the completely identical description was provided. It is possible that previous experience and knowledge of driving automation affected their answer to the question about whether they actually trusted and wanted to use the automation or not. In particular, several literatures have described that operators' trust are highly dependent on their individual differences (e.g., Wintersberger et al., 2017). Therefore, accurate trust calibration with respect to this issue should be highlighted in next study.

However, the current study discussed about the dynamics of trust with empirical data analysis, and the result suggested how three dimensions of trust interact each other and what should be concerned according to the levels of driver knowledge in terms of driving automation. The findings described in this paper will crucial to have a wider perspective in automation design with respect to driver trust and current social context. Further, the current suggestion in terms of key dimension determining driver trust could be applied to the driver mental model development or training design.

Chapter 4. REVISITING MUIR AND MORAY (1996) WITH STUDENTS IN ENGINEERING MAJOR

4.1 Purpose of Replication

STUDY I replicated the control process experiment of Muir and Moray (1996) to observe evidence for the model of human-machine trust which operator trust developed from faith, initially the best predictor, then dependability, and finally predictability, the opposite order of the predicted trust development. Instead, dependability predicted overall trust towards the automatic pump system the most convincingly among the three proposed constructs. STUDY II applied the model of Muir and Moray (1996) to the context of partial vehicle automation regarding relatively familiar automation for people in general populations and obtained identical results with the STUDY I. In spite of different levels of knowledge in terms of system failure in driving automation and different types of system failure, the attribution of dependability was the largest among the three dimensions of trust for the development of driver trust. STUDY II has implications for effects of operators' gender and occupation on trust in automation. As aforementioned in the Chapter 4, because all student participants in STUDY II were from the University of Tsukuba, students are likely to adopt new technologies considering the characteristics of university and areas. Here, one may suspect that the results occurred because the current participants were different from the original participants in terms of their expertise. In Muir and Moray's (1996) original study, graduate students were experienced in process control or thermodynamics (Muir's Experiment 1) or Engineering graduate students (Muir's Experiment 2; Muir, 1989) were recruited for the experiment. It is possible that operators' different educational background, affinity for technology, contributes the formation of trust.

STUDY I remains the possibility that first impression may strongly relate to trust formation and the use of automatic pump mode in the course of interaction with machines. Most of participants reported concerns about the complexity of plant operation after the first experience of the plant during the tutorial. The participants in the STUDY I spent more time in automatic controllers compared to those of the original study. This issue may be tangled with operators' self-confidence in system operation. STUDY II which collected drivers' propensity to trust and willingness to use vehicle automation observed that increases in the number of drivers who are prone to trust and use the automation in the course of experiment. The result indicated that experience with automation leads users to find benefits of automation. However, the numbers regarding propensity to use automation was relatively lower than those of trust. That is, other factors, not trust, may make a large attribution in terms of operators' decision making to rely on automation.

Operators' willingness to rely on automatic aids is determined by several factors, and one of them is self-confidence in system operation (Lee & Moray, 1992; Lee & Moray, 1994; Riley, 1996; Kantowitz et al., 1997; De Vries et al., 2003). In general, low self-confidence leads operators to rely on the performance of automatic controller, on the contrary, high self-confidence leads increased manual controls resulting in high workload. The combination of low self-confidence and low trust also yields increased workload (Lee et al., 1999). Self-confidence can be changed by the occurrence of system fault and the type or magnitude of system fault like trust, and it affects operators' decision making about the allocation between manual control and using automatic mode in supervisory control situations (Lee & Moray, 1994). Given that reliability is a basis of trust (e.g., Chancey et al., 2017; Sato et al., 2019), self-confidence influencing automation usage may be formed by the first impression on automated systems.

Based on previous literatures and findings of STUDY I, it can be assumed that the first impression with reference to the pasteurizer degraded participants' self-confidence and led frequent use of the automatic pump in the STUDY I. It needs to clarify the reason of significantly high proportion of automatic controllers in comparison with the study of Muir and Moray (1996). self-confidence can be a key to identify the reason why operators in STUDY I frequently used automated pump modes compared to those of the original study. Additional replication is expected to identify the reason why willingness to rely on automation regarding several choices, historical background, gender (dispositional quality), educational background (may closely correspond to pre-existing knowledge), and self-confidence (the first impression on machines).

STUDY III repeated Muir and Moray's (1996) process control tasks to address issues about trust development with only undergraduate students majoring in Engineering to more closely replicate Muir and Moray's (1996) original study, demographic information, and reliance on automation regarding self-confidence based on findings of STUDY I and II. Accordingly, the objective of investigating self-confidence is to look into not only the relationship between trust in automation and self-confidence but also how operator's first impression functions trust formation and automation usage. Following research hypotheses were established:

- H1: Automation failure in the display of machine negatively impacts operators' trust.
- H2: The affinity for technology from engineering backgrounds leads less automation usage compared to results in STUDY I. However, current social background leads relatively high propensity to adopt automatic aids compared to those of the original study.
- H3: Responsibility can account for the meaning of trust in machines rather than competence.
- H4: Trust in machines is shaped by the feeling of dependability because of social background regardless of the engineering background.
- H5: The first impression on machines forms low self-confidence in the plant operation, resulting in the propensity to rely on automatic controllers.
- H6: There is difference in trust ratings between female and male students.

4.2 Methodology

This research complied with the University of Tsukuba's ethics code and was approved by the ethical review board at the University of Tsukuba.

4.2.1 Participants and Apparatus

Twelve undergraduate students (6 females; M_{age} = 21.15 years, SD_{age} = 1.08 years) who major in Engineering at University of Tsukuba, Japan were recruited and participated in the study. None had participated in STUDY I. They were paid 2,460 JPY for each session.

4.2.2 Apparatus, Pasteurizer Plant, and Experimental Design

Apparatus, pasteurizer plant, experimental task, and experimental design were identical to those of STUDY I.

4.2.3 Procedure

Participants first completed an informed consent and a demographics form before beginning the study. Participants were then trained to become proficient in using the milk pasteurization system. The experimenter trained participants by introducing them to the pasteurization system with a manual, practicing manual pump control, practicing switching between modes, practicing automatic control mode. This replicates the training procedure described by Muir and Moray (1996). Unlike Muir and Moray (1996) that participants should engage with the pasteurizer at least 7 sessions and lasted until they achieved mean performance score 80%, free-choice sessions lasted until mean score reached 80% or they completed seven training during the training program after the fourth step (mean number of training sessions = 4.5, *SD* = 1.24).

After reaching the performance threshold, participants interacted with nine pasteurization systems in a randomized order as dictated by the experimental conditions (Table 2.1). The participants had no information about the system's accuracy. Each participant completed one simulation run with 80 iterations for each of the nine experimental conditions. After each run, participants completed the subjective trust questionnaire which used in STUDY I.

4.2.4 Dependent Variables

Data on task performance were collected, including performance score, pump control actions, and proportion of time in automatic pump mode, to ensure participants were engaged in the task. In addition to the task performance measures, participants also completed a subjective trust questionnaire assessing the constructs outlined in Table 2.2.

STUDY I remains several limitations. For example, it is unclear whether the participant actually could be aware of error in the pump system as well as recognize the type of error (constant or variable error, and display or control property error). Therefore, additional questions in terms of their impression, impression, and whether they were aware of error or not were presented to participants in STUDY III. First, after the fourth step, where the participant finished the first operation with the automatic controller, the first impression about the pasteurizer was collected with the trust questionnaire (Table 2.2) except for a participant. This data collection was used to understand how their first impression has impacts on shaping trust and relates to the dimension of trust. For instance, if they regarded the pump system as the automatic controller or the pasteurizer itself, the experiment asked them to consider the pump system as the pump system itself from now. Questions in terms of self-confidence and understanding of two subsystems were adapted from Lee and Moray (1994), “How high was your self-confidence in controlling the pump system/active heating system?” and “How much did you understand the mechanism of the pump system/active heating system?” Whether the operator actually trust the pump or not and which control type between manual and automatic control the operator prefers were checked. Further, they reported if they really understand the mechanism of the pump and active heating system. From the experimental session, the experiment asked them to report whether there is error in the pump or not, and if there is the error, describe error type and phenomena that they observed in this plant. Lastly, after the last operation, their overall impression of the experiment was collected.

4.2.5 Statistical Analyses

Identical statistical analysis methods to Muir and Moray (1996) and STUDY I were applied to assess data collection for STUDY III. To strengthen statistical results, Bayes factors (*BF*) and variance inflation factors (VIF) were also calculated. As STUDY III did not consider run factors within experimental sessions, there were two experimental factors: control properties of the pump (exact, constant error, variable error) and display properties of the pump (honest, constant error, variable error). To measure operators’ performance on pasteurization plants, this study used three performance measures: scores, the number of pump control actions, and the proportion of time in automatic pump mode. The two experimental factors were also analysed on these performance measures. Two-way (3 x 3) repeated-measures ANOVAs were conducted to analyse subjective ratings of trust in the pump system (overall trust, trust in the pump, and trust in the pump’s display) and the three performance measures across the three experimental factors. For all post-hoc tests, this study utilized the Bonferroni correction against alpha inflation due to multiple comparisons. This study assessed correlations between overall trust and three performance measures (performance score, manual control pump actions, proportion of time in automatic pump mode). In this study, independent samples t-test compared performance variables of STUDY III with not only the results of Muir and Moray (1996) but also STUDY I. To examine the meaning of trust in machines, simple linear regression and stepwise regression were conducted. To look into differences between operators who controlled the plant with only automatic pump modes and with only manual operation, independent-samples t-test and linear regression were performed. Further, stepwise regression was used to assess hypothesized models in terms of trust development. Two participants failed to keep the main vat volume, resulting in system crashes. Thus, the data of failed participants in the experimental program were eliminated. All statistical analyses were conducted in R (R Core Team, 2020).

4.3 Results

The statistical analysis results partly refer the work of Long et al., (under review). Table 4.1 describes the average of subjective ratings of trust and performance with standard deviation.

Table 4.1 Average of subjective ratings of trust and performance with standard deviation

	Subjective Ratings of Trust			Performance		
Experimental Condition	Overall Trust	Trust in Control	Trust in Display	Score	Manual Control Action	Proportion of Time in Automatic Pump Mode (%)
C1	86.93	85.20	76.87	90.93	19.10	30.00
	(9.88)	(12.49)	(24.16)	(5.28)	(29.40)	(48.30)
C2	79.40	81.13	75.40	88.24	37.20	26.54
	(16.01)	(11.90)	(21.83)	(8.75)	(38.05)	(43.76)
C3	77.74	71.00	61.13	89.73	44.10	25.80
	(16.54)	(17.40)	(32.68)	(9.11)	(62.18)	(43.09)
C4	78.47	78.14	68.33	85.58	42.50	30.96
	(13.24)	(14.29)	(29.10)	(11.96)	(84.14)	(46.93)
C5	77.40	74.53	68.47	84.98	29.70	24.07
	(18.31)	(15.82)	(23.03)	(6.48)	(33.31)	(41.98)
C6	71.13	66.73	56.60	80.17	31.00	26.67
	(17.28)	(18.16)	(24.02)	(13.14)	(38.55)	(43.89)
C7	85.47	79.07	79.07	90.79	20.50	29.56
	(12.09)	(16.42)	(19.22)	(7.66)	(22.63)	(47.61)
C8	83.20	80.80	75.93	88.24	30.50	23.07
	(14.96)	(14.03)	(21.69)	(5.81)	(50.07)	(41.66)
C9	75.33	69.40	61.20	88.42	49.40	25.06
	(21.01)	(22.42)	(32.97)	(3.81)	(64.40)	(40.98)
Overall	79.45	76.22	69.22	87.45	33.78	26.86
	(15.80)	(16.50)	(25.82)	(8.74)	(49.24)	(42.36)

4.3.1 Overall Trust in Pump System

Figure 4.1 illustrates mean ratings of trust of the pump system. Data provided substantial evidence for the main effect of display properties, $F(2, 18) = 4.22$, $p = .04$, $\eta_G^2 = .06$, $BF_{10} = 2.3$. Overall trust tended to be significantly lower when the automation made variable errors (M = 84.19) than no errors (M = 88.33). However, remaining effect was not significant, $F(2, 18) = 2.28$, $BF_{10} = .51$, and the interaction effect, $F < 1$, $BF_{10} = .15$.

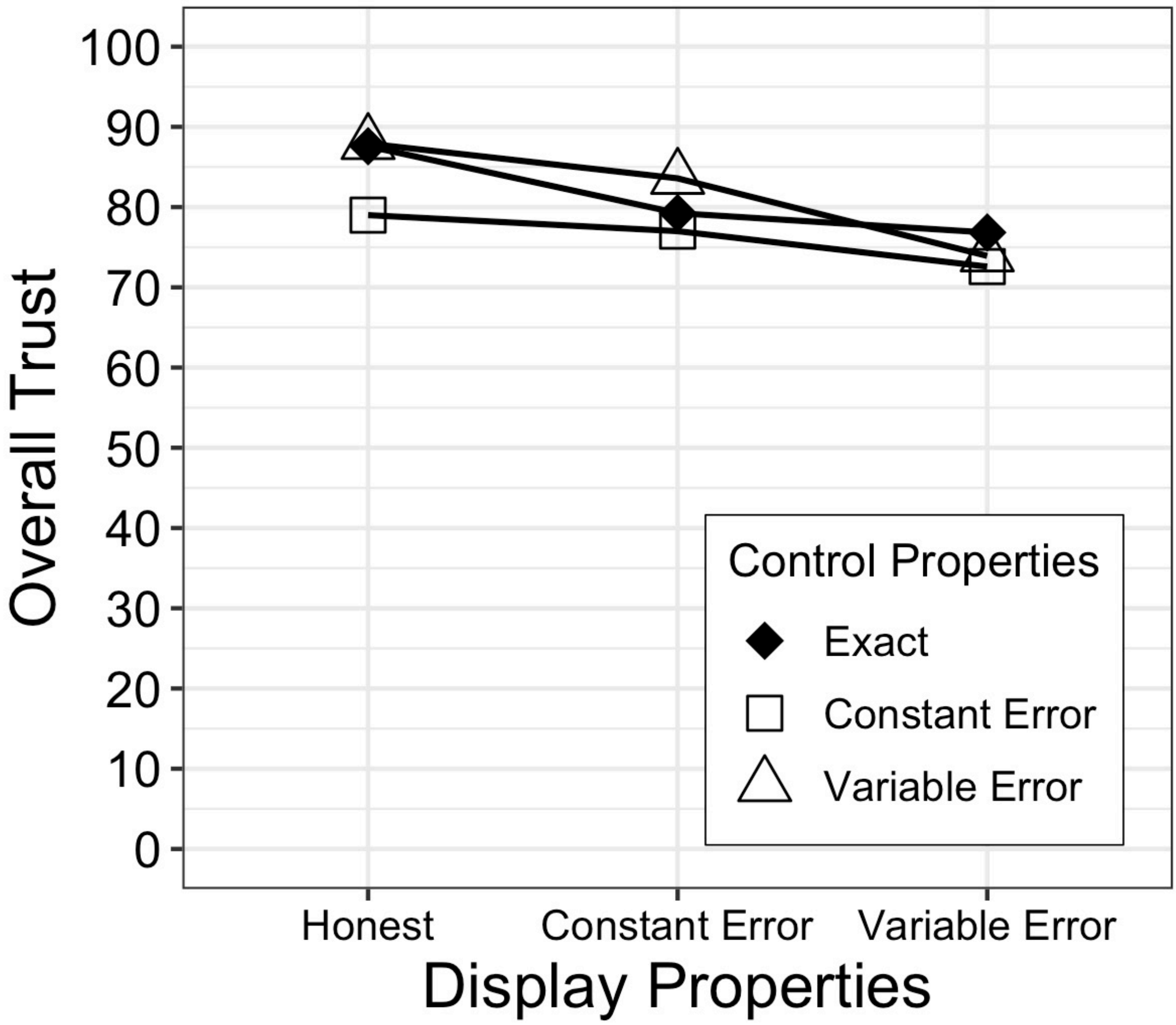


Figure 4.1 Mean ratings of overall trust in pumps as a function of control and display properties of the pump

4.3.2 Trust in Pump Control and Display Properties

Figure 4.2 illustrates mean ratings of trust of the pump control. ANOVA found the main effect of display properties, $F(2, 18) = 6.18$, $p < .01$, $\eta_G^2 = .1$, $BF_{10} = 31.23$. Operators' trust in pump control was significantly higher when they interact with honestly displaying pump ($M = 80.8$) and pump with constant display error ($M = 78.82$) than when interacting with the pump including variable display error ($M = 69.01$). However, remaining effect was not significant, $F(2, 18) = 1.75$, $BF_{10} = .31$, and the interaction effect, $F < 1$, $BF_{10} = 1.32$, all ps > .2.

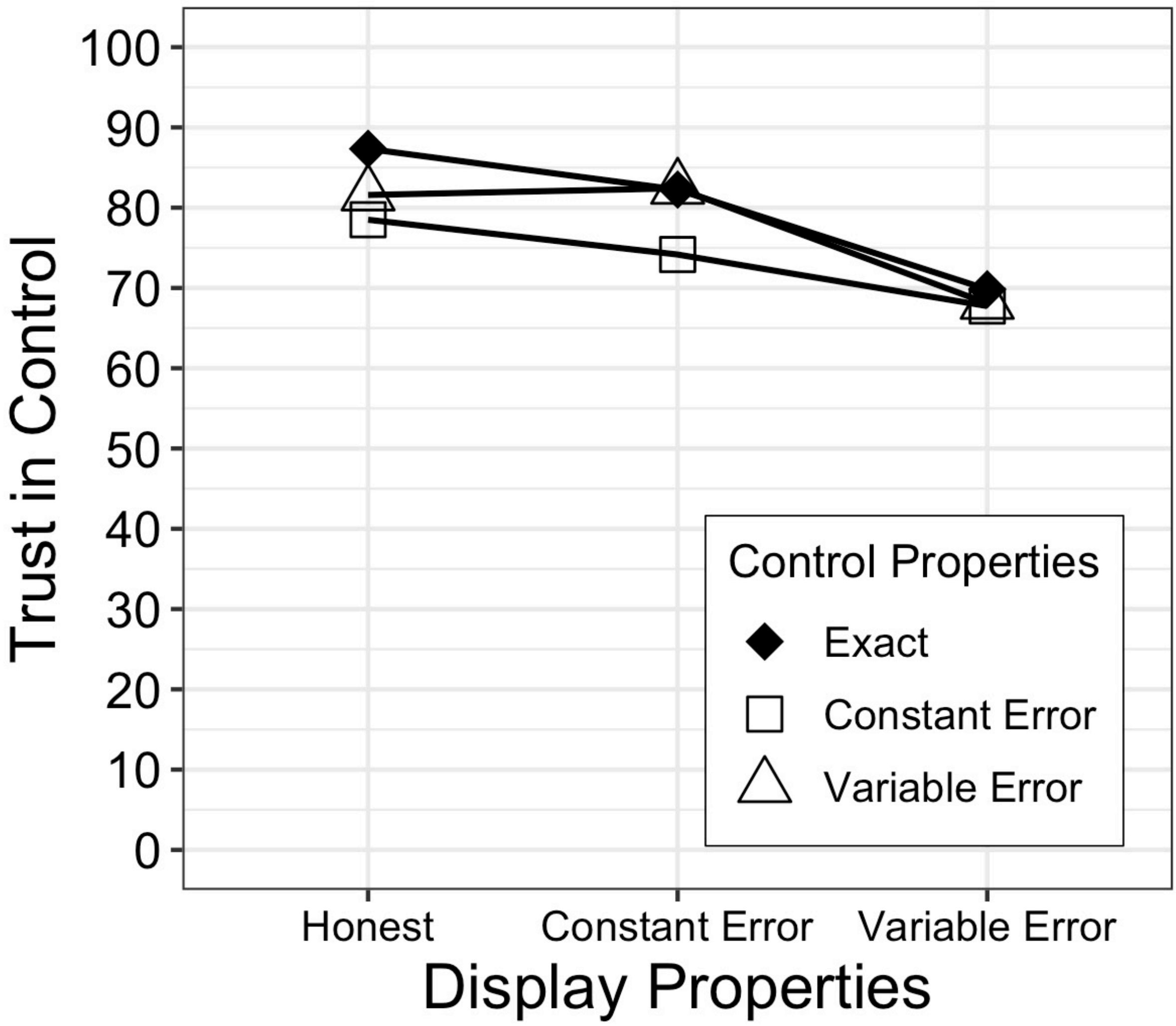


Figure 4.2 Mean ratings of trust in pump control as a function of control and display properties of the pump

Figure 4.3 shows mean ratings of trust of the pump display. Data provided substantial evidence for the main effect of control properties, $F(2, 18) = 5.1$, $p = .01$, $\eta_G^2 = .01$, $BF_{10} = 0.34$. Trust in the pump display tended to be significantly lower when the automation made constant control errors ($M = 64.47$) than no errors ($M = 71.13$) and variable control errors ($M = 72.07$). The main effect of display properties was also found, $F(2, 18) = 4.18$, $p = .03$, $\eta_G^2 = .07$, $BF_{10} = 23.47$. Trust in the pump display tended to be significantly lower when the automation made variable display errors ($M = 59.64$) than no errors ($M = 74.75$) and constant control errors ($M = 73.27$). However, the interaction effect was not observed, $F < 1$, $BF_{10} = .89$.

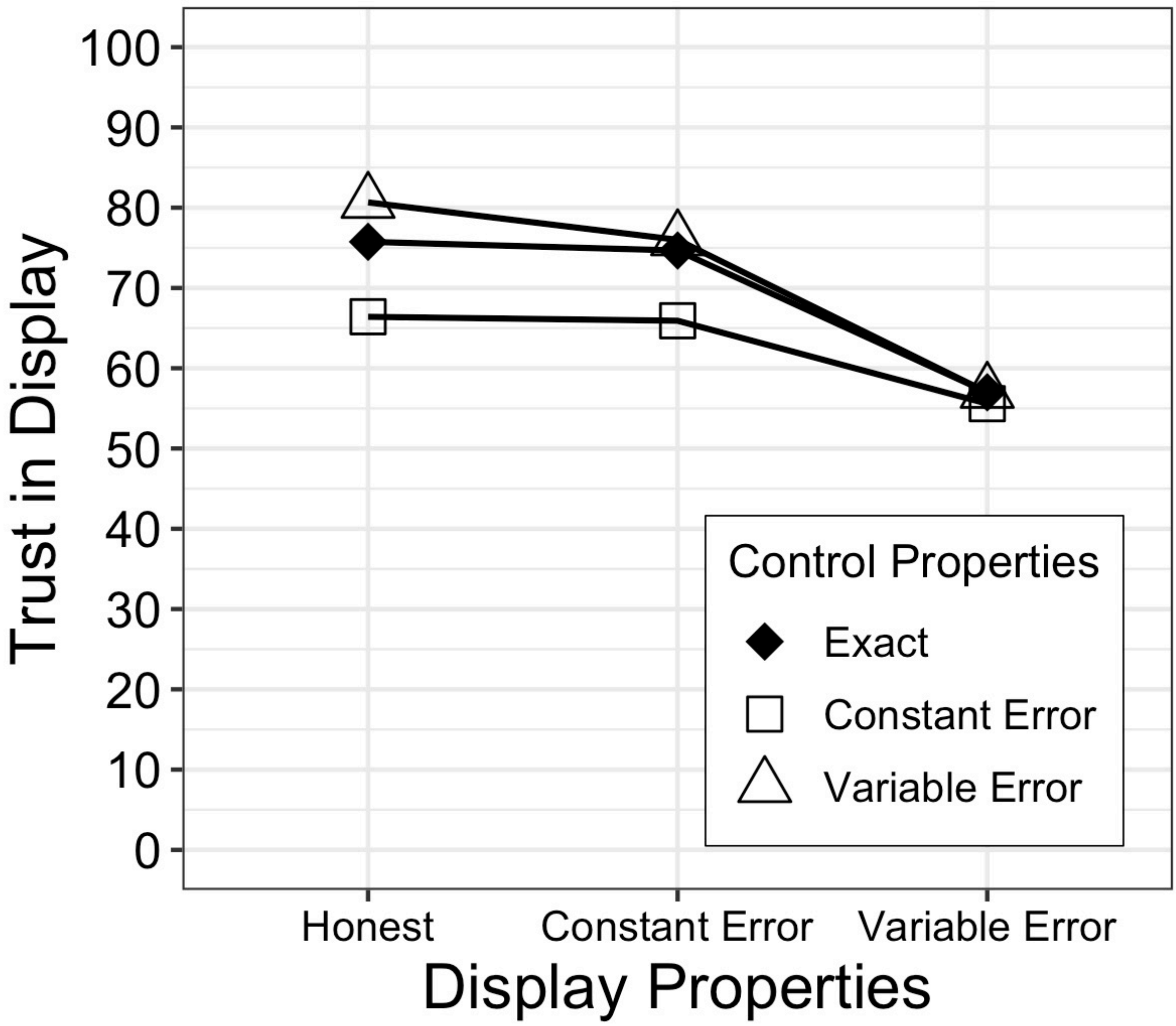


Figure 4.3 Mean ratings of trust in the pump display as a function of control and display properties of the pump

4.3.3 Self-Confidence

ANOVA did not show the main effect of both control and display properties on self-confidence in pump and active heating system operations, all Fs < 2.61, ps > .1, BF_{10} < .22, and the significant interaction effect on self-confidence in operation of both systems, all Fs < .31, ps > .87, BF_{10} < .01.

4.3.4 Performance

Figure 4.1 presents the mean performance score, manual pump control actions, and proportion of time in automatic pump mode of Muir and Moray (1996), STUDY I and this study (STUDY III) separately.

Performance score. Data provided substantial evidence for the main effect of control properties, $F(2, 18) = 5.55$, $p = .01$, $\eta_G^2 = .1$, $BF_{10} = 6.75$. Performance score tended to be significantly lower when the automation made constant control errors ($M = 83.58$) than no errors ($M = 89.63$) and variable control errors ($M = 89.15$). The main effect toward display properties and interaction effect were not found, all Fs < 1, ps > 0.42, and BFs = .22.

Independent-samples t-test revealed that scores of STUDY III were significantly lower than those of Muir and Moray (1996), $t(16) = 8.37$, $p < .001$, $BF_{10} = 3.01 \times 10^4$ and of STUDY I, $t(16) = 4.78$, $p < .001$, $BF_{10} = 106.53$. These imply that participants in STUDY III could be not able to figure out how to achieve high performance score like participants in both the original study (Muir and Moray, 1996) and STUDY I.

Manual pump control actions. Data gave substantial evidence against the main effect of both control and display properties, all Fs < 1.3, ps > .32, BF_{10} < .32, and the interaction effect, $F(4, 36) = 1.06$, $p = .39$, $BF_{10} = .01$.

Independent-samples t-test also revealed that manual control numbers of STUDY III did not significantly differ from those of two previous studies, Muir and Moray (1996), $t(16) = -1.03$, $p = .31$, $BF_{10} = .59$ and STUDY I, $t(16) = -1.58$, $p = .14$, $BF_{10} = .93$.

Proportion of time in automatic pump mode. Data gave substantial evidence against the main effect of both control properties and the interaction effect, all Fs < 1, ps > .79, BF_{10} < .11. A favourable trend of the main effect for display properties was found, $F(1, 18) = 2.73$, $p = .09$, $BF_{10} = .65$, but BF did not support the trend.

Independent-samples t-test revealed that the proportion of STUDY III were significantly higher than those of Muir and Moray (1996), $t(16) = -12.73$, $p < .001$, $BF_{10} = 6.19 \times 10^6$ and significantly lower than those of STUDY I, $t(16) = 31.98$, $p < .001$, $BF_{10} = 3.18 \times 10^{12}$.

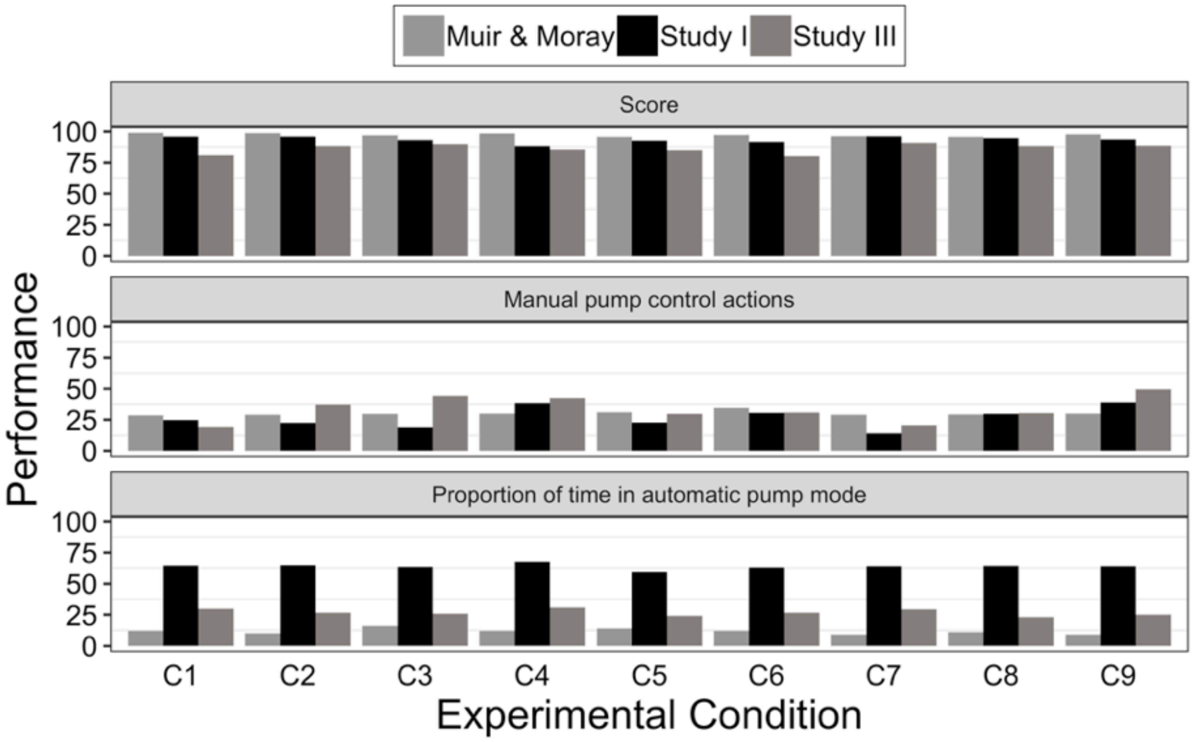


Figure 4.4 Mean performance score, manual pump control actions, and proportion of time in automatic pump mode of Muir and Moray (1996), STUDY I and STUDY III

4.3.5 Correlation between Performance and Trust

Figure 4.5 illustrates charts of bivariate correlation between overall trust and performance variables. The correlation between overall trust in the pumps and performance scores was statistically significant, $r(88) = .36$, $p < .01$. The correlation between overall trust and the number of manual pump operation was not significant, $r(88) = -0.08$, $p = .44$, as well as use of the automatic controller, $r(88) = .09$, $p = .42$.

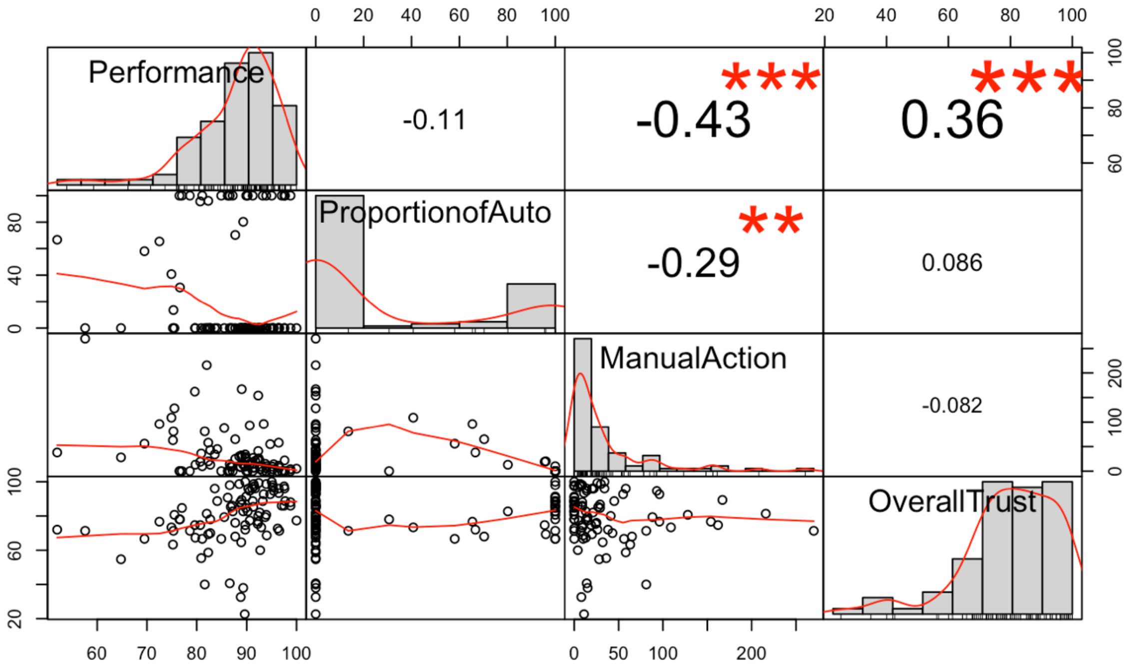


Figure 4.5 Chart of correlation between ratings of overall trust in the pump and three performance measures.

4.3.6 The Meaning of Trust in Machines

Stepwise regression analyses were performed to examine whether competence and/or responsibility predict overall trust in machines. Muir and Moray (1996) found that competence is a better predictor of trust accounting for the meaning of trust than responsibility. However, data from STUDY I showed that responsibility was the most predictive in all the experimental conditions except for the C6 condition (constant error-variable error) as shown in Table 2.4. The results from data of STUDY III were summarized in Table 4.2. In the C2 (exact control-constant display error) condition, ANOVA result was not observed. All models did not include the multicollinearity problem.

Table 4.2 Summary of analyses of the meaning of trust. *C* = competence, *Rs* = responsibility, *T* = overall trust.

Experimental Condition	R^2 C	BF_{10} C	R^2 Rs	BF_{10} Rs	R^2 C, Rs	Stepwise Best Predictor	F (Best Model)	p
Total	0.21	5.74 x 10^{12}	0.67	4.48 x 10^{17}	0.68	*Rs, C*	82.31	< .001
C1	-0.54	0.75	0.84	5.38	0.54	*Rs*	11.47	< .001
C2	0.07	0.60	0.48	0.53	0.34	*C, Rs*	3.28	.1
C3	0.26	1.54	0.97	11.45	0.64	*Rs*	1.54	< .001
C4	0.23	12.45	0.91	2199.07	0.93	*Rs*	118.64	< .001
C5	0.18	16.14	0.92	44.40	0.77	*Rs*	31.00	< .001
C6	0.30	3.34	0.86	4.14	0.49	C	9.77	< .001
C7	0.22	25.36	0.64	483.35	0.93	*Rs*	63.12	< .001
C8	0.78	239.99	0.29	60.94	0.86	*Rs*	57.56	< .001
C9	0.06	7.56	0.93	57.73	0.79	*Rs*	34.35	< .001

4.3.7 The Development of Trust in Machines

Table 4.3 shows the results of a stepwise regression entering 1) Faith, 2) Dependability, and 3) Predictability. This analysis showed that dependability was the best predictor overall, with only dependability the first training and experimental conditions. Faith predicted trust in the last training condition. Bayesian regression analyses corroborated the findings using the stepwise regression, and concorded with the findings of STUDY I, that dependability was the best predictor after the first interaction with the automatic controller (faith, BF_{10} = 4.91; dependability, BF_{10} = 64.79; predictability, BF_{10} = 11.38), the first training (faith, BF_{10} = 2.07; dependability, BF_{10} = 65.17; predictability, BF_{10} = 1.17), and in the experimental condition (faith, BF_{10} = 1.44; dependability, BF_{10} = 272.60; predictability, BF_{10} = 8.60). Unlike in STUDY I, though, dependability least predicted overall trust among the three predictors (faith, BF_{10} = 50.08; dependability, BF_{10} = 8.31; predictability, BF_{10} = 44.57).

Table 4.3 Summary of analyses of the development of trust.

Session	R^2 P	R^2 D	R^2 F	Stepwise Best Predictors	F (Best Model)	p
Automatic Pump	0.31	0.86	0.16	D	28.13	< .001
First Training	0.14	0.71	0.22	D	28.19	< .001
Last Training	0.46	0.24	0.72	F	25.80	< .001
Plant (C1)	-0.092	0.81	0.13	D	43.85	< .001

Figure 4.6 illustrates bivariate correlation between three dimensions of trust in automation at each point. VIFs at all points were not larger than 10. That is, all regression models did not entail the multicollinearity problem.

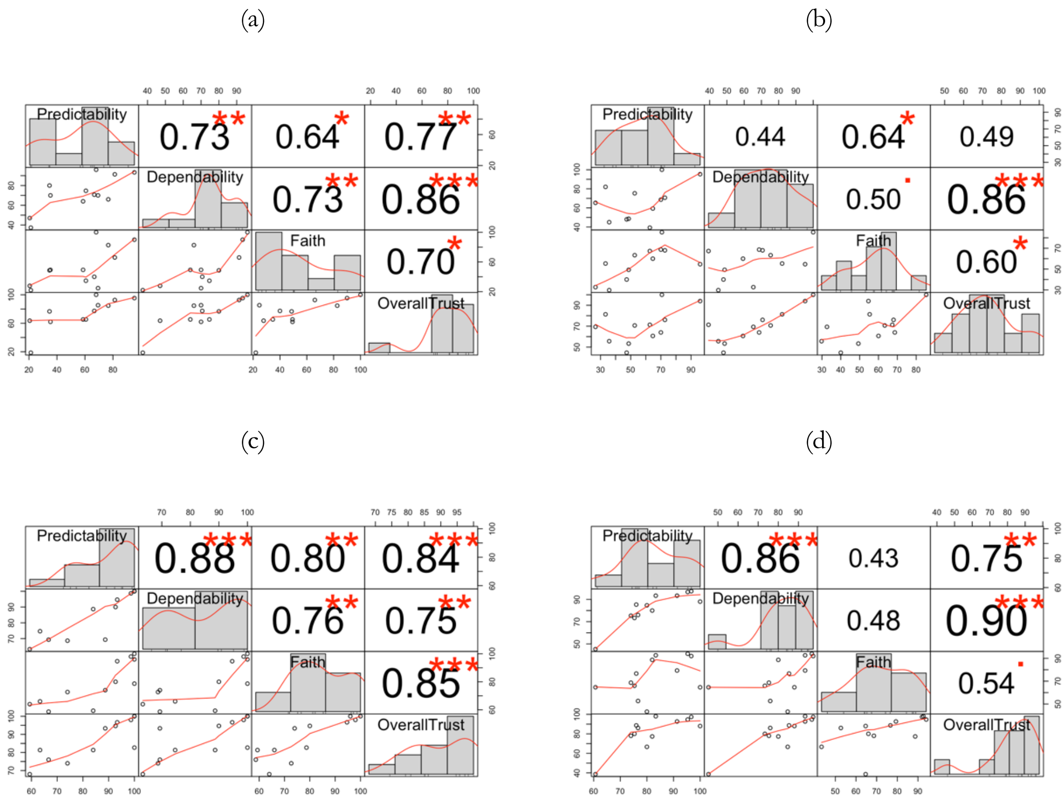


Figure 4.6 Bivariate correlation between all dimension of trust at all time points: (a) first interaction with automatic pump mode, (b) first training, (c) last training, and (d) experimental session of C1 plant.

Note. *** denotes $p < .001$, ** denotes $p < .01$, and * denotes $p < .05$.

4.3.8 Operators' Trust When Understanding Mechanisms of Pump System and Active Heating System

Inconsistent with the finding of STUDY I, which dependability consistently dominates operators' trust in automation, faith became a dominant contributor of human-machine trust after when the operators finished the last training. Considering this result, additional regression analyses were performed. As aforementioned, the operators were asked to check when they understand the mechanism of two subsystems. The points differed depending on participant, and the participants did not answer if they did not feel the lack of understanding in terms of the mechanisms. Stepwise regression showed dependability is the best predictor of trust when operators understood the mechanism of the pump system (faith, BF_{10} = 2.47; dependability, BF_{10} = 899.62; predictability, BF_{10} = 190.64), and faith best accounts for operators' trust when understanding the active heater's mechanism (faith, BF_{10} = 104.89; dependability, BF_{10} = 5.31; predictability, BF_{10} = 2.47).

4.3.9 Qualitative Analyses of Self-Confidence in System Operation and The Use of Automation

Self-confidence in operating the pump system and active heating system was rated by each participant. After the first interaction with automation, 9 of 12 participants (5 females, 4 males) reported that "I have no confidence in understanding this pump system completely." There was positively moderate correlation between overall trust and self-confidence in pump operation during training (r = .63) and experimental session (r = .59). Figure 4.7 presents changes of subjective ratings of overall trust (black) and self-confidence in pump control (blue) during the training session. Data was collected from the first interaction with automatic controllers except for a female student. Her self-confidence data were collected from her 10th run, but her first impression about difficulty in plant operation was collected. As shown in Figure 4.7, operators' trust and self-confidence in system operation develop in similar ways throughout experiencing the pasteurizer.

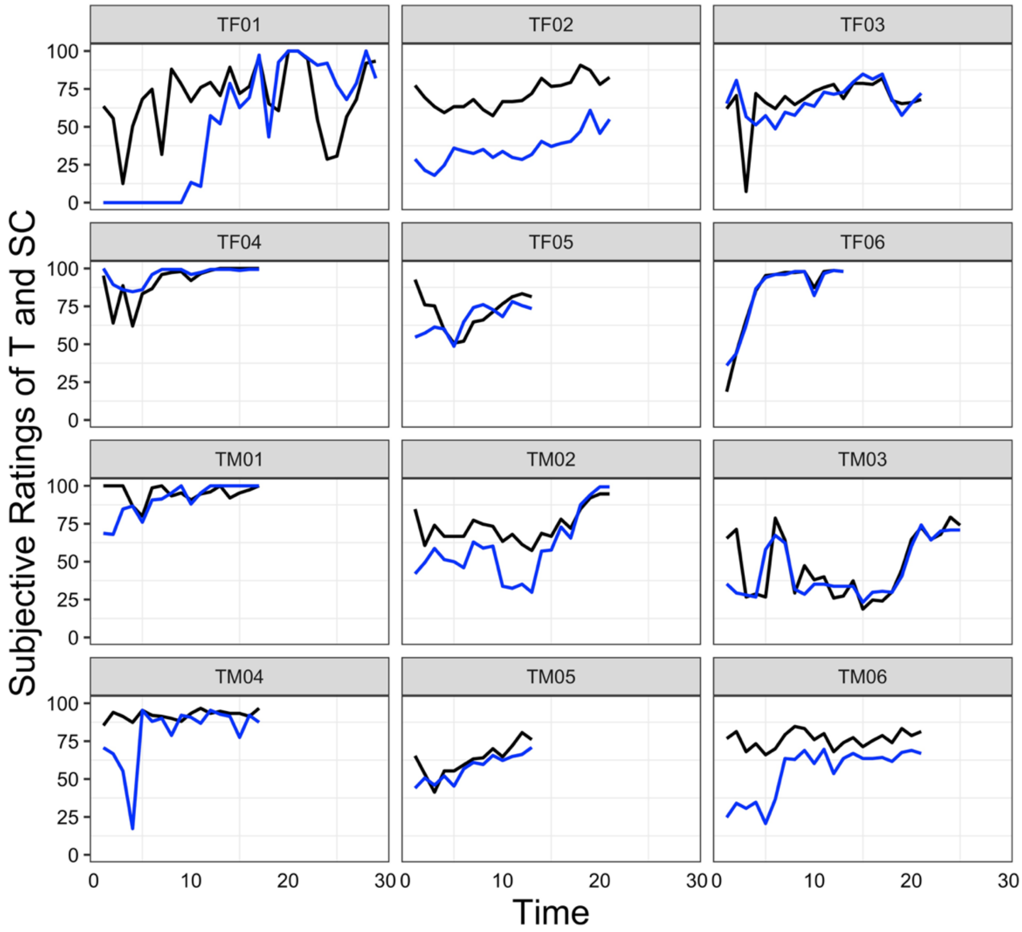


Figure 4.7 Subjective ratings of overall trust (black) and self-confidence (blue) for each participant.

According to the difference between subjective ratings of overall trust and self-confidence in pump control, three types were classified. Analyses in terms of self-confidence in operating the active heating system were also carried out in the identical way to self-confidence in operating the pump system. In this analysis, because two participants made failures during the experimental program, their data were disregarded from this analysis. That is, there were 90 cases for the analysis (9 experimental conditions x 10 participants).

- If the difference is positive, it indicates that the operator is likely to prefer taking the automatic controller to manual control.
- If the difference is negative, it indicates that the operator is likely to prefer manual control to taking the automatic controller.

Table 4.4 presents the mean values of overall trust, performance score, the proportion of time in automatic mode, and manual control action frequency, and the number of operators according to the classification. As shown in Table 4.4, the number of participants who rated overall trust lesser than self-confidence (the difference is negative) was larger than who rated overall trust greater than self-confidence (the difference is positive). Operators' trust, score, and proportion of time in automatic mode were higher when the difference is positive than negative. In addition, a clear evidence was observed that when the self-confidence is larger than overall trust, the operator prefers taking manual control to adopting the automatic controller. There are 8 cases that the difference is zero, and all of them in this case responded "Yes" to the question whether they trust the pump system or not with showing high levels of trust and self-confidence over 90.

Table 4.4 Summary of results from questionnaire according to the classification of self-confidence in pump operation

	N	Overall Trust	Preference in the use of automatic controller	Trust the pump	Score	Proportion of time in automatic mode	Manual control actions
Positive	33	82.73	12	32	88.14	31.47	19.39
Negative	49	74.72	13	35	86.45	25.27	45.41
Equal	8	94.92	2	8	90.79	17.59	21.88

Table 4.4 presents the mean values of overall trust, performance score, the proportion of time in automatic mode, and manual control action frequency, and the number of operators according to the classification. Here, similar trends with self-confidence in the pump operation were observed. Operators who rated overall trust more than or equally self-confidence answered, "I do trust the pump system." One interesting point is difference of manual control actions between participants who gave equal points to overall trust, self-confidence in operating the pump and active heater. Whilst the number of manual controls was 21.88 with regard to the pump system operation, the number was 10.17 with respect to the active heater operation. Overall, there results imply that overall trust and self-confidence in the plant operation relates to the use of automatic controllers.

Table 4.5 Summary of results from questionnaire according to the classification of self-confidence in the active heater operation

	N	Overall Trust	Preference in the use of automatic controller	Trust the pump	Score	Proportion of time in automatic mode	Manual control actions
Positive	30	85.82	11	30	89.25	32.59	17.20
Negative	54	78.33	15	39	85.90	24.81	45.61
Equal	6	97.89	2	6	92.42	16.67	10.17

4.3.10 Gender Difference in Trust

Independent samples t-test did not show difference of overall trust between female and male operators ($t(10) = -.67$, $p = 0.52$, $BF_{10} = 0.54$). Figure 4.6 illustrates scatter plots in terms of relationships between the use of the automatic controllers and self-confidence. Figure 4.8-(a) and (b) mean those of females and males respectively. There seems no clear distinction between females and males. As shown in Figure 4.8-(a), relatively more female participants displayed strong self-confidence when they used the automation.

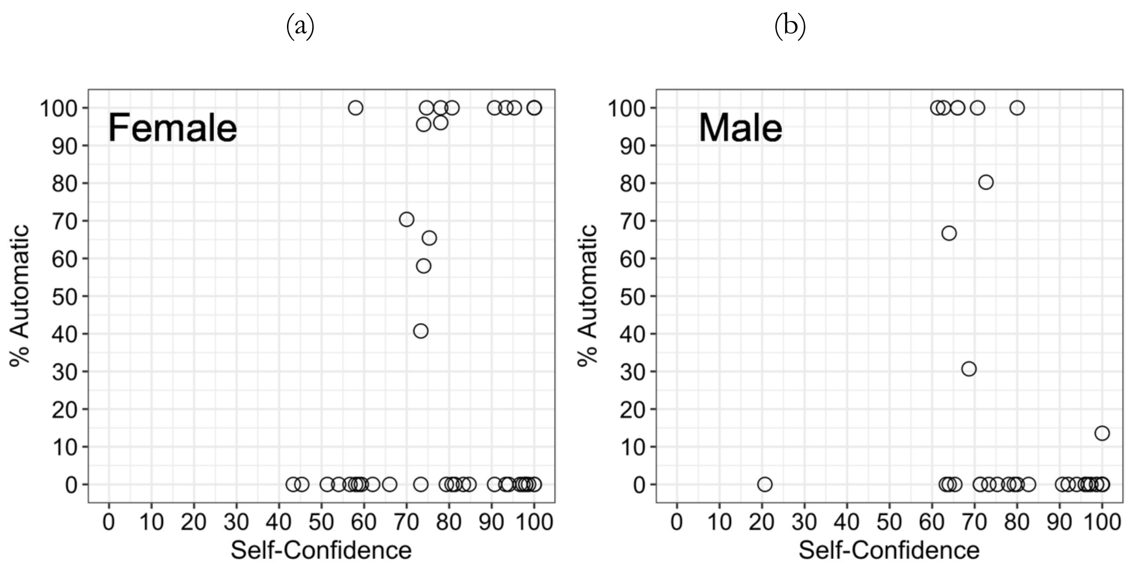
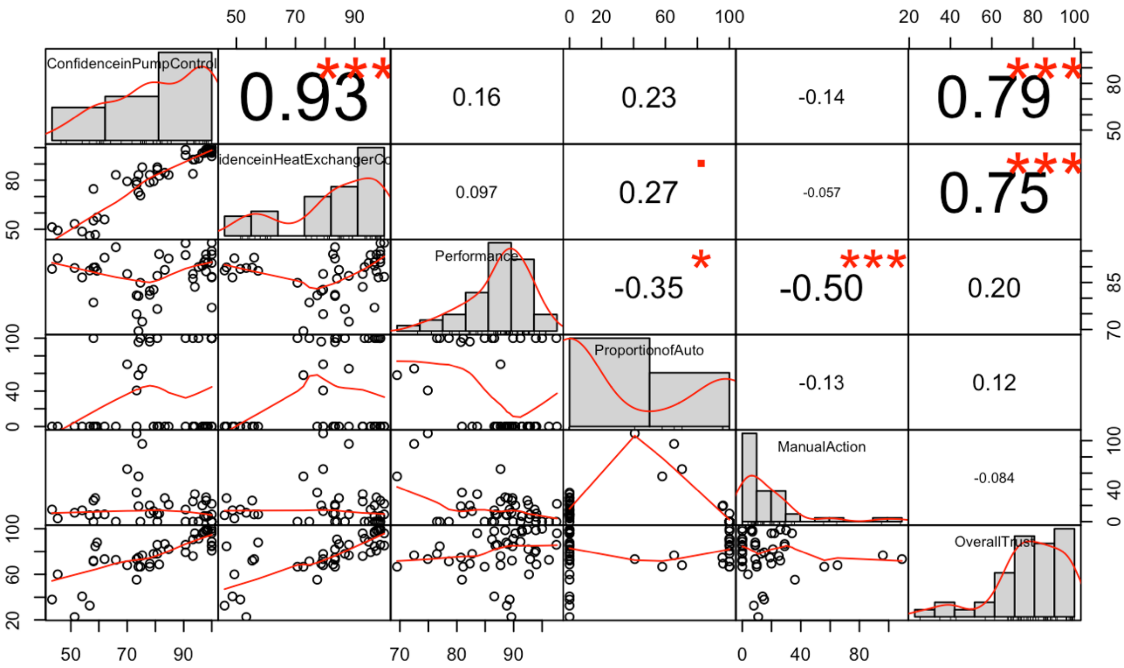


Figure 4.8 The use of the automatic controller of the pump as a function of self-confidence for (a) female and (b) male participants.

Figure 4.8-(a) and (b) are correlation charts between female and male operators' overall trust, self-confidence in pump and heater operation, performance score, proportion of time in automatic pump control mode, and the number of manual actions. In both participant groups, the positively strong correlation between self-confidence in pump and heater operation and the negatively moderate correlation between performance score and manual control actions were observed. Female operators' trust and self-confidence in operating both subsystems were positively correlated. For male operators, the proportion of time in automatic pump mode was negatively correlated with self-confidence in pump and heater operation. Male operators' overall trust was positively correlated with self-confidence in heater operation and performance score.

(a)



(b)

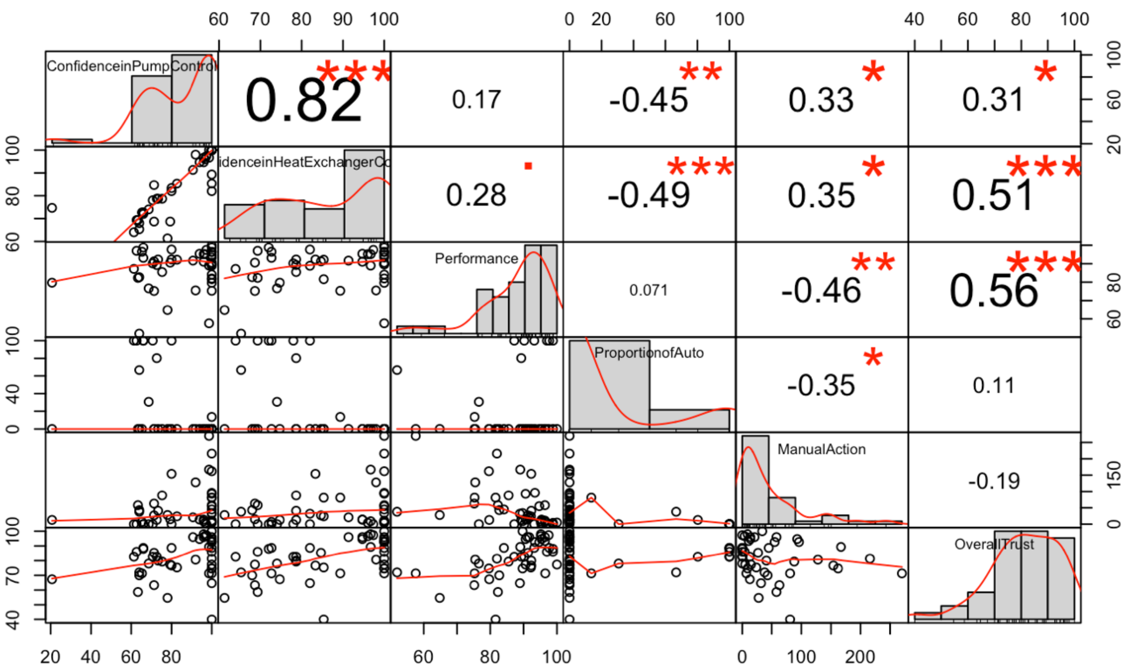


Figure 4.9 Bivariate correlation female and male operators' overall trust, self-confidence in pump and heater operation, performance score, proportion of time in automatic pump control mode, and the number of manual actions.

4.4 Discussions

Table 4.6 describe differences in experimental design and results among three studies: the original study of Muir and Moray (1996), STUDY I, and STUDY II.

	Muir and Moray (1996)	STUDY I	STUDY III
N	6 male students from engineering major	12 male students (3 X; 7from engineering)	6 female & 6 male eng students (2 X)
Bonus	O	x	X
Default pump rate	Unknown	20 litre/sec	20 litre/sec
Training program	7 sessions & 80%	7 sessions & 80%	7 sessions or 80%
Trust in the pump	Control: E > C > V	Display: H > V	Display: H > V
	Display: H > C > V	Run: 4 > 1	
	Run: 4 > 3 > 2 > 1		
Trust development	F -> D -> P	D	D -> D -> F -> D
Score	97.3	93.58	87.45
Manual control	30.13	26.76	33.78
Auto use %	11.67	63.87	26.86

Table 4.6 Differences in experimental design and statistical results among three pasteurization studies: Muir and Moray (1996), STUDY I, and STUDY III.

4.4.1 Subjective Ratings of Trust on Pump System

Muir and Moray (1996) observed that control and display properties of pump systems significantly impacts operators' overall trust in the pump system. STUDY I showed the main effect of the display property on operators' trust, and specifically, variable display error in the pump system decreased subjective ratings of overall trust in the pump compared to no display error condition. Results of STUDY III were identical to those of STUDY I as expected that the error in display property impacts operators' trust ratings (H1). Only display property has a significant effect on overall trust, with variable display error leading to decreased degrees of overall trust. Whilst no effect of both properties on ratings of trust in pump's control and display was found in STUDY I, STUDY III found that the display property affects operators' trust in both control and display, with variable display error leading to decreased trust in both pump's control and display in comparison with no display error and constant display error. These findings indicate operators' sensitivity to the display property of pump system (Muir & Moray, 1996). Despite completely identical pasteurization plant, a clear effect of display property on trust was found in only this study. The length of session for each experimental condition can be considered as a reason of the different result between STUDY I and STUDY III. The original study (Muir & Moray, 1996) and STUDY I led all participants to experience 4 runs for each condition. STUDY III shortened the length from 4 to 1 run. Considering that the accumulation of data may yield different results, ANOVA was performed again with data of the first run for each condition from STUDY I. However, there was no main effect of both properties on trust in both pump's control and display. This result indicates that engineering background of participants in this study may bring sensitivity to the display property.

The requirement assigned for the participants was to achieve performance score as higher as possible, but the exact value, 80%, was presented to them. Even though the statistical analyses results indicate that scores of participants in STUDY III is significantly lower than those of Muir and Moray (1996) and STUDY I, it is clear that participants in STUDY III could be also able to reach 80%. Here, consistent with STUDY I performance scores when interacting with pumps including constant control error (M = 83.57) were significantly lower than no control error (M = 89.63) and variable display error (M = 89.15). One possible interpretation for low performance scores in the constant control error conditions corresponding to C4, C5, and C6 is that the magnitude of constant error is relatively larger than variable error. The magnitude of constant error is determined by the pump value, however, the magnitude of variable error is depending on the input of next pump target. The control property directly impacts the adjustment of pump values unlike indirect impacts of display property on pump values. Correlation analysis did not show significant relation between overall trust and score in both STUDY I and this study, and it supports distinct impacts of system fault on trust and performance of system. This finding may confirm that the magnitude of system fault can affect operators' trust and performance independently (Lee & Moray, 1992). This study observed that decrease in operators' trust depends not how system functions but how system exhibits information, and performance is depending on the direct impacts of systems' control property. That is, display and control properties of system can affect trust and performance respectively. This finding has implication that the projection of information within systems greatly contributes on trust. The examination of how highly transparent systems affect human trust in machines and trust development is an interesting avenue for future research.

4.4.2 The Meaning of Trust in Machines

Competence and responsibility are components for explaining the meaning of trust in machines. Muir and Moray (1996)'s finding showed competence is the best predictor of the meaning of human-machine trust. STUDY I which obtained opposite finding from Muir and Moray (1996) addressed this issue with the time that participants spent in automatic pump controllers. The time in STUDY I was significantly longer than that of Muir and Moray (1996), and participants in this study also spent longer time in the automatic controller than those of Muir and Moray (1996), but lesser than those of STUDY I (H2). After the first run with the automatic controller, 10 of 11 participants in this study responded, "I would like to use this automatic controller to maximise my performance score and convenience.", then 6 of 11 participant controlled the plant manually during the experimental session. However, it is difficult to generalize the effects of gender and engineering background on automation usage because of small sample size. Further study needs to investigate this issue with large sample size considering background and gender.

STUDY I remained an issue that it may be that how the pump is integrated in the plant is perceived more important for them than the direct comparison between their actual control behaviour and the operation of system. To address this issue, their first impression of the pump system was collected, and according to the interview after the first interaction with automatic controllers, 7 of 11 participants regarded achieving high score as the most important factor contributing their trust in the pump system, and the others regarded understanding the mechanism of systems as the most important determinant of trust. Bivariate correlation between overall trust and each component showed the largest correlation between trust and responsibility for all participants who prioritized maximizing performance score. For the 4 participants who prioritized understanding the mechanism, half of them showed competence as the best predictor, and the other showed responsibility as the best predictor of meaning of human-machine trust. This result provides an evidence that why responsibility can account for the meaning of human-machine trust in this study as expected (H3).

One interesting result is operators' trust in the C6 (constant control error-variable variable error) condition in the original study, STUDY I, and STUDY III. The original study reported that responsibility can account for the meaning of trust in machines in only C6, but the results of STUDY I and III were opposite. That is, competence can predict the meaning of trust in only C6 condition. Despite non-clear values of trust ratings in the original study, subjective ratings of overall trust, trust in pump's control and display were the lowest in the C9 (variable control error-variable control error) among all experimental conditions. The results of STUDY I and III seem that trust ratings in the C6 are relatively lower than those of other conditions, however, there seems no considerable differences in ratings among C3, C6 and C9 conditions. Like this, evidence is insufficient to address this issue because significant differences in overall trust and performance and variables between C6 and other conditions. This result remains a possibility of various interpretations, such as effects of different cultural backgrounds between Canada and Japan or different social backgrounds between 1980s and the present. Further, the result of STUDY III could highlight distinction between competence and responsibility as a component for the meaning of trust in machines. Muir and Moray (1996) concluded with suggesting the importance of both components, and Merritt and Ilgen's (2008) empirical finding described that both components are approximately same with regard to trust in automation. The result of STUDY III provides different roles of competence and responsibility in the meaning of trust in machines now (2018-2020) depending on cultural contexts.

4.4.3 The Development of Trust in Machines

STUDY III used participants with background similar to those of the participants in Muir and Moray (1996) study, and provides strong evidence that dependability is the key construct that predicted overall trust at the first trial of and after the extensive practice trials. The results are similar to those of STUDY I that showed that dependability governed overall trust throughout the course of human-machine interaction. The present study is also designed to confirm the finding with reference to participants' engineering background. Despite the participants' background in Engineering, their initial trust was controlled by dependability, not faith, at the beginning of their interaction with the automatic pump system and immediately following their practice, indicating strong evidence against the model proposed by Muir and Moray (1996). This result partly supports the H4 which trust may be governed by dependability.

STUDY III failed to replicate the finding of STUDY I which dependability consistently dominates operators' trust throughout human-machine interaction. Except for the last training, dependability was the best overall predictor for predictor for trust development across different time points of their interaction. One unexpected result is that faith governed operators' trust when they finished the last training session. Additional regressions considering their understanding of system mechanisms resulted dependability and faith were the best predictors when the operator felt that they understood the mechanisms of the pump system and active heating system respectively. In the next study, it is important to clarify what drives faith to a most predictive dimension at that point. Operators first figured how the pump system works out, then understood how the active heating system works. One possible concern is that if operators felt that I completely understand all mechanisms of systems, faith can be emerged as a dominant dimension of human-machine trust. STUDY II found a similar result that faith governed trust after dependability governed it and suggested a possibility that faith incorporates dependability for drivers who received relatively less information in terms of automation. Completely same information with STUDY I was presented to participants in this study, thus it is difficult to consider the findings as same issue. However, one possible interpretation is that the feeling of understanding based on experience perhaps leads feeling of faith. Another consideration is difference of control type between pump and active heater system. Whilst the pump system was semi-automated with the free choice between manual and automatic mode, both the steam flow and temperature in the active heating system should be adjusted by manual control. Difficulty in operating each system may be depending on control type, and it affects the evolution of faith. This study addressed the finding with self-confidence and suggested two remaining questions. Consistent with STUDY II, next study should consider separate questions in terms of faith need to observe the evolution of faith in further study (e.g., Sanbonmatsu et al., 2018; Balfe et al., 2018).

4.4.4 Self-Confidence in System Operation

Self-confidence is one of main attitude influencing human-machine trust (e.g., Bandura,1982; Gist & Mitchell, 1992). Participants in STUDY I reported less confidence in system operation in the post-experiment interview, but it was insufficient to explain the effects of the first impression on trust and automation usage. This study thus considered self-confidence, in addition, participants' gender. When operators' self-confidence is low, they are likely to spend more time in automatic controllers, and when their self-confidence is high, they are prone to control system manually. Participants were asked to rate overall trust in the pump system and self-confidence in pump operation from 0 to 100 scale.

As aforementioned, 9 of 12 participants showed low self-confidence in understanding the pump system completely. 11 of 12 participants showed the willingness to rely on the automatic pump for convenience and achieving high performance score and all participants preferred the automatic mode than manual control in the early stage of training, then 6 of 11 participant (3 female, 2 males) controlled the plant manually during the experimental session. All participants' self-confidence increased with the experience of plant during the training session, and the operators who controlled it from with automation to in manual was increased. Further, in pre-training interview, 4 of 5 participants who used the automatic controllers during the experimental sessions prioritized convenience in the plant operation to high performance for the reason why they decided to use the automatic aids. This result indicates that the H4 which a long-term effect of initial impression on machines which results reliance on automation is partly supported. The continuous experience of automated machines offsets effects of the first impression on trust and self-confidence, and it can address individual differences and changes of social contexts between 1980s and the present considering that participants easily relied on the ability of automated systems in the early stage of interaction with systems. Another interesting thing is that some participants who adopted automatic pump mode during the training session attempted to do manual control with the start of experimental sessions, then the consecutive failures of experiments caused by immature manual control ability led very low ratings of trust and the reuse of automatic controller. This is indicative of the importance of training for appropriate automation usage.

There were 49 of 90 cases when ratings of self-confidence were larger than those of trust. In this case, the proportion of time in the automatic pump mode was 26%, and it was higher than those of cases which self-confidence ratings were smaller than those of trust (36%). This result seems consistent with previous researches – e.g., Lee and Moray (1994), but this should be interpreted carefully because it might reflect individual differences. For example, in the 8 cases which ratings of both trust and self-confidence is same, two participants rated like this to 2 and 3 cases. That is, a participant rated in this way to 3 of 9 experimental conditions. In addition, this study found that participants' self-confidence in pump operation did not differ across control and display properties of the pump system. Self-confidence in system operation is an important attitude in automation usage, but should consider it with familiarity with machines. That is, self-confidence can be increased and decreased along with trust in automation unlike three dimensions of trust. It is insufficient to explain dynamics of human-machine trust by only self-confidence, this issue could be regarded with operators' reliance on automation (e.g., Lee & See, 2004).

As trust varies by gender, with females typically trusting less than males (Hillesheim et al., 2017), the time in automatic pump mode and self-confidence was observed by gender. However, the statistical analyses of trust between female and male engineering students did not find any difference in trust contrary to the hypothesis (H6). One interesting point is that two female operators responded that "I could not understand the mechanism of the active heater", and a female operator did not report the understanding of pump system mechanism even though they could meet the requirement to move on the experimental session. However, all male students reported that "I could understand the mechanisms of both systems". In addition, one interesting thing is some female participants showed high degrees of self-confidence in the pump operation when they engaged with the automatic controllers throughout a whole simulation. This trend was not observed from the male group. There was no difference of performance scores during the experimental program (Female mean = 87.42; Male mean = 87.48). Bivariate correlation between self-confidence and overall trust indicates that trust is relevant to self-confidence in operating pump and active heating systems for female operators. For male group, self-confidence in active heating system control and performance score are highly related to their trust. Both groups showed self-confidence in heater operation is an important element. This issue can be discussed with the meaning of trust in machines. Correlation between overall trust and responsibility for all male students and 3 female students was larger than those between trust and competence except for two female students. For a female student, correlation between overall trust and two components were same. This is indicative of gender difference. Dispositional qualities affect what operators view machine as meaningful. However, 3 of 6 female students showed significant correlation between responsibility and trust, and competence and responsibility are significantly intercorrelated. Thus, how responsibility can capture the meaning of human-machine trust as well as how distinguish two elements should be considered with self-confidence in the next study.

4.4.5 Limitations

STUDY III made an attempt to address insufficient theoretical issues of human-machine trust revealed by STUDY I and II. This study should be also interpreted with several limitations which already addressed in the STUDY I and II, such as cultural difference between Western countries and Japan (see, Yamagishi & Yamagishi, 1994; Bliss et al., 2019). This study indicates a gap of attitudes between users in laboratory experiments and real-world. Student participants in this study relatively neglected to understand the mechanism of pasteurizer plant throughout the whole experiment. For enthusiastic participation, future research should use expert operators who have worked as supervisory controllers regardless of automation domains to examine the generalizability of the current finding to real-world. Changes of protocols can also be considered to be a solution of this problem. In this study, participants required to achieve performance scores not 100% but 80%. It is likely to lead them to be not aware of occurrence of error, specific error type, and the locus of error. As aforementioned, questionnaires for three dimensions of trust should be newly developed to capture the feelings of predictability, dependability, and faith.

Chapter 5. CONCLUSIONS

5.1 Overall Findings

The objectives of this thesis were to clarify factors shaping operators' trust in machine under supervisory control situations where requires human adaptive allocation ability between taking manual control and automatic aids. As figures out the mechanism of how human trust evolves throughout human-machine interaction, it is expected to explore approaches for designing operators' trust in automated machines. Theoretical model which trust can be developed with the interaction among three dimensions of trust: predictability, dependability, and faith conceptualized by Muir (1994) was adapted to capture main determinants influencing human-machine trust, and evaluation of theoretical model was implemented by replicating Muir and Moray's (1996) process control experiment which conducted in 1980s. In order to reinterpret the framework in the context of modern societies where automated machines are increasingly becoming ubiquitous, this thesis research attempted to replicate Muir and Moray's (1996) experiment in STUDY I and III as well as to apply the framework to the context of contemporary automated vehicles in STUDY II. The present result indicates that dependability initiates human trust in machines as well as most governs trust in the course of interaction with machines despite automation domain. The individual findings on operators' trust in machines investigated in this thesis are reported in detail in Chapter 2, Chapter 3, and Chapter 4.

The following general conclusions follow from STUDY I, II and III as described above. Briefly:

- The feeling of dependability initiates trust in automation for current users in general population.
- Social background where automated machines are ubiquitous is the main contributor of trust in automation in comparison with other dispositional qualities or the levels of pre-existing knowledge about automation.
- Continuous exposure to automated machines leads an increase in trust toward automation and self-confidence in the system operation compared to the first human-machine interaction.

- The first impression on machines shaped by interacting with the machine leads low self-confidence in the operation and high propensity to rely on automation, but it can be dissolved by the continuous human-machine interaction.
- The critical factor determining trust development is dependability in general regardless of automation domain, difficulty in the system operation, and levels of knowledge in advance of actual interaction with the machine.
- Trust is affected by not the presence of automation failure but criticality caused by automation failure.
- Current users are sensitive to changes of display property in the machine.

5.1.1 Initial Trust in Machines

Initial trust can be considered as the most important to shape further trust in automation as trust is dynamically changed by automation experience. Figure 1.2 summarized which factor determines trust with reference to time points regarded as initial trust in automation. Considering the accumulating knowledge of users about machines, (1) dispositional qualities which shaped by human affiliation, gender, age, or cultural context, (2) pre-existing knowledge, (3) specific knowledge about system before actual interaction with machines, and (4) first interaction with machines can contribute to shaping initial trust. In the series of experiments, specific knowledge about automated systems before using it was presented to all participants to equalize their knowledge levels, and the result indicates that trust is initiated from the feeling of dependability.

Rempel et al. (1985) stated that "a more common understanding of trust is probably captured by the second component of our model, dependability." In the context of human-machine trust, operators will then decide whether they can make a generalization about the machine's ability based on perceived dependability under circumstances that involve risk. There is implication that changes of society leads dependability to dominate initial trust for general users. Perception of technology may be depending on not dispositional quality but historical background. To clarify the most critical factor beginning trust in machines with dependability, the series of experiments considered demographic factors, such as age, gender, nationality, and background on trust formation. To have comprehensive understanding of how dependability fosters users' trust in machines for practical implication, social mood could be interpreted with the present result.

Modern automation designers should be aware that dependability might be a principal contributor of initial trust development for general users. This indicates that conveying information in terms of how well systems complete the assigned task before and during use of automation so that the users can develop dependability on the system might be important to prompt users' trust. As the first impression can be changed by continuous exposure to automation and its failure and might be relevant to automation usage, design features should be highlighted to be an important component for developing dependability on sophisticated systems, such as autonomous cars (Yamani & Horrey, 2018), home service robots (Mitzner et al., 2018), or even peacekeeping robots (Bliss et al., 2020). Recent automated systems are still less transparent and provide few or no explanation for how the system works or how decisions are made. Further, users of these machines, particularly those purchased off-the-shelf, typically receive little to no training on how the automation functions. Thus, this could lead to low trust in and potentially misuse of a system (Lyons et al., 2017; Sadler et al., 2016). The present finding highlights the importance of guiding appropriate levels of trust initially to help promote successful human interaction with machines in novel technologies in future.

5.1.2 The Development of Trust in Machines

It has been largely investigated that human-machine interaction increases the levels of trust across several domains. The present studies also indicate that experiencing automated machines can increase operators' subjective ratings of trust in automation. Continuous experiencing automation with exactly controlled system led increased levels of trust (Chapter 2). In the context of vehicle automation, drivers' self-reported trust after a practice drive was significantly higher than before (Chapter 3). This is why capturing factors forming trust in the course of experiencing machines. What develops trust and how it works in different supervisory control domain are considered to be important to have better understanding of mechanism of trust development. The series of experiment indicates that dependability initiates human-machine trust when operators begin to interact with unfamiliar systems as well as can account for trust in automation throughout human-machine interaction compared to predictability and faith. Despite levels of knowledge about automation, affinity for technology, automation domain, and gender, trust develops from dependability.

This result entails several theoretical implications. As trust is widely considered to be an important factor for interactions with automation (Kohn et al., 2018; Payre et al., 2016), several frameworks of trust in automation have been based, at least in part, on the findings of Muir and Moray (1996; e.g., Jian et al., 2000; Lee & See, 2004). The premise that trust grows with the interactions of three dimensions of trust helped develop frameworks for understanding trust calibration and development over time, specifically on the dimension of dependability, which Lee and See (2004) liken to process which is the middle stage of trust development. The results of the thesis study challenge the long-accepted view of the trust development initially proposed by Muir and Moray (1996) and encourage further examination of the development of trust in automation for general users across several automation domains.

The present results may reflect that the psychological structure of trust towards automation for the participants in 2020 are different from those in 1989 (e.g., cohort effect). Along with the advancement of technology since 1980s, general users of technologies that involve some level of automation might be more literate about the technologies that surrounds them than users in 1980s, calling dependability as a dominant attributor of trust. The unexpected results imply that different dimensions, not dependability, guided trust depending on participants' understanding of systems (Chapter 3 and 4), exploring the possibility that faith or predictability may govern trust development after exposure to system failure. Therefore, future studies should be investigated with longitudinal aspect of trust development towards imperfect automated systems including system failure with reference to what environmental, cultural, and psychological factors guide trust development for general users who are unlikely to be trained as highly as professional users. Additionally, these studies should carefully consider the findings of our study. Although this study found that trust develops from and along with dependability, trust might evolve differently if participants encounter novel technology, such as unmanned aerial vehicles in the modern society and the magnitude of risk due to system failures. Future research should examine whether dependability predicts overall human trust in machines during interacting with automated systems in different tasks and automation technologies. These studies are of increasing importance as new, more novel technologies are introduced to users in general population.

5.1.3 Automation Failure

Supervisory controllers should make an appropriate decision in unexpected situations to enhance the performance of system. The occurrence of error during interacting with machines has been known as a critical factor leading to decreased trust. Further, system error can lead to disuse of automation (Parasuraman & Riley, 1997; Madhaven et al., 2006). The considerable distinction between error in pasteurizer study (Chapter 2 and 4) and driving automation is whether participants could be aware of the error or not. Most operators in the pasteurizer study could not be aware of system error occurrence during the experimental session. Further, some of them perceived the existence of error, but they could be not able to characterize which and where error occurred specifically. Vehicle automation is relatively familiar automation compared to the raw milk pasteurization plant. Contrary to the pasteurizer study, participants did not have extensive training for being accustomed to the vehicle automation, and there was no specific goal to interact with the automation.

The occurrence of intervention task due to system limitation and malfunction can lower the levels of driver trust, however subsequent experience of error-free automation can rebuild the decreased trust (Hergeth et al., 2016; Kraus et al., 2019). Also, the understanding of chronic failure and how to accommodate it can recover the decreased trust (Lee & Moray, 1992; Itoh et al., 1999). For example, participants in the study of Muir and Moray (1996) reported lowered levels of trust when encountering error in the pump system, then they recovered trust after figuring out how the constant error was functioned in the pump system. It means that the exposure to variable error that they could not characterize keeps lowering trust. This finding seems that system transparency is crucial for system design. It is true, however, STUDY II and STUDY III indicate that it is not necessarily. Levels of knowledge in terms of automation did not make difference of self-reported trust ratings and determinant of initial trust (Chapter 3). No participants in STUDY III were able to figure out exact default values of pump rate for each experimental condition (Chapter 4).

Based on the findings, the purpose why operators use automated machines can be considered to be one of important factors determining whether operators could recognize system error. Another factor is whether the system error leads operators to encounter risky situation or not. Whilst drivers encounter a collision on the curved highway if they could not respond to the system limitation and malfunction in the vehicle automation, operators in the pasteurizer study did not encounter such risky situation even though they could not perceive the error and recognize which error occurred. It might also affect different attitude toward system error. Here, difficulty in system operation can be also considered as the factor leading to unawareness of system error. However, it is weak to be the main reason because all participants in the pasteurizer had adequate training for being skilled at operating systems. This result cannot conclude the misuse of automation and the failure of role in supervisory controllers because operators were able to achieve high performance (performance score over 80%) regardless of clear perception and recognition of system error. As considering that automation increases the complexity and still requires human intervention, the present studies are indicative of practical application, such as appropriate training or education in terms of all possibilities of system failures regardless of how critical and risky for people who are prone to interact with automated systems.

5.1.4 Implication for Practice

What the human-machine trust studies focus on is to investigate appropriate levels of trust leading appropriate levels of reliance on unfamiliar automation beyond trust. This thesis study challenged to have valuable insights into identifying the dimension of trust toward automation with reference to current social contexts. The empirical findings from the above three experiments have an impact on designing automation and training methods for novice users who never interacted with automated machines.

Trust training. This study confirmed that dependability is the best predictive contributor determining trust as well as predictability and faith could seize human trust with dependability depending on what users experience in the course of human-machine interaction. Continuous trial-and-error dynamically changes trust in automation. As reliance on automation and the willingness to use automation are not directly affected by the specific dimension shaping trust, users need to have training to raise their trust through trial-and-error, particularly when using automation involving critical risk, such as system-malfunction in vehicle automation.

What dependability accounts for trust in automation can be interpreted that current users consider the entire automated machine as the most important compared to the specific component of machine. As recent social background creates the tendency, it is expected that users may purchase automated machines considering not specific ability of machine but the machine itself. This study found that the knowledge levels in terms of automation calls predictability and faith as a significant determinant of trust with dependability. In related vein, clear description should be prepared for novice users considering that it is impossible to have face-to-face training for all users who attempt to interact off-the-shelf automated machine products.

Highly transparent system design is the most important design issue to address human-machine trust. Regarding situations where users could not have an extensive knowledge and understanding for operating machines – e.g., lack of time or unclear descriptive manual, the automation should be transparent to have no misunderstanding or misuse of the system. Particularly, designing human-machine interface which projects what automation is doing and other circumstances should be highly transparent to understand these information easily.

Feedbacks from automation could lead the match between mental model development and a good system image. In supervisory control, operators are obligated to observe the system behaviour. Whether users' feeling of understanding about system mechanism calls different dimensions not dependability is found in this thesis study with the observation that imperfect understanding of system may lead inappropriate reliance on automatic aids. To prevent such situations, explicit feedback in multi-modality, such as in visual, auditory, or haptic, should be prepared to help understanding of the system and anticipation of future situations for users.

5.2 Further Research

This thesis study originally aimed to confirm the psychological structure of Muir and Moray (1996)'s framework which trust evolves from faith, then dependability, lastly, predictability by replicating their control process experiment. Three experiments obtained same results in terms of dependability leading to initial trust for general users, there should be remaining works to clarify the mechanism of trust development.

Questionnaire items of Muir and Moray (1996) were adapted to examine operators' trust in automation for all series of experiment. The original study conducted by Muir and Moray (1996) clearly observed the difference with reference to trust development with three dimensions, and this questionnaire has been widely used in current studies for trust calibration. However, this thesis study did not find difference consistent with the original findings. Even though another replication study in the US that used completely same questionnaire and pasteurizer program also found dependability predicts trust for general users, it needs to consider how to capture more exact meaning of trust dimensions. For example, Balfe et al. (2018) applied Muir and Moray's (1996) questionnaire to examine participants' dependability with a question that "I can count on Automated Route Setting (ARS) to do its job" and prepared additional item to evaluate faith with items that "If ARS makes a routing decision which I am uncertain about I have confidence that ARS is correct." and "Even if I have no reason to expect that ARS will be able to deal with a situation, I still feel certain that it will." Items for faith examination by Balfe et al. (2018) include word of uncertainty whilst Muir and Moray's (1996) item assumed future circumstance with the machine. The more detailed questionnaire items, the more accurate trust calibration may be expected. Further, accurately calibrated trust can provide new view for automation design.

In related vein, unexpected results which faith best predicted trust with the experience of automation (Chapter 3 and 4) remains new research questions in terms of the psychological structure of trust. After the exposure to system failure in vehicle automation, drivers who have relatively less levels of knowledge rebuilt their trust based on feeling of faith (Chapter 3), and after operators comprehended all mechanisms of subsystems in the pasteurizer, faith was the best predictive dimension of trust (Chapter 4). To explore the possibility of faith dominating trust in such situation, more detailed questionnaire items should be developed. Because supervisory control encompasses several human factors problem in appropriate function allocation, it is important to clarify what does change the dominant contributor – e.g., the comprehension of how a system that operators should control manually works, of how system that automatic controller works, of whole system mechanism, and having full self-confidence in system operation.

Objective measurement can be facilitated to look into the relationship between automation usage and trust, and self-confidence. Operators who have high self-confidence in system operation are likely to prefer manual control to applying automatic aids (Lee & Moray, 1994). High levels of automation trust lead users to neglect monitoring status (Hergeth et al., 2016). As automation usage is closely relevant to operators' self-confidence in system control and trust development, human behaviour that objectively collected and evaluated, such as gaze behaviour could be considered to explore understanding of human-machine trust and trust dimensions.

Lastly, as aforementioned several times, this framework of human-machine trust and findings from this thesis study could be highlighted with other automation domain. STUDY I replicated the original study of Muir and Moray (1996) to confirm the previous finding with respect to modern society, STUDY II applied the obtained finding from STUDY I to domain of driving automation which needs supervisory controllers, and STUDY III partly replicated the pasteurizer study to look into relationship between trust and dispositional quality, and the attitude of self-confidence. Further application of this finding with human-machine trust framework to other domains should be investigated to offer invaluable insight into automation design regarding characteristics of automation and to generalize this tendency.

The future works as mentioned above is briefly:

- Developing questionnaire items to seize the meaning of trust dimensions should be considered.
- Objective measurements of operators' behaviours could be investigated with the degrees of their subjective ratings to observe the clear relationship between trust and user behaviour.
- The validation of Muir's (1994) framework to other automation domains should be examined.

REFERENCES

Abraham, H., Lee, C., Brady, S., Fitzgerald, G., Mehler, B., Reimer, B., & Coughlin, J.F. (2017). Autonomous vehicles and alternatives to driving: trust, preferences, and effects of age. *Proceedings of the Transportation Research Board 96th Annual Meeting*. Washington D.C.: Transportation Research Board.

Autor, D. H. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, 29(3): 3–30.

Baba, M. L., Falkenburg, D. R., & Hill, D. H. (1996). Technology management and American culture: Implications for business process redesign. *Research Technology Management*, 39(6), 44–54.

Bansal, P., Kockelman, K., (2015). Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies. *Proceedings of the 95th Annual Meeting of Transportation Research Board and for publication in Transportation Research Part A*. <http://www.caee.utexas.edu/prof/kockelman/ public_html/TRB16CAVTechAdoption.pdf> (September 4, 2015).

Bansal, P., Kockelman, K. M., & Singh, A. (2016). Assessing public opinions of and interest in new vehicle technologies: An Austin perspective. *Transportation Research Part C: Emerging Technologies*, 67, 1–14.

Barber, B. (1983). The logic and limits of trust. New Brunswick, NJ: Rutgers University Press.

Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour*, 18, 47–57.

Beller, J., Heesen, M., & Vollrath, M. (2013). Improving the driver-automation interaction: An approach using automation uncertainty. Human Factors, 55(6), 1130–1141.

Bentley, R., Hughes, J. A., Randall, D., & Shapiro, D. Z. (1995). Technological Support for Decision Making in a Safety Critical Environment. *Safety Science*, 19 (2/3), 149–156.

Billings, C. E. (1996). Human-centered aircraft automation: Principles and guidelines. Technical mem. no. 110381. NASA Ames Research Center, Moffett Field, CA.

Bliss, J. P. (2009). Faith versus Trust: The Influence of Situational Experience on Perceived Signal Credibility, and a Distinction Between Constructs. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 53(4), 269–273.

Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behaviour in response to alarms of varying reliability. *Ergonomics*, 38(11), 2300-2312.

Bliss, J. P., Q. Gao, X. Hu, M. Itoh, S. Long, Y. Papelis, and Y. Yamani. "Cross-cultural trust of robot peacekeepers as a function of dialog, appearance, responsibilities and onboard weapons." In. C. Nam & J. Lyons (Eds.), Trust in human-robot interaction: Research and applications. Elsevier. (in press)

Cahour, B., & Forzy, J. (2009). Does projection into use improve trust and exploration? An example with a cruise control system. *Safety Science*, 47, 1260–1270.

Chancey, E. T., Bliss, J. P., Yamani, Y., & Handley, H. A. H. (2017). Trust and the Compliance-Reliance Paradigm: The Effects of Risk, Error Bias, and Reliability on Trust and Dependence. *Human Factors*, 59(3), 333–345.

Chavaillaz, A., Wastell, D., & Sauer, J. (2016). System reliability, performance and trust in adaptable automation. *Applied Ergonomics*, 52, 333–342.

Chien, S. Y., Semnani-Azad, Z., Lewis, M., & Sycara, K. (2014). Towards the development of an inter-cultural scale to measure trust in automation. In Lecture Notes in Computer Science (Vol. 8528 LNCS, pp. 35–46). Basel, Switzerland: Springer International.

Choi, J. K., & Ji, Y. G. (2015). Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, 31(10), 692–702.

Cummings, M. L., & Clare, A. S. (2015). Holistic modelling for human-autonomous system interaction. *Theoretical Issues in Ergonomics Science*, 16(3), 214–231.

Davis, F.D., (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 319–340.

Damos, D., & Wickens, C. D. (1980). The acquisition and transfer of time-sharing skills. *Acta Psychologica*, 6, 569–577.

de Vries, P., Midden, C., & Bouwhuis, D. (2003). The effects of errors on system trust, self-confidence, and the allocation of control in route planning. *International Journal of Human Computer Studies*, 58(6), 719–735.

Deb, S., Strawderman, L., DuBien, J., Smith, B., Carruth, D. W., & Garrison, T. M. (2017). Evaluating pedestrian behavior at crosswalks: Validation of a pedestrian behavior questionnaire for the U.S. population. *Accident Analysis and Prevention*, 106(May), 191–201.

Deutsch, M. (1960). The effect of motivational orientation upon trust and suspicion. *Human Relations*, 13, 123–139

Deutsch, M. (1973). The resolution of conflict: Constructive and destructive processes. New Haven, CN: Yale University Press.

Dikmen, M., & Burns, C. M. (2016). Autonomous driving in the real world: Experiences with Tesla autopilot and summon. In Proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications (Automotive'UI 16), (pp. 225–228).

Domeyer, J., Venkatraman, V., Price, M., Lee, J. D., Collaborative, T., & Arbor, A. (2018). Characterizing Driver Trust in Vehicle Control Algorithm Parameters. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1), 1821–1825.

Doney, P. M., Cannon, J. P., & Mullen, M. R. (1998). Understanding the influence of national culture on the development of trust. *Academy of Management Review*, 23, 601–620.

Du, N., Haspiel, J., Zhang, Q., Tilbury, D., Pradhan, A. K., Yang, X. J., & Robert, L. P. (2019). Look who's talking now: Implications of AV's explanations on driver's trust, AV preference, anxiety and mental workload. *Transportation Research Part C: Emerging Technologies*, 104(September 2018), 428–442.

Dzindolet, M. T., Pierce, L. G., Beck, H. P., & Dawe, L. A. (2002). The perceived utility of human and automated aids in a visual detection task. *Human Factors*, 44(1), 79-94.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors*, 37(1), 65–84.

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77, 167–181.

Feldhütter, A., Gold, C., Hüger, A., & Bengler, K. (2016). Trust in Automation as a Matter of Media Influence and Experience of Automated Vehicles. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1), 2024–2028.

Gao, J., & Lee, J. D. (2006). A dynamic model of interaction between reliance on automation and cooperation in multi-operator multi-automation situations. *International Journal of Industrial Ergonomics*, 36, 512–526.

Ghazizadeh, M., Lee, J. D., & Boyle, L. N. (2012). Extending the Technology Acceptance Model to assess automation. *Cognition, Technology & Work*, 14(1), 39–49.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, 71, 147–177.

Hartwich, F., Witzlack, C., Beggiato, M., & Krems, J. F. (2019). The first impression counts – A combined driving simulator and test track study on the development of trust and acceptance of highly automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, (65), 522–535.

Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep Your Scanners Peeled: Gaze Behavior as a Measure of Automation Trust during Highly Automated Driving. *Human Factors*, 58(3), 509–519.

Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior Familiarization With Takeover Requests Affects Drivers' Takeover Performance and Automation Trust. Human Factors, 59(3), 457–470.

Hillesheim, A. J., Rusnock, C. F., Bindewald, J. M., & Miller, M. E. (2017). Relationships between User Demographics and User Trust in an Autonomous Agent. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1), 314–318.

Ho, G., Wheatley, D., & Scialfa, C. T. (2005). Age differences in trust and reliance of a medication management system. *Interacting With Computers*, 17, 690–710.

Hoff, K. A., & Bashir, M. (2014). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434.

Hoffman, R. R., Johnson, M., Bradshaw, J. M., & Underbrink, A. (2013). Trust in automation. *IEEE Intelligent Systems*, 28(1), 84–88.

Hulse, L. M., Xie, H., & Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science*, 102, 1–13.

Igbaria, M., & Iivari, J. (1995). The effects of self-efficacy on computer usage. *Omega*, 23(6), 587–605.

Itoh, M., Abe, G., & Tanaka, K. (1999). Trust in and Use of Automation- Their Dependence on Occurrence Patterns of Malfunctions. In *IEEE SMC '99 Conference Proceedings* (pp. 715–720).

Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4(1), 53-71.

Kabacoff., R. I. (2020). R in Action (3rd Ed). Shelter Island, NY: Meanning Publications.

Kantowitz, B. H., Hanowski, R. J., & Kantowitz, S. C. (1997). Driver acceptance of unreliable traffic information in familiar and unfamiliar settings. *Human Factors*, 39(2), 164–176.

Kaur, K., & Rampersad, G. (2018). Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars. *Journal of Engineering and Technology Management*, 48(May), 87–96.

Kato, S., Hashimoto, N., Ogitsu, T., & Tsugawa, S. (2010). Driver Assistance Systems with Communication to Traffic Lights – Configuration of Assistance Systems by Receiving and Transmission and Field Experiments, *Journal of Robotics and Mechatronics*, 22(6), 737–744

Kelly, C., Boardman, M., Goillau, P. and Jeannot, E. (2001). Principles and Guidelines for the Development of Trust in Future ATM Systems: A Literature Review, European Organisation for the Safety of Air Navigation: 48pp.

Kohn, S. C., Quinn, D., Pak, R., de Visser, E. J., & Shaw, T. H. (2018). Trust Repair Strategies with Self-Driving Vehicles: An Exploratory Study. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 62(1), 1108–1112.

König, M., & Neumayr, L. (2017). Users' resistance towards radical innovations: The case of the self-driving car. *Transportation Research Part F: Traffic Psychology and Behaviour*, 44, 42–52.

Körber, M., Baseler, E., & Bengler, K. (2018a). Introduction matters: Manipulating trust in automation and reliance in automated driving. *Applied Ergonomics*, 66, 18–31.

Körber, M., Prasch, L., & Bengler, K. (2018b). Why Do I Have to Drive Now? Post Hoc Explanations of Takeover Requests. Human Factors, 60(3), 305–323.

Kramer, A.F., Larish, J., & Strayer, D.L. (1995). Training for attentional control in dual-task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied*, 1, 50–76.

Kraus, J., Scholz, D., Stiegemeier, D., & Baumann, M. (2019). The More You Know: Trust Dynamics and Calibration in Highly Automated Driving and the Effects of Take-Overs, System Malfunction, and System Transparency. *Human Factors.*

Krueger, R., Rashidi, T. H., & Rose, J. M. (2016). Preferences for shared autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, 69, 343–355.

Lacson, F. C., Wiegmann, D. A., & Madhavan, P. (2005). Effects of Attribute and Goal Framing on Automation Reliance and Compliance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 49(3), 482–486.

Lee, C., Seppelt, B., Reimer, B., Mehler, B., & Coughlin, J. F. (2019). Acceptance of Vehicle Automation: Effects of Demographic Traits, Technology Experience and Media Exposure. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1), 2066–2070.

Lee, J. D., & Kolodge, K. (2019). Exploring trust in self-driving vehicles with text analysis. *Human Factors*, 62(2), 260–277.

Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *International Journal of Human-Computer Studies*, 40(1), 153–184.

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, 46(1), 50–80.

Lee, J., & Moray, N. (1992). Trust, Control Strategies and Allocation of Function in Human-Machine Systems. *Ergonomics*, 35(March), 1243–1270.

Lewandowsky, S., Mundy, M., & Tan, G. P. A. (2014). The dynamics of trust: Comparing humans to automation. *Journal of Experimental Psychology: Applied*, 4(2), 104–123.

Long. S., Lee. J., Yamani, Y., Unverricht., J., & Itoh. M. (2020). Examining Human-Machine Trust: A Replication Study of Muir and Moray (1996). (Under review)

Lopatovska, I., Rink, K., Knight, I., Raines, K., Cosenza, K., William, H., Sorscher, P., Hirsch, D., Li, Q. & Martinez, A. (2018). Talk to me: exploring interactions with the Amazon Alexa. Journal of Librarianship & Information Science. Advance online publication.

Louw, T., Markkula, G., Boer, E., Madigan, R., & Carsten, O. (2017). Coming Back into the Loop: Drivers' Perceptual-Motor Performance in Critical Events after Automated Driving. *Accident Analysis and Prevention*, 108, 9–18.

Lyons, J. B., Sadler, G. G., Koltai, K., Battiste, H., Ho, N. T., Hoffmann, L. C., & Shively, R. (2017). Shaping trust through transparent design: theoretical and experimental guidelines. In *Advances in Human Factors in Robots and Unmanned Systems* (pp. 127-136). Springer, Cham.

Madhavan, P., & Wiegmann, D. A. (2007b). Similarities and differences between human-human and human-automation trust: An integrative review. *Theoretical Issues in Ergonomics Science*, 8, 277–301.

Madhavan, P., Wiegmann, D. A., & Lacson, F. C. (2006). Automation Failures on Tasks Easily Performed by Operators Undermine Trust in Automated Aids. *Human Factors*, 48(2), 241–256.

Madsen, M., & Gregor, S. (2000). Measuring human-computer trust. In G. Gable & M. Vitale (Eds.), In *Proceedings of the 11th Australasian Conference on Information Systems* (p. 53). Brisbane, Australia: Information Systems Management Research Centre.

Manzey, D., Reichenbach, J., & Onnasch, L. (2012). Human Performance Consequences of Automated Decision Aids: The Impact of Degree of Automation and System Experience. *Journal of Cognitive Engineering and Decision Making*, 6(1), 57–87.

Merritt, S. M., & Ilgen, D. R. (2008). Not All Trust Is Created Equal: Dispositional and History-Based Trust in Human-Automation Interactions. *Human Factors*, 50(2), 194–210.

Metzger, U., & Parasuraman, R. (2001). The role of the air traffic controller in future air traffic management : An Empirical Study of Active Control versus Passive Monitoring. *Human Factors*, 43(4), 519–529.

Meyer, J. (2001). Effects of warning validity and proximity on responses to warnings. *Human Factors*, 43(4), 563-572.

Mitzner, T. L., Tiberio, L., Kemp, C. C., & Rogers, W. A. (2018). Understanding healthcare providers' perceptions of a personal assistant robot. *Gerontechnology*, 17(1), 48.

Muir, B. M. (1994). Trust in automation: 1. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics*, 37, 1905–1922.

Muir, B. M., & Moray, N. (1996). Trust in automation: 2. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, 39, 429–460.

Moray, N., & Inagaki, T. (1999). Laboratory studies of trust between humans and machines in automated systems. *Transactions of the Institute of Measurement & Control*, 21(4–5), 203–211.

Moray, N., Hiskes, D., Lee, J., & Muir, B. M. (1995). Trust and human intervention in automated systems. In J.-M. Hoc, P. C. Cacciabue & E. Hollnagel (Eds.), Expertise and technology: Cognition & human-computer cooperation (pp. 183-194). Hillsdale, NJ: Erlbaum.

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology: Applied*, 6(1), 44–58.

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), 381–410.

Parasuraman, R., & Miller, C. A. (2004). Trust and etiquette in high-criticality automated systems. *Communications of the ACM*, 47, 51–55.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253.

Parasuraman, R., Sheridan, T. B., & 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.

Payre, W., Cestac, J., & Delhomme, P. (2014). Intention to use a fully automated car: Attitudes and a priori acceptability. *Transportation Research Part F: Traffic Psychology and Behaviour*, 27, 252–263.

Payre, W., Cestac, J., & Delhomme, P. (2016). Fully automated driving: Impact of trust and practice on manual control recovery. *Human Factors*, 58(2), 229-241.

Rajaonah, B., Tricot, N., Anceaux, F., & Millot, P. (2008). The role of intervening variables in driver–ACC cooperation. *International Journal of Human-Computer Studies*, 66, 185–197.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13, 257–266.

Rasmussen, J., Pijtersen, A.M., Goodstein, L.P., 1994. Cognitive Systems Engineering. Wiley, New York.

Reeves, B., & Nass, C. (1996). The media equation: How people treat computers, television, and new media like real people and places. New York: Cambridge University Press.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, 49(1), 95.

Riley, V. (1994). A theory of operator reliance on automation. In M. Mouloua & R. Parasuraman (Eds.). Human performance in automated systems: Recent research and trends (pp. 8-14). Hillsdale, NJ: Erlbaum.

Riley, V. (1996). Operator reliance on automation: Theory and data. In R. Parasuraman & M. Mouloua (Eds.), Automation and human performance: Theory and applications (pp. 19–35). Hillsdale, NJ: Lawrence Erlbaum.

Rödel, C., Stadler, S., Meschtscherjakov, A., & Tscheligi, M. (2014). Towards Autonomous Cars: The Effect of Autonomy Levels on Acceptance and User Experience. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1-8). (AutomotiveUI'14). New York, NY, USA: Association for Computing Machinery.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, 35, 651–665.

Rouder, J. N., & Morey, R. D. (2012). Default Bayes Factors for Model Selection in Regression. *Multivariate Behavioral Research*, 47(6), 877–903.

Sadler, G., Battiste, H., Ho, N., Hoffmann, L., Johnson, W., Shively, R., Lyons, J. & Smith, D. (2016). Effects of transparency on pilot trust and agreement in the autonomous constrained flight planner. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)* (pp. 1-9). IEEE.

Sanbonmatsu, D. M., Strayer, D. L., Yu, Z., Biondi, F., & Cooper, J. M. (2018). Cognitive underpinnings of beliefs and confidence in beliefs about fully automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 55, 114–122.

Sanchez, J., Fisk, A. D., & Rogers, W. A. (2004). Reliability and Age-Related Effects on Trust and Reliance of a Decision Support Aid. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 48(3), 586–589.

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. In G. Salvendy (Ed.), Handbook of human factors and ergonomics (2nd ed., pp. 1926–1943). New York, NY: Wiley.

Sato, T., Yamani, Y., Liechty, M., & Chancey, E. T. (2019). Automation trust increases under high-workload multitasking scenarios involving risk. *Cognition, Technology & Work*, 22, 399–407

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems. *Human Factors*, 58(3), 377–400.

Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2015). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*. Advance online publication.

Seppelt, B. D., & Lee, J. D. (2007). Making adaptive cruise control (ACC) limits visible. *International Journal of Human-Computer Studies*, 65, 192–205.

Seppelt, B. D., & Lee, J. D. (2019). Keeping the driver in the loop: Dynamic feedback to support appropriate use of imperfect vehicle control automation. *Journal of Human Computer Studies*, 125(November 2018), 66–80.

Sheridan, T. B. (1988). Trustworthiness of command and control systems. *IFAC Proceedings Series*, 21(3), 427–431.

Sheridan, T. B. & Hennessy, R. T. (Eds.). (1984). Research and modeling of supervisory control behauior. Washington, D. C.: National Academy Press.

Sheridan, T. B., & Ferrell, W. (1974). Man-machine systems: Information, control, and decision models of human performance. Cambridge, MA: MIT Press.

Sheridan, T. B., & Verplank, W. L. (1978). Human and computer control of undersea teleoperators. Man-Machine Systems Laboratory report, Massachusetts Institute of Technology, Cambridge.

Sheridan, T. B., Fischhoff, B., Posner, M., & Pew, R. W. (1983). Supervisory Control Systems. Research Needs for Human Factors (pp. 49-77). Washington, D.C.: National Academy Press.

Skjerve, A. B. M., & Skraaning, G., Jr. (2004). The quality of human–automation cooperation in human-system interface for nuclear power plants. *International Journal of Human–Computer Studies*, 61, 649–677.

Society of Automotive Engineers International J3016 (2018). Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Warrendale, PA: SAE International.

Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors*, 60(8), 1095–1116.

Walker, F., Boelhouwer, A., Alkim, T., Verwey, W. B., & Martens, M. H. (2018). Changes in Trust after Driving Level 2 Automated Cars. *Journal of Advanced Transportation*, 2018(August), 1–9.

Ward, C., Raue, M., Lee, C., D'Ambrosio L., & Coughlin, J. F. (2017). Acceptance of automated driving across generations: the role of risk and benefit perception, knowledge, and trust. In: Kurosu M. (eds) Human-Computer Interaction. User Interface Design, Development and Multimodality. HCI 2017. Lecture Notes in Computer Science, Springer, Cham, 10271, 254-266.

Wickens, C. D., & Hollands, J. G. (2000). Engineering psychol- ogy and human performance (4th ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Wiegmann, D. A., Rich, A., & Zhang, H. (2001). Automated diagnostic aids: The effects of aid reliability on users' trust and reliance. *Theoretical Issues in Ergonomics Science*, 2(4), 352–367.

Wiener, E. L., & Curry, R. E. (1980). Flight-deck automation: Promises and problems. *Ergonomics*, 23(10), 995-1011.

Wintersberger, P., Frison, A., Riener, A., & Boyle, L. N. (2016). Towards a Personalized Trust Model for Highly Automated Driving. *Mensch Und Computer*, (September 2016), 2–7.

Xu, A., & Dudek, G. (2015). OPTIMo: Online Probabilistic Trust Inference Model for Asymmetric Human-Robot Collaborations. *ACM/IEEE International Conference on Human-Robot Interaction*, 2015, 221–228.

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, 18, 129–166.

Yamani, Y., & Horrey, W. J. (2018). A theoretical model of human-automation interaction grounded in resource allocation policy during automated driving. *International Journal of Human Factors and Ergonomics*, 5(3), 1–15.

Zuboff, S. (1988). *In the age of smart machines: The future of work technology and power*. New York: Basic Books.

APPENDIX A INSTRUCTION SHEET FOR STUDY I

実験についての説明

本日は，実験に参加していただき，誠にありがとうございます．まず，実験の目的と，行っていただく作業の概要，全体のスケジュールを説明します．

はじめに

あなたは生乳低温殺菌加工工場のオペレータであるとします。 あなたの仕事は生乳を75℃～85℃の温度に加熱してバクテリアを殺菌することです。

プラントは完全に自動化されているため、自動的に実行できます。 オペレータは、プラントを手動で制御することも、プラントを手動と自動を組み合わせて制御することもできます。

現実の世界と同様に、プラントでは、時折不具合が起こることがあります。 はじめにシステムの操作に慣れていただくためにトレーニングを行います。プラントが自動的に稼働しているときに何が起こるのか、手動で制御を引き継ぐときに何が起こるのかを学習していただきます．本実験は合計 3 日間（9 時間）の作業を行います。

私たちはプラントに関するあなたの態度や気持ちに関心があります。ポンプがどれほど信頼できるものであるのかを評価します。各実験の後、あなたは機械への信頼の主観的推定値を評価していただきます。プラントの動作には多少の変動があることがあります。それを制御する最も良い方法を試してください。

1

実験で行っていただく作業

本実験では，あなたの仕事は生乳を 70℃～85℃の温度に加熱して低温殺菌をすることです．オペレータの務めは，安全性を保った上で，システム性能を最大にしてより多くの殺菌された生乳を作り出すことです．

図 1 の四つの黒い三角は三方弁（THREE-WAY VALVE）を表し，生乳の輸送先が決定されます．生乳の温度が 70～85℃の場合，生乳は出力タンク(目標のタンク)に移送されます．温度が 85℃を超えると，生乳の品質が損なわれるため，廃棄タンクに運ばれます．生乳が規定の温度よりも低い場合(70℃以下)，再びメインタンクに戻されて再度，低温殺菌システムに入ります．最大出力を出すためには，ポンプと熱交換システムの 2 つのサブシステムを適切に操作する必要があります．

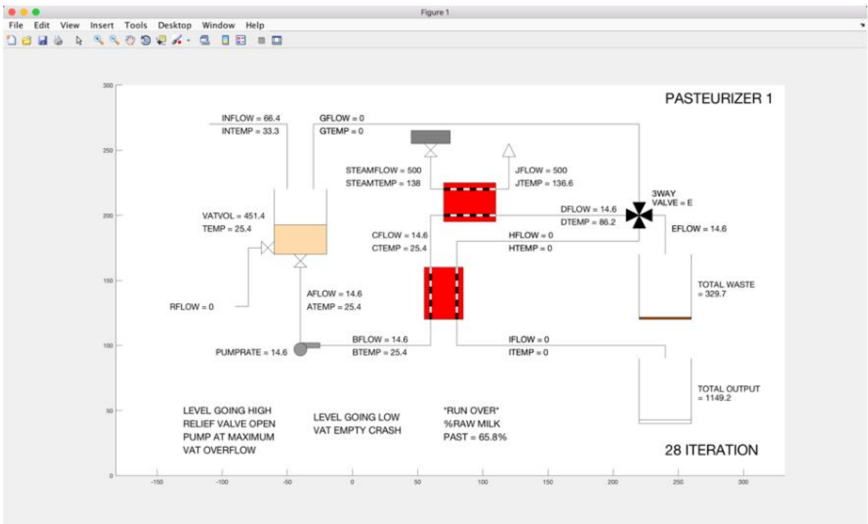


図 1. この実験での低温殺菌プラントシミュレーション

・ポンプシステム

　ポンプサブシステムは半自動的に操作されます．あなたは自動および手動の操作モードを切り替えることができます．手動の際は，あなたはコマンドを入力して新しいポンプ目標への流量を調整する必要がありますが，自動モードではシステムは自動的に生乳の流量とシステムに合った流量をセットします．

　手動の際は，ポンプの流量を調節できます．ポンプは 0~100L/分の範囲で調節を行えます．

・熱交換システム

　スチームバルブの制御を行うことができます．バルブの調節により，スチームの流量と温度を調節できます．バルブは 0~100L/分の範囲で調節を行うことができます．

質問はありませんか．

スケジュール

実験全体はトレーニングと本実験に分かれており，トレーニングではシステム操作に慣れていただきます．あなたの進行具合により時間が前後します．その後，本実験を 9 時間行います．1 回の参加につき 3 時間の実験を実施します．実験は約 3 週間行い，毎週月曜日，水曜日，金曜日に実験を行います．

実験中に体調が悪くなった場合は待機している実験者に声をかけていただければ，休憩時間を設けます．休憩をしても体調が回復しない場合は実験を中止しますが，実験が中止されても参加時間に応じた謝金は支払われます．

No.	内容	時間長
1	実験説明	15 分
2	トレーニング&アンケート	7.5 時間
3	本実験 1&アンケート	1 時間
4	本実験 2&アンケート	1 時間
5	本実験 3&アンケート	1 時間
6	本実験 4&アンケート	1 時間
7	本実験 5&アンケート	1 時間
8	本実験 6&アンケート	1 時間
9	本実験 7&アンケート	1 時間
10	本実験 8&アンケート	1 時間
11	本実験 9&アンケート	1 時間

何か質問があれば，そばにいる実験者にいつでも問い合わせてください．

質問はありませんか．

Appendix B Instruction Sheet for STUDY I and III

Welcome to Human- Machine Trust Experiment

本実験では，あなたは，別紙図1に示されている，生乳の低温殺菌プラントのシュミュレータの操作を行っていただきます．あなたが行うことは，出力タンク (TOTAL OUTPUT) の量を最大にすることです．その量は，低温殺菌が正しく行われた量からメインタンクからポンプを通して処理された総生乳量を割ったものです．あなたのパフォーマンスは，それぞれのシミュレーションとディスプレイ下部の(“%RAW MILK PAST = XX.X%”)という表示から確認することができます．そこで，あなたはトレーニングセッションで 80%以上のパフォーマンスを達成していただきます．

シミュレータ上のプラントでは，生乳が流入するパイプはディスプレイの左上に表示されます．牛乳はメインタンクに流れていきます．ダイアグラムの左下にあるフィードストックポンプは，メインタンクからの生乳をパッシブおよびアクティブ熱交換システムに送ります．アクティブ熱交換システムは生乳の温度を上昇させます．アクティブ熱交換システムにおけるスチームポンプの速度とスチームの温度を変更して，熱交換システムを通過する生乳の温度を制御することができます．熱交換システムに続いて，THREE-WAY VALVE が自動的にミルクの流れを決めます．アクティブ熱交換システムから出る生乳の温度が規定の温度より高すぎる場合，それは商品としては使用できないために破棄タンクに送られます．生乳の温度が規定の温度より低すぎる場合は，さらに低温殺菌を必要とし，メインタンクに再度戻されます．

あなたは，キーボードの割り当てられたボタンを使用して，ポンプ速度 (PUMPRATE)，スチームの流量 (STEAMFLOW) およびスチームの温度 (STEAMTEMP) を制御することができます．さらに，ポンプシステムを自動モードで動作させたり，解除したりすることができます．また，ポンプシステムモードの状態（手動と自動）は，ディスプレイの右上隅に表示されます (Pump System Mode = Auto/Manual)．

各条件は 80 回の反復を4トライアルで構成され，1トライアル約 8 分かかります．それぞれの実行が完了したら，低温殺菌装置に関する信頼性アンケートに記入していただきます．

メインタンクが空きになったら，シミュレーションは中止されます．

何か質問はありますか．

私たちの実験への協力ありがとうございます．

5

トレーニング

あなたは，これからトレーニング用低温殺菌装置のシミュレートを行っていただきます．ここでは，あなたは手動でのポンプ操作，自動でのポンプ操作、およびポンプ操作の切り替えについての練習を十分に行っていただき，現実のオペレータと同様に操作ができるようにしていただきます．1 時間のセッションで，それぞれ 80 回の反復を 4 トライアル， 1 トライアルにつき約 10 分で構成されています．また，トレーニングは 5 段階あります．それぞれについて以下に示します．

・練習1：練習マニュアルを使用して操作を行う
・練習2：手動でポンプの制御を行う
・練習3：手動と自動の切り替えを練習する
・練習4：自動でポンプの制御を行う
・練習5：自動または手動制御を自由に切り替えながら操作を行う

練習5のシミュレート終了時の達成度が 80%を超えた段階で，練習を終了とします．また，シミュレート実行ごとに，アンケートへの回答をお願いします．

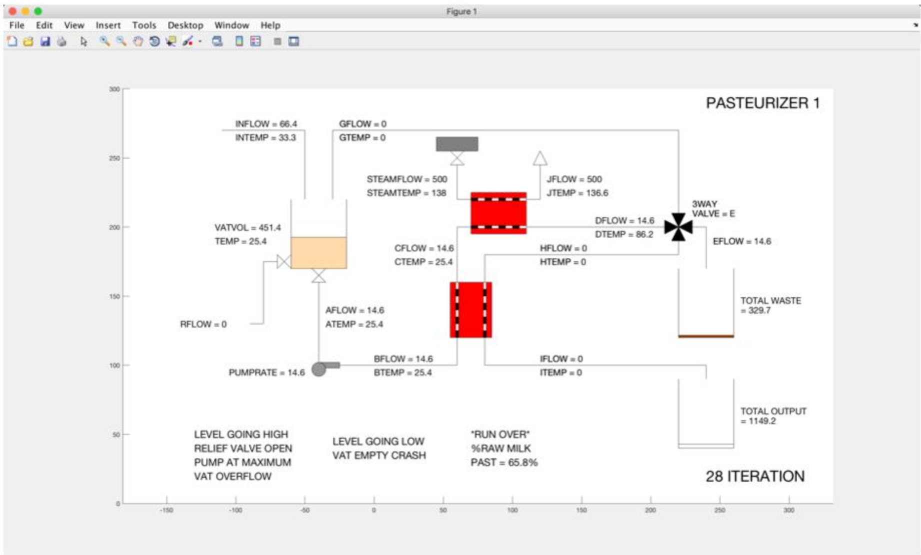


図 1. この実験での低温殺菌プラントシミュレーション

表 1. パラメーターと内容

パラメーター	内容
INFLOW	メインタンクに入る生乳の流速
INTEMP	メインタンクに入る生乳の温度
VATVOL	メインタンクに入っている生乳の量
TEMP	メインタンクに入っている生乳の温度
RFLOW	メインタンクの生乳の量が 950 L を超えた時に捨てられた生乳の総量
PUMPRATE	ポンプの速度
AFLOW/BFLOW/CTEMP	ポンプの速度
ATEMP/BTEMP/CTEMP	メインタンクから出てきた生乳の温度
STEAMFLOW	熱交換システムに関与するスチームの流速
STEMMTEMP	熱交換システムに関与するスチームの温度
DFLOW	熱交換システムを通った生乳の流速
DTEMP	熱交換システムを通った生乳の温度
3WAYSVALVE	熱交換システムを通った生乳の温度から輸送先を決定するバルブ $\mathrm{3WAYSVALVE} = \begin{cases} EPipeLine, & \mathrm{DTEMP} > 85 \\ GPipeLine, & 75 < \mathrm{DTEMP} \\ HPipeLine, & 75 \leq \mathrm{DTEMP} \leq 85 \end{cases}$
EFLOW	廃棄タンクに捨てられる品質が損なわれた生乳の流速
ETEMP	廃棄タンクに捨てられる品質が損なわれた生乳の温度
TOTAL WASTE	損なわれた生乳の総量
GFLOW	メインタンクに戻される生乳の流速
GTEMP	メインタンクに戻される生乳の温度
HFLOW	出力タンク(目標のタンク)に流される低温殺菌された生乳の流速
HTEMP	出力タンク(目標のタンク)に流される低温殺菌された生乳の温度
IFLOW	パシブヒーターを通った後の生乳の流速
ITEMP	パシブヒーターを通った後の生乳の温度
TOTAL OUTPUT	低温殺菌された生乳の総量

シミュレーション操作

‘0’ = シミュレーションスタート

ポンプシステム操作

‘1’ = 手動/自動モード切り替え

‘Y’ = ポンプの流量増加

‘U’ = ポンプの流量減少

熱交換システム操作

‘H’ = スチームの流量増加

‘J’ = スチームの流量減少

‘N’ = スチームの温度増加

‘M’ = スチームの温度減少

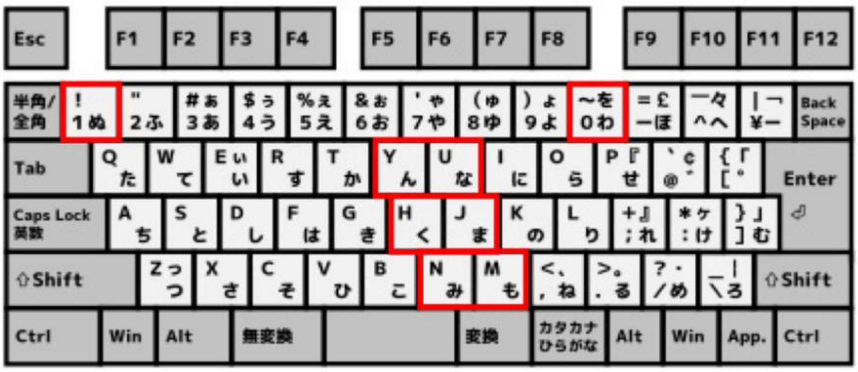


図 2. シミュレーションで操作するキー

アンケートでは以下のような評価を行います．

1. ポンプはどの程度適切に機能を果たしていますか？
2. ポンプの挙動はその時々においてどの程度予測できますか？
3. このポンプを作業遂行にあたりどの程度頼りにすることができますか？
4. このポンプは低温殺菌システムにおける役割をどの程度果たしていましたか？
5. 時刻が違っても同じような状況であれば，このポンプは同様に応答しますか？
6. 将来，このポンプがこれまで経験したことのないシステム状態に対してどの程度対応できると思いますか？
7. このポンプが正確に応答するとどの程度信頼できますか？
8. ポンプの表示精度をどの程度信頼していますか？
9. 全体としてあなたはこのポンプをどの程度信頼していますか？

ある実験の評価が，他の実験の評価に影響を与えないようにしてください

質問はありませんか．

よろしければ，トレーニングを始めます．

本実験

これから生乳の低温殺菌工程のシミュレータの制御を行っていただきます．80 回反復のシミュレーションの制御を 4 トライアル行っていただきます．また，本セッション終了後にアンケートへの記入を宜しくお願いします．

また，プラント内の装置の問題により，トライアルごとに多少の違いがあり，手動または自動制御の効果が少なくなる場合があります．

実験中は低温殺菌された生乳の量が最大になるように留意してください
メインタンクの流量は約 500L に維持できるように留意してください

質問はありませんか．

よろしければ，本実験を開始します．

Appendix C Trust Questionnaire for STUDY I

アンケート

参加者番号：
トレーニング / 実験 _____回目・トライアル_____目の反復

1. **ポンプはどの程度適切に機能を果たしていますか？**

0 ―――――――――――――――――――――――――――― 100

全く機能してない　　　　　　　　　　　　　　　　　　　　　　　　　　完全に機能している

2. **ポンプの挙動はそのときどきにおいてどの程度予測できますか？**

0 ―――――――――――――――――――――――――――― 100

全く予測してない　　　　　　　　　　　　　　　　　　　　　　　　　　完全に予測できる

3. **このポンプを作業遂行にあたりどの程度頼りにすることができますか？**

0 ―――――――――――――――――――――――――――― 100

全く頼りにできない　　　　　　　　　　　　　　　　　　　　　　　　　完全に頼りにできる

4. **このポンプは低温殺菌システムにおける役割をどの程度果たしていますか？**

0 ―――――――――――――――――――――――――――― 100

全く果たしてない　　　　　　　　　　　　　　　　　　　　　　　　　　完全に果たしている

5. **時刻が違っても同じような状況であれば，このポンプは同様に応答しますか？**

0 ―――――――――――――――――――――――――――― 100

全く違う　　　　　　　　　　　　　　　　　　　　　　　　　　　　　完全に同様である

6. 将来、このポンプがこれまで経験したことのないシステム状態に対してどの程度対応できると思いますか？

0 ―― 100

全くできない　　　　　　　　　　　　　　　　　　　　　　　　　　　完全にできる

7. このポンプが正確に応答するとどの程度信頼していますか？

0 ―― 100

全く信頼してない　　　　　　　　　　　　　　　　　　　　　　　　完全に信頼している

8. ポンプの表示精度をどの程度信頼していますか？

0 ―― 100

全く信頼してない　　　　　　　　　　　　　　　　　　　　　　　　完全に信頼している

9. 全体としてあなたはこのポンプをどの程度信頼していますか？

0 ―― 100

全く信頼してない　　　　　　　　　　　　　　　　　　　　　　　　完全に信頼している

Appendix D Instruction Sheet for STUDY II

自動運転に対するドライバの信頼に関する
実験の教示書

筑波大学
認知システムデザイン研究室

1

実験のタイムスケジュール

順番	内容	所要時間
1	実験説明	15分
2	アイトラッカーの環境確認	10分
3	運転シミュレータの練習走行	10分
4	実験その１	20分
5	休憩	10分
6	実験その2	15分
7	最終アンケート作成	5分

2

- この実験では，自動走行システムを利用して高速道路を走行していただきます．
- 自動走行システムを利用しているときに，皆様がどのようなことを感じるかを調査します．
- まず，自動走行システムの使い方をご説明します．

3

自動運転システムの使い方の説明

- 運転席の左側にあるハンドルの自動運転スイッチ（○）を操作すると自動運転モードになります．

4

自動運転システムの使い方の説明

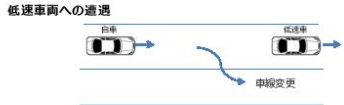
- 自動運転モード中は，以下に示す表示が出ます．
- 下記表示点灯中は，ご自身によるペダル，ハンドルの操作をしなくても自動で走行できます（80km/h）．
- 自動運転システムは様々な交通場面に対応して走行することができます．
- 基本的に自動運転システムで走行を行いますので，システム作動中はハンドルとペダルには触れなくても大丈夫です．
- また，運転引き継ぎ要請（以降に説明）があった時には，ブレーキペダル，アクセルペダル，あるいはハンドルのいずれかを操作することによって運転の引き継ぎが可能です．

自動運転モード（先行車なし）　自動運転モード（先行車あり）　自動運転モード（合流車あり）

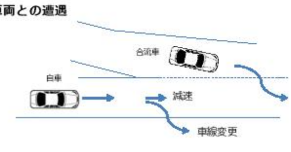
5

<u>通常運転で対応できる状況（具体例）：</u>

低速車両への遭遇



合流車両との遭遇



6

➢ 実験装置内の問題によって自動運転走行中にときどきハンドルの揺れが発生します．自動運転システムの問題ではないので，ご心配なさらないでください．

➢ また，装置内の問題によってウィンカーの点滅なしで車線変更を行います．同じく，自動運転システムの問題ではないので，ご心配なさらないでください．

➢ 以上で自動運転に関する説明は終わりです．

➢ 何か質問はありますか？

7

走行中
実施してもらうこと

ここから先，研究結果への影響を考慮し，
質問内容によってはお答えできないことがあります．
ご了承ください．

8

実験中にすること１：安全運転

➢ 高速道路を想定した環境で走行します．

➢ “実際に自分の車に乗る感覚”で 運転席に座って下さい．
自動走行中は、ペダルから足を離し，ハンドルから手を放すことができます．

9

実験中にすること２：車の制御の引き継ぎ

➢ 通常走行中，状況によっては自動走行システムが動作をしなくなる場合があります．
その時は，自分で運転を行う(車の制御を引き継ぐ)必要があります．

10

システムが運転手に車の制御引き継ぎを要請したときの表示

➢ 交通状況や天候などによって，自動では安全走行できない場合があります．

➢ ポーポーポーの音と同時に，
「自動運転解除」の表示がパネル内に点滅します
この時，あなたは直ちに車の制御をシステムから引き継いで，自分で運転を行う必要があります．

11

運転の引継ぎが
必要な場面
(具体例)

12

場面1
(例) 環境要因（濃霧，大雨等）によるセンサ不良

13

場面2
(例) 自動運転システムの故障：センサー等が機械的に壊れる

14

実験中にすること3：アンケート調査その1

実験者の教示に従って，以下の質問にお答え下さい．

質問：あなたはこの自動運転がどのように動作するかどの程度予測できますか？

全く予測できない　　どちらでもない　　完全に予測できる

質問：あなたはこれまでの経験において，この自動運転をどの程度頼りにできると思いますか？

全く頼りにできない　　どちらでもない　　完全に頼りにできる

質問：あなたはこの自動運転に対して，これまでに経験したことのない走行場面であっても安全に走行するとどの程度思いますか？

全くそう思わない　　どちらでもない　　完全にそう思う

質問：総合的に判断して，あなたはこの自動運転をどの程度信頼できますか？

全く信頼してない　　どちらでもない　　完全に信頼している

15

実験中にすること4：アンケート調査その2

実験者の教示に従って，以下の質問にお答え下さい．

質問1：先ほどのような，本線に合流してくる車が存在した場合にあなたであればどのような運転をしますか？

1．自分が減速して本線に合流してくる車両を入れる．2．自分が右車線に車線変更する

質問2：先ほどのような，本線に合流してくる車が存在した場合に，自動運転であればどのような運転をしてほしいですか？

1．自動運転が減速して本線に合流してくる車両を入れる．2．自動運転が右車線に車線変更する

16

実験中にすること4：アンケート調査その3

実験者の教示に従って，以下の質問にお答え下さい．

自動運転システムに対する私の判断は…　（チェックマークをつけて下さい）

1	(useful) 役立つ	_	_	_	_	_	役立たない (useless)
2	(pleasant) 感じが良い	_	_	_	_	_	不愉快である (unpleasant)
3	(bad) 悪い	_	_	_	_	_	良い (good)
4	(nice) 素晴らしい	_	_	_	_	_	煩わしい (annoying)
5	(effective) 効果がある	_	_	_	_	_	不適切である (superfluous)
6	(irritating) しゃくにさわる	_	_	_	_	_	好ましい (likeable)
7	(assisting) 助けになる	_	_	_	_	_	助けにならない (worthless)
8	(undesirable) 望ましくない	_	_	_	_	_	望ましい (desirable)
9	(raising alertness) 注意力を高める	_	_	_	_	_	眠気を誘う (sleep-inducing)

17

以上で説明は終わりです．

- **実験中に不安を感じたり，ご気分が悪くなった場合には，いつでも中止（走行中であっても）して頂いて結構です．**
- **その場合には，運転席から口頭でお知らせ下さい．**
- **実験を中止したことによって，ご協力者の方に不利益となることは一切ありません．**

18

APPENDIX E TRUST QUESTIONNAIRE FOR STUDY II

自動運転システムに関するアンケート

時刻：
参加者番号：#
実施回：実験説明後 / 練習走行後 / 本走行Ⅰ後 / 本走行Ⅱ後 / 本走行Ⅲ後 / 本走行Ⅳ後

1. **あなたはこの自動運転がどのように動作するかどの程度予測できますか？**

0 — 100
全く予測できない — 完全に予測できる

2. **あなたはこれまでの経験において，この自動運転をどの程度頼りにできると思いますか？**

0 — 100
全く頼りにできない — 完全に頼りにできる

3. **あなたはこの自動運転に対して，これまでに経験したことのない走行場面であっても安全に走行するとどの程度思いますか？**

0 — 100
全くそう思わない — 完全にそう思う

4. **総合的に判断して，あなたはこの自動運転をどの程度信頼できますか？**

0 — 100
全く信頼してない — 完全に信頼している

5. **先ほどのような，本線に合流してくる車が存在した場合にあなたであればどのような運転をしますか？**

1. 自分が減速して本線に合流してくる車両を入れる．　2. 自分が右車線に車線変更する

6. **先ほどのような，本線に合流してくる車が存在した場合に，自動運転であればどのような運転をしてほしいですか？**

1. 自動運転が減速して本線に合流してくる車両を入れる．　2. 自動運転が右車線に車線変更する

7. **あなたはこの自動運転を信頼していますか？**　信頼している / 信頼していない

8. **あなたは今後もこの自動運転を使いたいですか？**　使いたい / 使いたくない

自動運転システムに対する私の判断は…　（チェックマークをつけて下さい）

1	(useful) 役立つ	–	–	–	–	–	役立たない (useless)
2	(pleasant) 感じが良い	–	–	–	–	–	不愉快である (unpleasant)
3	(bad) 悪い	–	–	–	–	–	良い (good)
4	(nice) 素晴らしい	–	–	–	–	–	煩わしい (annoying)
5	(effective) 効果がある	–	–	–	–	–	不適切である (superfluous)
6	(irritating) しゃくにさわる	–	–	–	–	–	好ましい (likeable)
7	(assisting) 助けになる	–	–	–	–	–	助けにならない (worthless)
8	(undesirable) 望ましくない	–	–	–	–	–	望ましい (desirable)
9	(raising alertness) 注意力を高める	–	–	–	–	–	眠気を誘う (sleep-inducing)

APPENDIX F INSTRUCTION SHEET OF EXPERIMENTAL SESSION FOR STUDY III

本実験

これから生乳の低温殺菌工程のシミュレータの制御を行っていただきます．80回反復のシミュレーションの制御を行っていただきます．また，本セッション終了後にアンケートへの記入を宜しくお願いします．

また，プラント内の装置の問題により，トライアルごとに多少の違いがあり，手動または自動制御の効果が少なくなる場合があります．

実験中は低温殺菌された生乳の量が最大になるように留意してください
メインタンクの流量は約 500L に維持できるように留意してください

質問はありませんか．

よろしければ，本実験を開始します．

APPENDIX G TRUST QUESTIONNAIRE FOR STUDY III

アンケート

日付：
時刻：
参加者番号：
パフォーマンス：________%

1. ポンプはどの程度適切に機能を果たしていますか？

0 — 100
全く機能してない — 完全に機能している

2. ポンプの挙動はそのときどきにおいてどの程度予測できますか？

0 — 100
全く予測してない — 完全に予測できる

3. このポンプを作業遂行にあたりどの程度頼りにすることができますか？

0 — 100
全く頼りにできない — 完全に頼りにできる

4. このポンプは低温殺菌システムにおける役割をどの程度果たしていますか？

0 — 100
全く果たしてない — 完全に果たしている

5. 時刻が違っても同じような状況であれば，このポンプは同様に応答しますか？

0 — 100
全く違う — 完全に同様である

6. 将来、このポンプがこれまで経験したことのないシステム状態に対してどの程度対応できると思いますか？

0 100

全くできない 完全にできる

7. このポンプが正確に応答するとどの程度信頼していますか？

0 100

全く信頼してない 完全に信頼している

8. ポンプの表示精度をどの程度信頼していますか？

0 100

全く信頼してない 完全に信頼している

9. 全体としてあなたはこのポンプをどの程度信頼していますか？

0 100

全く信頼してない 完全に信頼している

10. あなたが好むポンプの操作方法はどちらしょうか？ 手動操作 / 自動操作

11. あなたはこのポンプを信頼していますか？ 信頼している / 信頼してない

12. あなたはこのポンプの仕組みをどの程度理解していると思いますか？

0 ―――――――――――――――――――――――― 100

全く理解できてない　　　　　　　　　　　　　　完全に理解している

13. あなたはこの熱交換システムの仕組みをどの程度理解していると思いますか？

0 ―――――――――――――――――――――――― 100

全く理解できてない　　　　　　　　　　　　　　完全に理解している

14. あなたは，あなたのポンプ操作能力にどの程度自信を持っていますか？

0 ―――――――――――――――――――――――― 100

全く自信がない　　　　　　　　　　　　　　完全に自信がある

15. あなたは，あなたの熱交換システム操作能力にどの程度自信を持っていますか？

0 ―――――――――――――――――――――――― 100

全く自信がない　　　　　　　　　　　　　　完全に自信がある

□　もしあなたがポンプの仕組みが理解できたと思いましたら，チェックを入れてください.

□　もしあなたが熱交換システムの仕組みが理解できたと思いましたら，チェックを入れてください.

Questionnaire for first impression

- あなたの「ポンプ」に対するイメージは？
  ポンプそのもの / 自動ポンプ / このプラント（PASTEURIZER）自体
- このポンプを完全に理解する自信がありますか？
  自信ある / 自信ない
- 自動ポンプを使ったとき，手動操作より高いパフォーマンスが達成できると思いますか？
  はい / いいえ
- これからも自動ポンプに頼りたいと思いますか？
  はい / いいえ
- これからも自動ポンプに頼りたいと思う理由は何ですか？
  高いパフォーマンス達成 / 手動操作より便利だから / その他
- 高いパフォーマンスを達成できるなら，ポンプの仕組みが理解できなくてもポンプを信頼できますか？
  はい / いいえ
- 自動ポンプが手動操作より高いパフォーマンスを達成できても，自動ポンプの仕組みが完璧に理解できないとポンプを信頼できないと思いますか？
  はい / いいえ
- 以下の中，あなたのポンプに対する信頼感により影響するのは？（順位を決めてください）
  パフォーマンス / システムへの理解 / その他

Questionnaire about system failures during experimental sessions

- あなたは，このプラントのポンプに問題があったと思いますか？　　はい / いいえ

  もし，あると思いましたら問題があったと思われる項目にチェックを入れてください（複数応答可）.
  - □ ポンプの表示精度
  - □ ポンプの応答精度

  また，チェックを入れた項目の問題でどのような現象が確認できましたか？

本実験終了後アンケート (Final)

参加者番号：

- あなたは，パスチャライザー（シミュレーション）は自動化システムだと思いますか？

そう思う / そう思わない

その理由は？

- あなたは，ポンプシステム自動化システムだと思いますか？

そう思う / そう思わない

その理由は？

- あなたは，アクティブ熱交換システムは自動化システムだと思いますか？

そう思う / そう思わない

その理由は？

- ポンプの属性の中，どこに問題が生じたときがあなたにとってもっと致命的ですか？

ポンプ表示精度 / ポンプ応答精度

その理由は？

- パスチャライザー制御の難易度をお答えください.

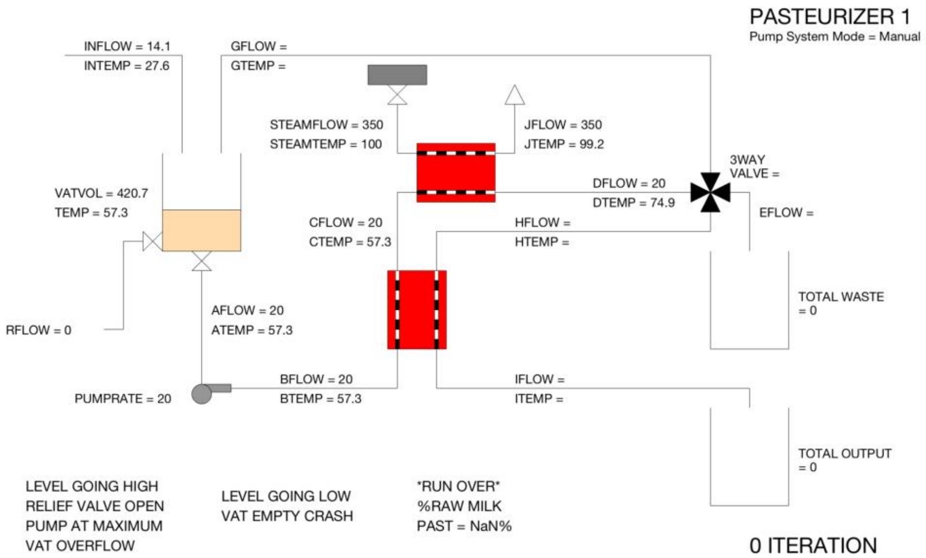
難しくない / 難しい

- パスチャライザーに含まれているエラーを気づきましたか？

気づいた / 気づけなかった

- 気づいた場合，どこにどのようなエラーが含まれているか理解できたと思いますか？

そう思う（完全に理解できた） / そう思わない（あまり区別できなかった）



- エラーによるポンプシステムの上がり・下がりは割合で設定されています.
- 上記のデフォルト値が 20 ではなく 80 であったら，どの属性にどのようなエラーが含まれているかわかりやすかったと思いますか？

そう思う / そう思わない（やってみないと分からない）