THE INFLUENCE OF RISK IN TRUST AND AUTOMATION

by

Kelly Satterfield
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Psychology

Committee:

_______________________________________ Director

____

____

_______________________________________ Department Chairperson

_______________________________________ Program Director

_______________________________________ Dean, College of Humanities and Social Sciences

Date: ___________________________________ Fall Semester 2016
George Mason University
Fairfax, VA

The Influence of Risk in Trust and Automation

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Kelly Satterfield
Master of Arts
George Mason University, 2013
Bachelor of Science
University of Dayton, 2009

Director: Tyler Shaw, Professor
Department of Psychology

Fall Semester 2016
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Tyler Shaw for his guidance, expertise, and mentorship on this project, as well throughout my graduate career. I'd also like to thank the other members of my committee, Dr. Carryl Baldwin, and Dr. Ewart de Visser for their invaluable help. And finally I'd also like to thank my husband, Michael, for supporting me through this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

THE INFLUENCE OF RISK IN TRUST AND AUTOMATION

Kelly Satterfield, Ph.D.

George Mason University, 2016

Dissertation Director: Dr. Tyler Shaw

Advances in technology have led to increases in the implementation of automated systems. In order for the full range of benefits of automation to be utilized, an understanding of how operators trust an automated system is vital. The level of risk in an environment is an important factor that might affect trust and one that has not been studied extensively. This dissertation aims to explore the effect differing levels of risk have on trust in an automated system. The first study examined whether trust in an autonomous teammate differs between situations with low and high risk. A second study determined if experience with the automation interacts with level of risk. Results support the hypothesis that trust decreases with experience of automation failures in training. However, findings were mixed regarding the hypothesis that trust decreases with increased risk. These experiments provide evidence that situational factors are important for training to combat complacency, as well as for appropriate calibration of trust in automation.

INTRODUCTION

Trust across Domains

Trust has been shown to be an important variable in many operational domains including sociology, economics, organizational management, finance, and e-commerce. For example, considerable research has been conducted examining how organizations trust each other in joint business ventures (Gill & Butler, 1996; Inkpen & Currall, 1998; McKnight, Cummings, & Chervany, 1998; Wicks, Berman, & Jones, 1999). This is especially true in multi-national organizations in which cultures differ and cross-cultural collaboration is still necessary (Doney, Cannon, & Mullen, 1998). Trust has also been found to be important in relationships between supervisors and employees (Tan & Tan, 2000), in increasing organizational productivity (Nyhan, 2000), and firms and customers alike have identified the value of trust in relationship management (Morgan & Hunt, 1994).

While each domain may put a slightly different emphasis on why trust is necessary and how it is formed, many agree that trust formation is a result of weighing benefits and consequences. As an example, economists highlight the calculation of costs and benefits in trust formation (Williamson, 1993). Within the finance domain, much research has been conducted on trust between lenders and borrowers and the emphasis has been that trust is formed from risk assessment and the potential of long-term benefits (Guth, 2002; Ferrary, 2003). Similarly, social psychologists have defined trust as based on rational costs and benefits of others' behavior (Lewicki & Bunker, 1995). Thus, it appears that the overall consensus is that a major factor in trust formation is the weighing the relative costs and benefits.

Aside from a consideration of costs and benefits, marketers have proposed that prior reputation and satisfaction are important concepts in trust formation, but may differ depending on the participant. For example, retailers may value a vendor's reputation more in trust formation, while vendors may rely on the satisfaction of the interaction with the retailer as in indicator of how much trust should be given (Ganesan, 1994). It is clear that some research areas have emphasized cost and benefit assessment and others have emphasized reputation, but what is key is the idea that trust is an important variable in relationships.

The purpose of this dissertation is to examine the influence of risk, or changing stakes, on an operator's willingness to trust an autonomous teammate. The first section of this introduction discusses trust in automation, the second section discusses factors the influence trust in automation, and the third section discusses the empirical evidence involving risk and trust.

Importance of Trust in Automation

The aforementioned research on trust in a variety of domains points to its relevance, particularly in human-human relationships. However, recently this research has been applied to another domain: automation and autonomous systems. Automated systems are prevalent in the world today in a wide range of fields including aviation, medicine, and driving. Automation is defined as a computer or machine that performs partially or fully a task previously performed by a human (Parasuraman & Riley, 1997; Parasuraman, Sheridan, & Wickens, 2000). Properly designed automation can offset the workload of an operator and improve safety, and it can be tempting to implement full automation whenever possible. However, there can be drawbacks to automation such as decreased situational awareness, skill loss, and complacency (Parasuraman et al., 2000). Therefore, it is important to fully examine the costs and benefits of the automated system before implementation. And in order to fully understand these costs and benefits, it is necessary to see how operators trust an automated system since this will have a strong effect on how effectively the automation is utilized.

Automation is often implemented to offset workload in environments in which task demands are high. Complex environments are often unstable and full of uncertainty and in order for individuals to adapt to this complexity, trust becomes important (Lee & See, 2004). The complexity of automation often makes a complete understanding of the system impractical, and operators must rely on the automation to perform their job effectively.

However, it is important that humans appropriately calibrate their trust to an automated system. Lee and Moray (1992) have shown that an operator's use of automation to control a simulated manufacturing plant was directly related to momentary trust in the automation, which in turn affected the frequency of errors. Hence, operators may trust automation even when it is not beneficial or the automation is prone to malfunctions. This extreme misuse of automation leads to operators relying on automation without monitoring its behavior or its limitations (Parasuraman & Riley, 1997). Blindly following automation when it is not perfectly reliable can lead to overreliance and failure to catch errors. This complacent behavior has been shown to be an effect of allocating attention away from monitoring automation and may result from high levels of trust in the automated system (Bagheri & Jamieson, 2004; Bailey & Scerbo, 2007; Parasuraman et al., 1993). On the other hand, operators may not trust automation even when it performs consistently and offloads workload. Operators are often slow to trust new technology and may dislike or mistrust a new automated system. However, complete disuse of automation, such as turning off alarms or disabling alerting systems can lead to serious accidents (Parasuraman & Riley, 1997). Because of these findings, it is imperative that trust in automation be examined to facilitate appropriate use of these systems—that is, to appropriately calibrate trust.

Defining Trust

While much research has stressed the value of trust in relationship formation, the diverse nature of this research has led to many differing definitions of trust. The first definitions of trust came from examining human-human relationships. These early researchers define trust as being associated with expectancy—that is, the expectancy that another person or organization will act in a certain manner (Rotter, 1967; Barber, 1983). High trust, then, would be associated with a high expectancy of a person who acts favorably. This expectancy may be a result of past experience or specific communication with the other person. Other researchers adopting the expectancy approach to trust suggest that trust is based on the necessity to rely on someone (or something) to perform as expected, but also add that trust includes the willingness to be vulnerable (Johns, 1996, Moorman et al., 1993). For trust to occur a certain behavior is expected, but failure of that behavior to occur results in some consequence.

Other researchers such as Deutsch (1958), a social psychologist, have proposed that trust is not just based on expectancy or predictability of a certain outcome, nor is it just based simply on the fact that a consequence is possible. Instead, trust is also dependent on the motivational consequences of the expected outcome. The weight of gains and losses is also taken into account, as well as the probability of that gain or loss.

Building on the preliminary ideas proposed by Deutsch (1958), the most widely used definition of trust was proposed by Mayer et al. (1995) stating that trust must involve an element of risk. Specifically, trust is not just the willingness to be vulnerable, but must include a behavior indicating a willingness to be vulnerable. Trust is the actual assuming of risk. However, not all risk-taking includes trust. It is specifically risk-taking behavior that includes a reliance or vulnerability on another party.

These definitions of trust demonstrate that there are inconsistencies in whether trust is a belief, attitude, or behavior. Ajzen and Fishbein (1980) provide a framework for reconciling these conflicting definitions. They propose that behaviors result from intentions which in turn are a function of attitudes that are based on beliefs. Therefore, within this framework, trust is an attitude and reliance is a behavior. Beliefs and perceptions underlie trust, which in turn affect a behavior, such as reliance on a system. (Lee & See, 2004). The beliefs about the characteristics of the automation motivate trust which then affects the reliance on the automation.

Rempel et al. (1985) suggested a way of categorizing the beliefs that motivate trust. Rempel et al. (1985) proposed that trust can be divided into three dimensions: predictability, dependability, and faith. Predictability is influenced by the consistency of behavior and a stable environment. This dimension is concerned with how stable performance is over time. The second dimension, dependability, involves dispositional characteristics of a person. Dependability refers to the reliability of the person. And the third dimension, faith, refers to the beliefs about the future behavior of the person. While the model of Rempel et al. (1985) defines trust in a human-human context, Lee and Moray (1992) have also identified three general dimensions of trust in the context of automation: performance, process, and purpose. Performance refers to the current and past performance of the automation and including its reliability, predictability, and ability. Performance concerns whether the automation is completing the goals it was intended to complete. Process refers to how the automation operators and whether the specific algorithms complete the automation's goals. Purpose refers to whether the automation is being used within the designer's intent. Performance and purpose can be thought to correspond to the dimensions of predictability and dependability in the Rempel et al. (1985) model, while purpose roughly corresponds to faith.

Distinguishing between human-human trust and human-automation trust

Important considerations that need to be taken into account are the similarities and differences between human-human and human-machine trust. With regard to similarities, it has been shown that people apply social rules, such as politeness, to machines (Nass, Steur, & Tauber, 1994). Also, neurological research has pointed to the fact that some of the same neural mechanisms are utilized in interpersonal trust games as well as in e-commerce interactions with real people (Dimoka, 2010; Riedl, Hubert, & Kenning, 2010). And while trust research in human-human interaction has emphasized the importance of intentionality, automated systems lack intentionality. However, automation is designed with a purpose and does reflect the intentions of the designers (Rasmussen, Pejterson, & Goodstein, 1994). Furthermore, automation can be given human characteristics and the anthropomorphism of automation may lead people to attribute intentionality (Nass & Lee, 2001; Nass & Moon, 2000).

While there are similarities in how trust is conceptualized in human-human interactions compared to human-automation interactions, differences exist in how human-automation trust is formed (Madhavan & Wiegmann, 2007). Rempel et al. (1985) explain that interpersonal trust is initially formed on the prediction of the other person's actions. Human-automation trust usually develops in the opposite direction. Research has shown that people often initially place high trust in automated systems (Dzindeolet et al., 2003). Research has also shown differences in how people assign responsibility in human-human interactions versus human-automation interactions. People perceive the ultimate responsibility in a human-human interaction to be shared, while as the responsibility in a human-automation relationship lies with the operator (Lewandowsky, 2001). Because of these differences, research is still needed that considers the unique elements of human-automation trust (Lee & See, 2004).

Factors That Influence Trust in Automation

While several models of trust exist, perhaps one of the most comprehensive has been advanced by Hoff & Bashir (2015). They have developed a model which attempts to describe the factors affecting trust in automation. Hoff and Bashir (2015) propose that the human operator, the environment, and the automated system account for the variability in human-automation trust. These three sources then map onto three different layers of trust: dispositional, learned, and situational trust. Dispositional trust reflects an individual operator’s tendency to trust automation and is influenced by factors such as culture, age, gender, and personality traits. Learned trust is dependent on past experiences with the automated system. Situational trust is influenced by the environment as well as the context-specific mental state of the operator.

Dispositional trust

Dispositional trust is an enduring trait and is stable over time. It involves long-term tendencies that are influenced from both biological and environmental factors. Age is one such biological factor. Research has shown that older adults rely on decision aids more than younger adults and do not adjust their trust after automation errors (Ho, Wheatley, & Scialfa, 2005). In contrast, Sanchez, Fisk and Rogers (2004) found that older adults were better at calibrating their trust with changing automation reliability. Both studies suggest that age has an effect on how people employ strategies when interacting with automation, thus suggesting that individual differences play a role in a person’s decision to trust. Culture is another dispositional factor that has been shown to affect trust in automation. Several studies have identified differences in how different cultures perceive social robots (Li, Rau, & Li, 2010); Rau, Li, & Li, 2009). For example, cultures defined as “low-context”, meaning these cultures prefer information be explicitly stated in a message, tend to be more engaged with a more social robot (Li, Rau, & Li, 2010). Stable personality traits have also been shown to affect trust in automation. Merritt and Ilgen (2008) showed that extroverts exhibit a greater propensity to trust compared to introverts.

Learned trust

The Hoff and Bashir model also suggest that past experiences strongly influence how people trust automated systems. This learned trust is influenced by the automation’s past performance, as well as the operator’s pre-existing knowledge of the system. According to Hoff and Bashir (2015) learned trust can then further be divided into initial learned trust and dynamic learned trust. Initial learned trust refers to trust prior to interacting with a system while dynamic learned trust refers to trust during an interaction. This distinction is made to reflect how trust can fluctuate with the system’s real-time performance.

Pre-existing knowledge is one factor that influences initial learned trust. It has been found that people trust automation more when it is portrayed as an expert with a good reputation (Madhavan & Weigmann, 2007). However, even if the system is reputable, dynamic learned trust can also be affected as trust may degrade faster when systems make obvious errors (Madhavan & Weigmann, 2005). It is also important to distinguish subject matter expertise, which is a situational trust factor, from past experience, which is a learned trust factor. Yuviler -Gavish and Gopher (2011) found that participants relied on a decision aid more after gaining more experience with the system.

So it would seem trust increases with more experience. However, Sanchez et al. (2014) found that experienced farmers relied on automation less in a collision avoidance task involving agricultural equipment, which would seem to contradict the findings of Yuviler -Gavish and Gopher (2011). But it's worth noting that the participants in Sanchez et al. (2014), while experienced farmers, had no experience with the particular automation used in the experiment. Therefore, Sanchez et al. (2014) should be classified as a study examining situational factors instead of learned factors. While some research has shown that experience with an automated system results in increased trust, there have been some studies demonstrating that experience with an unreliable system may lead to a decrease in trust (e.g. Manzey et al., 2012). Specifically, experience with automation failures leads to a stronger decrease in trust than positive experiences with a properly working aid leads to increases in trust.

Situational Trust

There are two broad sources of variability that affect situational trust: the external environment and the internal context-specific characteristics of the operator. Operator characteristics such as self-confidence, subject matter expertise, mood, and attentional capacity affect internal variability. Environmental factors like the system complexity, task difficulty, workload, as well as perceived risks and benefits all affect an external variability in trust.

Internal contextual characteristics of the operator present while performing the task influence situational trust. This is not to be confused with dispositional characteristics, which are stable over time. Instead, these characteristics may be transient and fluctuate depending on the task and environment. Self-confidence has been shown to affect how much trust operators place in an automated aid. Studies have found that when self-confidence is low, participants are more likely to use automation. In contrast, when self-confidence is high, participants trust in the automation less (de Vris, Midden, & Bouwhuis, 2003; Lee & Moray, 1994). When operators feel they are equipped to perform the task, they may prefer to take manual control. Additionally, mood is another temporary characteristic that may influence an operator's trust in an automated system. It has been found that positive moods lead to higher levels of trust in automation (Stokes et al., 2010; Merritt, 2011).

External environmental factors such as the system's complexity and reliability also affect trust in automation (Bailey & Scerbo, 2007). Increased task difficulty and workload affect the amount of time an operator can monitor automation and this in turn affects how much trust operators place in automation. For example, Biros et al. (2004) showed that with increased workload, trust and reliance on automation increased. This may be due to the fact that under a high workload, operators must rely on the automation more to accomplish the task within existing time constraints.

Risk and Trust

According to the model proposed by Hoff and Bashir (2014), perceived risk is another important situational trust factor. An environment always involves some sort of uncertainty and since trust relates to uncertainty, perceived risks and benefits in an environment exert a strong influence over trust in an automated system. However, to date, the effect of risk on trust in automation has not been thoroughly studied. Only two studies have directly looked at the effect of risk on trust in automation to mixed results. Perkins, Miller, Hashemi, and Burns (2010) investigated whether trust in a GPS navigation system would be affected by different levels of risk defined by an increase in the presence of hazards along the route. Results demonstrated that participants trusted the GPS less when more hazards were present. In contrast, participants in Lyons and Stokes' study (2011) relied on a human aid less compared to an automated aid in a high risk condition characterized by the probability of an attack along the suggested route, demonstrating an inclination toward automation with increased risk.

This type of research comes with its own set of unique challenges. It is difficult to manipulate risk in the laboratory environment and guarantee that the participants are fully engaged and have something at stake. Perhaps this is why there is such a dearth of research in this area. However, while the effect of risk on trust in automation has not been explored extensively, the economics literature has examined the effect of manipulating the stakes involved with trusting a partner. Trust games, such as the ultimatum game, require participants to trust a partner when exchanging money. These trust games involve two participants who are given a sum of money to begin. The first participant decides a portion of the money to send to the second participant. The second participant then simply accepts or rejects the proposed portion. If he or she accepts, both participants are paid, if he or she rejects neither participant is paid (Guth, Schmittberger, & Schwarze, 1982). In other variations, the second participant may then either choose to keep the money passed or choose to send a portion back to the first participant. The amount of money passed by each participant is thought to reflect the trust the participant has in his or her partner (Berg, Dickhaut, & McCabe, 1995; Camerer, 2003).

Johnson and Meslin (2011) provide a detailed meta-analysis of trust games and have identified the amount of stakes in the experiment as a variable of interest. By increasing the main sum participants are given at the beginning, researchers have examined whether raising the stakes affects how much trust participants place in their partner. Hoffman, McCabe, and Smith (1996) raised the beginning sum to $100 compared to the previously used $10 and found no difference in the proportion of money the first participant sent to the second participant. Similarly, other studies have found that increasing the stakes involved with trust games does not affect behavior (Forsythe et al., 1994; Carpenter et al., 2005). There is no difference in the amount of money passed from the first participant to the second participant.

In contrast, Johansson-Stenman, Mahmud, and Martinsson (2005) did find that the proportion of money sent decreased significantly with increased stake size. It should be noted, however, that even with large stakes, some participants sent nothing. These findings led Johannson-Stenman, Mahmud, and Martinsson (2005) to conclude that the first participant's decision may instead reflect risk preferences, rather than trust. Participants become more risk aversive with increased stakes but this does not reflect a difference in trust behavior. This idea is consistent with Holt and Laury's findings (2002) that participants become more risk averse with increased stakes in a lottery choice experiment. Similarly, Karlan (2005) argues that the first participant's behavior reflect propensity for risk and the second participant's behavior reflects trust.

To see if there were differing behaviors between first and second participants with increased stakes, Cameron (1999) also investigated the effect raising the beginning sum to $100 has on trust in the ultimatum game, but in an area in Indonesia where $100 has the equivalent of raising the stakes proportional to income much higher than what would be possible in the United States. This study found no difference in the proportion of money the first participant sent to the second participant. However, by also analyzing the amount of money the second participant was willing to accept, results demonstrated that with raised stakes, the second participant was more willing to accept lower proportions of money. Similarly, Munier and Zaharia (2002) found that with increased stakes, the lowest acceptable offers by the second participant were lower in the high stakes condition. So while Cameron (1999) and Munier and Saharia (2002) did not find a difference in the initial offer by the first participant with increased stakes, they did find different behavior by the second participant.

While most studies involving trust games have focused on trust between two people, some research has focused on examining the difference in trust in these games between people and computers. Sanfey et al. (2003) found that when differentiating between a computer and a person as the first player, participants rejected more low offers from a person compared to a computer. Results were explained as a function of punishment for unfairness. Within the game, any offer above $0 is a benefit to the second player, but often second players will reject low offers as a form of punishment for suggesting an unfair split. Interestingly, Sanfey et al. (2003) demonstrated that participants will reject low offers and punish a person, more often than a computer. While humans were punished more than computers, participants did punish computers slightly for low offers, indicating that humans could be attributing agency to the computer. Other research has demonstrated that participants may send less money in the trust game if the second player is a computer (Bottom et al., 2006). This line of research in economic trust games has relevance to the influence risk may have on trust in automation.

Study Rationale and Hypotheses

According to Hoff and Bashir (2014), risk is a situational factor that has an influence on trust in automation, but as mentioned above, few studies have empirically examined this relation. But since the environment largely defines the risks and benefits associated with automation, it is an important situational factor to examine. A study of this sort would have important implications for how operators trust automated systems in high stakes environments where consequences can be financially costly and may include loss of life.

Therefore, the aim of this dissertation is to investigate the effect differing levels of risk on trust in automation. However, in contrast to past studies, this dissertation will involve a tangible risk factor during the experiment. Furthermore, past studies looking at the effect risk has on trust in automation have relied on subjective measures of trust that come with certain disadvantages. Subjective questionnaires only provide trust information after the experiment is finished and are reliant on the ability of the participant to have self-awareness of their own behavior. In the current set of experiments, a behavioral measure of trust was utilized that had been evaluated through extensive pilot testing in addition to subjective measures.

This dissertation focuses on three research questions:

1. What effect do differing levels of risk have on trust in an automated system? (Experiment 1)
2. Does having prior experience with automation failures in training exacerbate the effect risk level has on trust? (Experiment 2)
3. Can we validate a behavioral measure of trust—specifically the degree to which humans intervene during autonomous functions (Experiment 1 and 2)

Previous Pilot Testing

Pilot testing was conducted in order to begin the validation of a behavioral measure of trust. Currently, the standard measure of trust in automation is through subjective questionnaires. However, these measures are sensitive to biases, assume participants are aware of what is driving their actions, and are only collected after the scenario is completed. Instead, a behavioral measure of trust can potentially provide continuous trust information throughout the entire scenario.

During pilot testing, 20 participants performed a 7 minute Distributed Dynamic Decision Making (DDD) scenario, the task that was also used for this dissertation. A point system was implemented and behavioral trust was defined by how many times the participant left their area of responsibility (AOR) and destroyed an enemy in their teammates AOR. In this case, the team member was fully autonomous (pre-programmed) and was responsible for protecting the green areas' red zone. Intervening in your teammate's AOR indicates a failure to trust the autonomous team member to destroy enemies on its own.

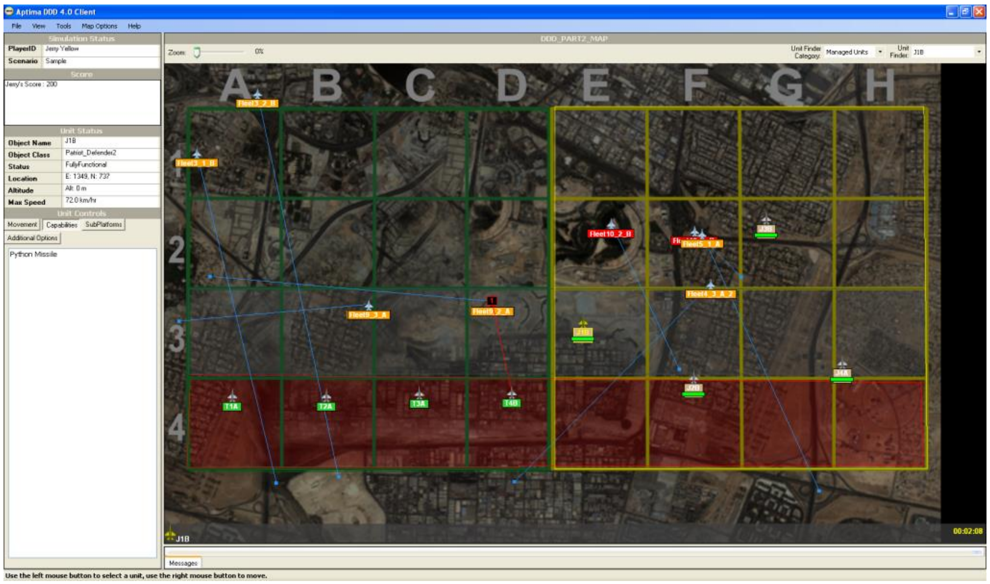

Figure 1: DDD simulation. Participant's AOR was the yellow grid. Automation's AOR was the green grid.

Results found variability in participants' tendency to intervene. Furthermore, this measure of behavioral trust was correlated with more established subjective measures of trust (Merritt, 2011). Results showed a significant negative correlation between subjective trust and the number of enemies destroyed in the green zone, r = -.49, $p < .05$. Based on these results, it was determined that the behavioral measure of trust would be included in the two experiments comprising this dissertation.

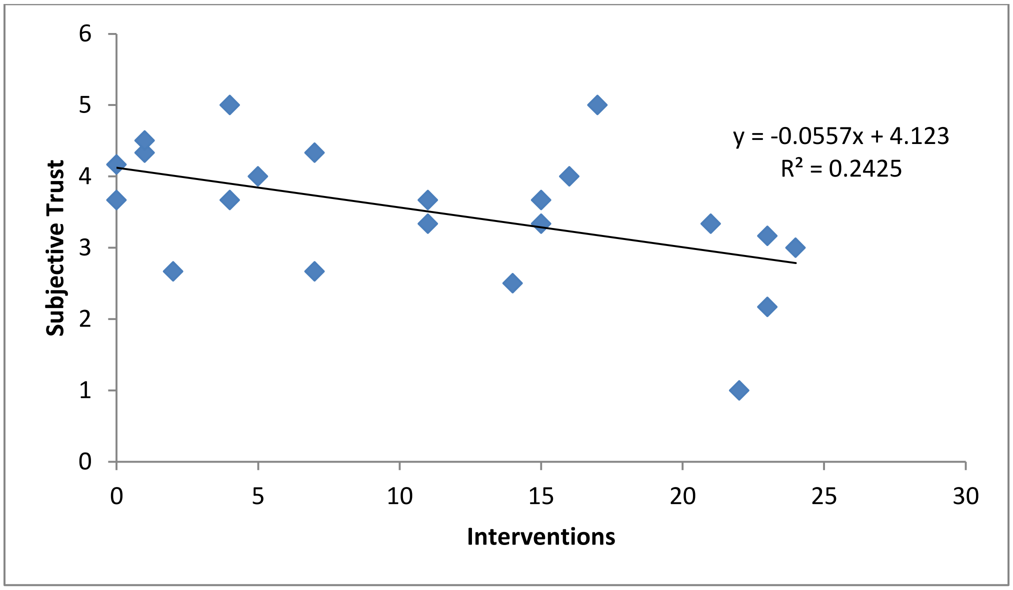

Figure 2: Relationship between subjective trust on the Self-Report Scale of Trust and the number of interventions in the automation's AOR

EXPERIMENT 1: METHOD

Design

The design of Experiment 1 was a between subjects design with the level of risk as the between subjects factor. The two levels of risk (low, high) were defined by the amount of money a participant stood to lose in the scenario.

Risk

Two levels of risk were distinguished by the amount of money a participant would lose for not obtaining a certain amount of points. All participants were given an amount of $50.00 before the scenario. Participants in the low risk condition stood to lose $10.00 for scoring below 1225 points during the DDD scenario. Participants in the high risk condition stood to lose $40.00 for scoring below 1225 points during the DDD scenario. These parameters were based on a study conducted by Harinck et al. (2007) which found that participants will feel loss aversion if the payout is at least $40.00. The point value was chosen from previous pilot testing.

Apparatus

The Dynamic Distributed Decision Making (DDD) 4.0 simulation developed by Aptima Inc. was used in this experiment. DDD is a tool for creating human-in-the-loop, distributed, multi-person and automation based scenarios. Participants performed the DDD scenario using a desktop computer and mouse.

The scenario was a seven minute simulated counter air operation in which enemy targets entered and immediately began moving towards a no-fly zone (red zone in this case). Enemy forces had the ability to attack and destroy all friendly UAV fighter assets. Participants controlled four UAV assets and an autonomous team member controlled four UAV assets. The participant and autonomous team member had separate, distinct zones for which they were responsible for protecting. The participant was responsible for protecting the yellow zone on the right and the autonomous team member was responsible for protecting the green zone on the left.

Participants were instructed that the goal of the scenario was to gain as many points as possible. A point system was explained as follows:

1. Participants lost 15 points for each orange enemy that entered the red zone.
2. Participants lost 30 points for each red enemy that entered the red zone.
3. Participants gained 50 points for destroying a red enemy while in the yellow zone. Participants gained 50 points when the automation destroys a red enemy in the green zone.
4. Participants gained 30 points for destroying a red enemy in the green zone.
5. Participants gained 25 points for destroying an orange enemy while in the yellow zone.
6. Participants gained 25 points when the automation destroys an orange enemy in the green zone.
7. Participants gained 15 points for destroying an orange enemy in the green zone.

This point system encouraged participants to make judgments about whether to intervene and destroy enemies in the autonomous team member's zone. Participants would lose points if the autonomous team member allowed an enemy to enter the red zone, but would gain more points if the autonomous team member destroyed an enemy compared to his or her self. There were a total of 48 enemies, 32 orange enemies, and 16 red enemies. The maximum number of points a participant can score is 1470 points. These parameters were utilized in pilot testing and were shown to result in variability in participants' preference to intervene even with no risk manipulation. The optimal strategy that would result in the highest number of points is to only destroy the enemies that the automation would fail to destroy. Since participants did not know which enemies the automation would fail to destroy, the next best strategy is to completely rely on the automation and not intervene in the automation's zone of responsibility. This strategy would result in more points than always intervening.

Subjective Questionnaires

The Self Report Scale of Trust which was developed by Merritt (2011) was used to measure trust in automated systems. It involves Likert ratings on six items. Sample questions include "I believe the automation is a competent performer" and "I can depend on the automation". The Checklist for Trust developed by Jian et al. (2000) was used to measure trust between people and automated systems. It involves Likert ratings to ten items. Sample questions include "The system is deceptive" and "I am wary of the system". The Propensity to Trust questionnaire developed by Merritt (2011) was used to measure dispositional trust in automated systems. It involves Likert ratings on six items. Sample questions include "I usually trust machines until there is a reason not to" and "For the most part I distrust machines".

Individual Differences

A number of individual difference measures will be used. The Balloon Analogue Risk Task (BART) as a measure of risk-taking (Lejuez et al., 2002). It is a computerized measure of real-world risk behavior that balances potential rewards versus losses. In the task, a participant is presented with a balloon and the chance to earn money by pumping the balloon up with a mouse click. Each mouse click inflates the balloon and the participant earns $0.05 with each pump. However, as some threshold, unknown to the participant, the balloon will overinflate and explode. At this point, participants will not earn the money from that particular trial. A participant can choose to cash-out before the balloon explodes and will collect the money earned for that trial up to that point. Therefore, each pump results in greater risk, but also has a greater potential reward.

The Complacency Potential Scale is a 20-item scale used to measure attitudes toward automated devices that reflect a potential for complacency (Singh, Molloy, & Parasuraman, 1993). Questions map onto one of four factors that form components of automation-induced complacency: Confidence-related, Reliance-related, Trust-related, and Safety-related Complacency.

The Mini-IPIP is a 20-item scale will be used to measure the Big Five Factors of Personality (Goldberg, 1999). This questionnaire includes four questions for each Big Five trait and has been shown to provide a reliable and valid measure of personality (Donnellan et al., 2006).

The Computer and Web Attitude Scale (Liaw, 2002) is a 16-item scale used to measure attitudes towards computers and technology. The National Technology Readiness Survey (Parasuraman, A., 2000) is a 36-item scale will be used to measure participant's readiness to accept new technologies.

Procedure

30 participants (F = 20, *M*= 23.19) participated in the experiment. Participants first answered the Merritt (2011) Propensity to Trust Questionnaire. They then completed the Balloon Analogue Risk Task (BART, Lejuez et al., 2002) to measure risk-taking behavior. Following the BART, participants viewed a self-paced PowerPoint presentation explaining DDD, the goals of the game, and how to specifically play the game. After viewing the PowerPoint, participants completed a five minute practice scenario to make sure they were competent in how to generally play DDD. Participants needed to demonstrate that they could move throughout the scenario as well as attack enemies. Participants then completed three practice scenarios that were seven minutes in length and involved the same parameters of the experiment. During this practice, the automation was 70% reliable. Participants were told that the automation may or may not be perfectly reliable. Following these practice scenarios, participants completed the seven minute experimental trial. During the experimental trial, the automation was also 70% reliable. 15 participants experienced the low risk condition and 15 participants experienced the high risk condition.

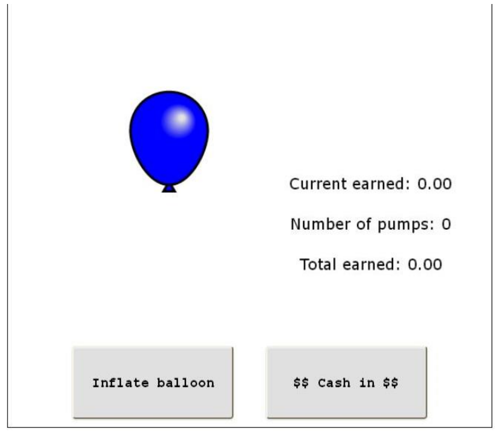

Figure 3: Balloon Analogue Risk Task (BART)

Following the experimental trial, participants completed two subjective trust questionnaires (Merritt, 2011; Jian et al., 2000), followed by a battery of individual differences questionnaires. Individual differences questionnaires included the Complacency Potential Scale (Singh, Molly, & Parasuraman, 1993), Mini-IPIP (Donnellan et al., 2006), Computer Attitudes Scale (Liaw, 2002), and the Technology Readiness Questionnaire (Parasuraman, A., 2000). This concluded the experimental portion of the session and participants were debriefed on the experiment and paid for their participation. The experiment lasted approximately an hour and a half.

It was hypothesized that subjective trust will be lower in the high risk condition compared to the low risk condition, similar to the results of Perkins et al. (2010).

Similarly, behavioral trust will be lower in the high risk condition compared to the low risk condition. Specifically, participants will destroy more enemies in the automation's zone in the high risk condition compared to the low risk condition.

EXPERIMENT 1: RESULTS

Performance

The percentage of enemy incursions into the red-zone was used to characterize operator performance, thus, lower numbers indicate better performance. An independent-samples t-test was conducted on the percentage of enemies allowed into the red-zone. The result of this test revealed a significant difference in the percentage of incursions between the High Risk condition and the Low Risk condition, $t(28) = -2.20$, $p < .05$, $d = .81$. The percentage of incursions was higher in the Low Risk condition (M = 11.1%, SD = 8.5%) compared to the High Risk condition (M = 5.3%, SD = 5.6%).

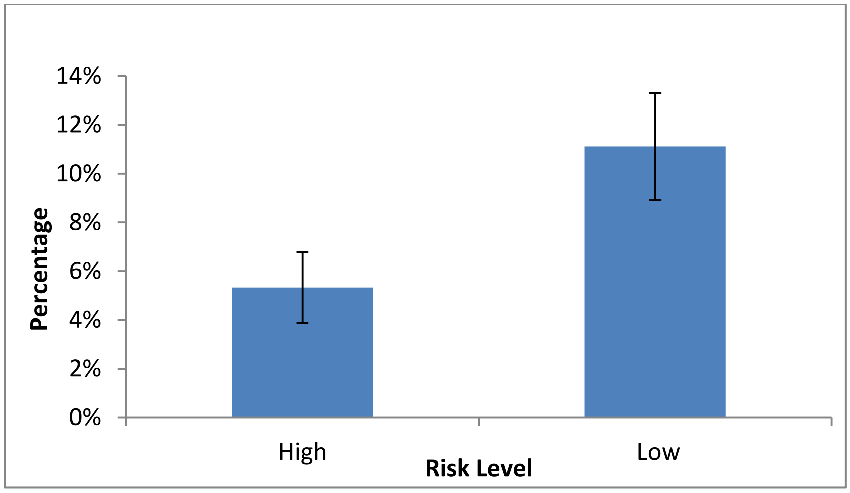

Figure 4: Percent Incursions

The percentage of incursions into the red zone on only the participant's zone of responsibility was also examined. An independent samples t-test revealed no significant difference between the Low (*M* = 4.3%, *SD* = 4.7%) and High risk (*M* = 2.4%, *SD* = 3.9%) conditions, $t(28) = -1.19$, $p = .25$.

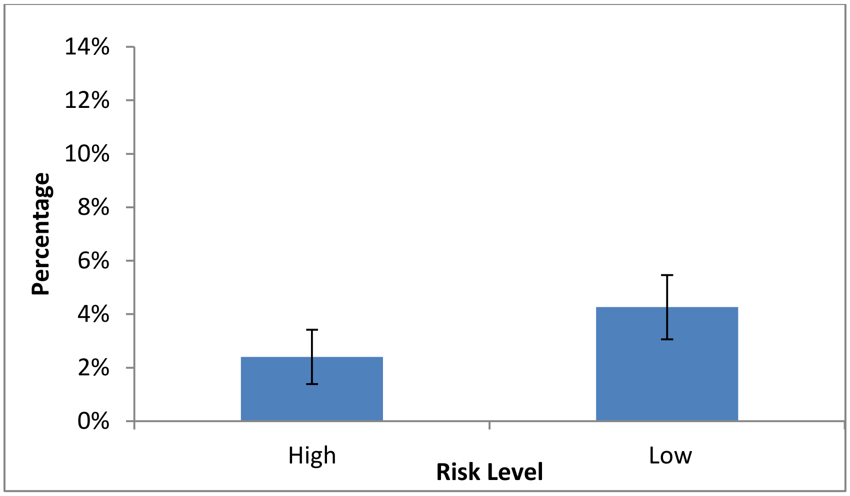

Figure 5: Percentage of Incursions in Participant's Zone of Responsibility

An independent samples t-test was also conducted to examine the difference in score between the two risk conditions. Results of this test revealed no significant difference in the score between the High risk (*M* = 1285.33, *SD* = 77.3) and theLow risk (*M* = 1267.33, *SD* = 74.2) conditions, $t(28) = .65$, $p = .52$.

Behavioral Trust

The number of enemies a participant killed in the automation's green zone was used to characterize behavioral trust (as indicated by results from the pilot data). Therefore, a high percentage of green zone interventions indicates low trust. An independent-samples t-test was conducted on the percentage of interventions in the green zone. Results indicated a significant difference in the percentage of interventions in the High Risk condition compared to the Low Risk condition, $t(28) = 2.183$, $p < .05$, $d = .80$.

Participants intervened at a higher percentage in the High Risk condition (*M* = 73.6%, *SD* = 30.6%) compared to the Low Risk condition (*M* = 47.7%, *SD* = 34.2%).

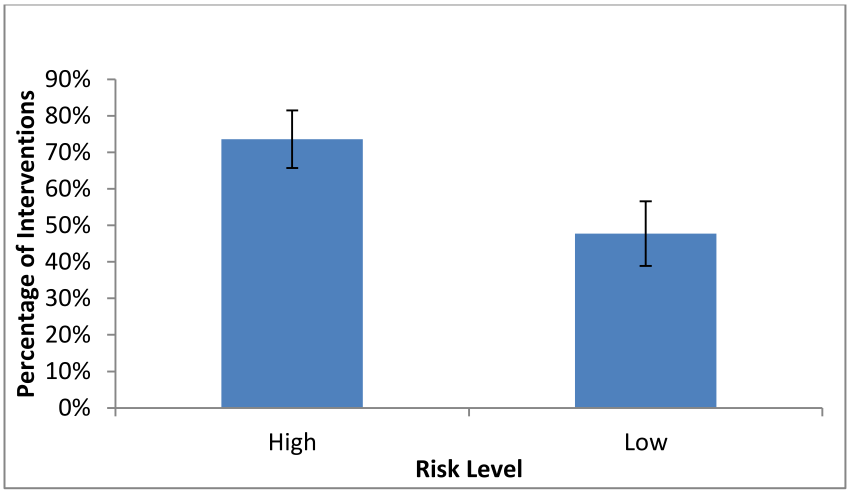

Figure 6: Percent Interventions

Subjective Trust

The Self-Report Trust Scale (Merritt, 2011) was one scale used to measure subjective trust. An independent samples t-test was used to compare subjective trust between the High and Low risk conditions. The result of this test revealed no significant difference between the High Risk (*M* = 2.94, *SD* = .99) and Low Risk (*M* = 2.81, *SD* = .94) conditions, $t(28) = .38$, $p = .71$.

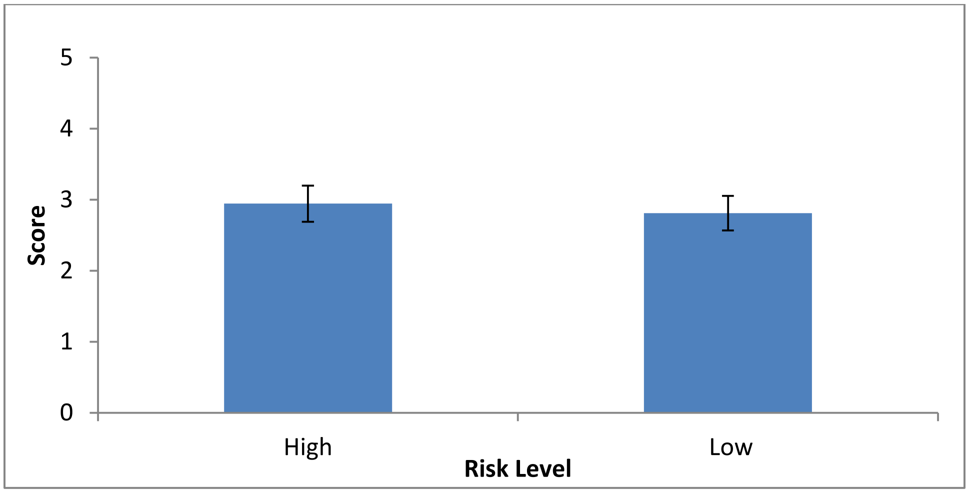

Figure 7: Subjective trust on Self-Report Scale

The Checklist for Trust (Jian et al. 2000) was also used to measure subjective trust. An independent samples t-test was used to compare subjective trust between the High and Low risk conditions. The result of this test revealed a difference approaching significance between the High risk (M = 4.3, SD = 1.2) and Low risk (M = 3.64, SD = .64) conditions, $t(21.13) = -1.83$, $p = .08$, $d = .67$.

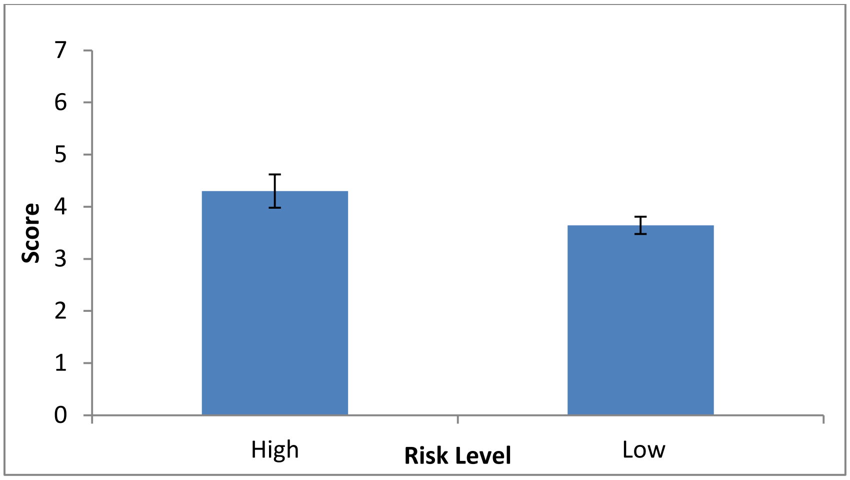

Figure 8: Subjective Trust on Checklist for Trust

Balloon Analogue Risk Task (BART)

A correlation analysis was conducted to examine the relationship between the scores on the Balloon Analogue Risk Task (BART) and the number of times a participant intervened in the automation's zone of responsibility. Results found a negative correlation approaching significance, $r = -.314$, n = 30, $p = .09$. This indicates a trend that low scores on the BART are correlated with a higher number of interventions.

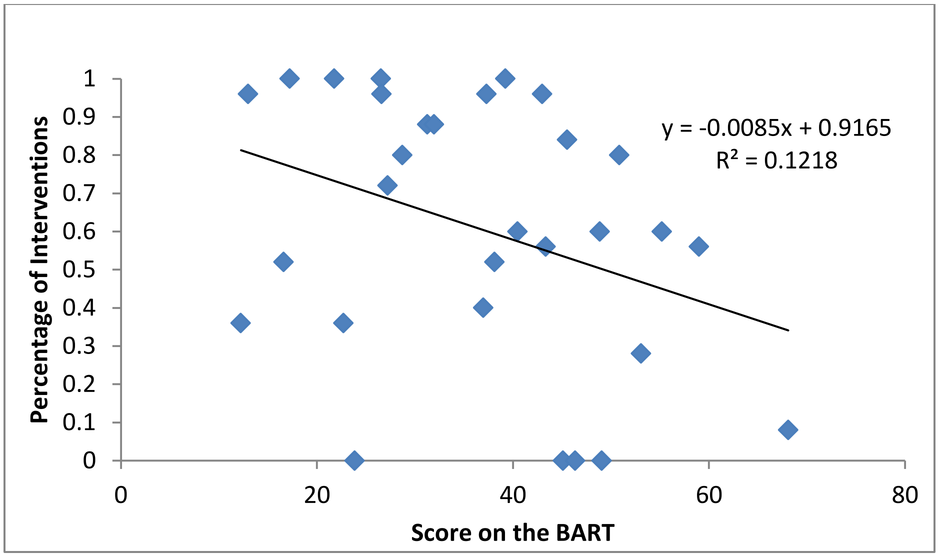

Figure 9: Relationship between BART and interventions

Individual Differences

Complacency Potential Scale

A correlational analysis was conducted to examine if overall complacency or if any of the subscales of the Complacency Potential Scale correlate with performance, subjective trust, or behavioral trust. Results demonstrated that the Confidence subscale had a significant positive correlation with the Self-Report Scale of Trust, $r = .36$, n = 30, $p < .05$. Results also demonstrated that the Reliance subscale had a significant positive correlation with score, $r = .41$, n = 30, $p < .05$.

Mini IPIP

A correlational analysis was conducted to examine if any of the personality subscales on the Mini-IPIP correlate with performance, subjective trust, or behavioral trust. All correlational relationships examined failed to reach significance.

Computer Attitudes Scale

A correlational analysis was conducted to examine if the scores on the Computer Attitudes Scale (CAS) correlate with performance, subjective trust, or behavioral trust. All correlational relationships examined failed to reach significance.

National Technology Readiness Scale

A correlational analysis was conducted to examine if overall technology readiness of if any of the National Technology Readiness subscales correlate with performance, subjective trust, or behavioral trust. Results demonstrated that the Discomfort subscale had a significant positive correlation with score, $r = .37$, n = 30, $p < .05$.

EXPERIMENT 1: DISCUSSION

Experiment 1 was designed to investigate the effect differing levels of risk has on trust in an autonomous system. It was initially predicted that trust in automation would be lower in a situation characterized by a high level of risk. This lower trust would be evidenced by reduced behavioral and subjective trust. The results of Experiment 1 partially support this hypothesis. Behavioral trust, as measured by the number of interventions into the automation's zone of responsibility, was lower in a high risk condition indicating less trust in the automation. However, subjective trust scores on both the Merritt (2011) Self-Report Scale and Jian et al. (2000) Checklist for Trust did not differ between the high and low risk conditions.

It is surprising that the behavioral trust results did not converge with the ratings of subjective trust. This could be due to the fact that the subjective trust ratings were obtained after the scenario was completed and participants then had a full picture of how well the automation performed. Subjective trust ratings may be more reflective of overall trust in hindsight, and not reflective of real-time trust. It could also be the case that because participants had experience with the automation failures in training, subjective trust did not differ because the automation was predictably unreliable for both conditions. These are reasons why behavioral indices of trust are preferable, as these measures can look at trust as on ongoing, developing process as opposed to an outcome.

Even though behavioral and subjective trust results differed, these findings are in support of the somewhat contradictory findings found by Perkins et al. (2010) which found that subjective trust was higher in a situation of higher risk, but that following a GPS's suggested route was lower in a situation of higher risk. Similarly, Experiment 1 found that behavioral trust was lower in a situation of high risk, but this effect was not present with subjective trust. Furthermore, Perkins et al. (2010) found that as risk increased, there was a positive correlation between subjective trust ratings and use in the GPS. The authors of the paper explain these findings as a distinction between common and uncommon hazards, which distinguished the risk levels. The uncommon hazards that characterized high risk were hazards such as drive-by shootings and riot outbreaks that participants are unlikely to have actually experienced outside of the experiment. Therefore, the inexperience with these hazards may have resulted in decreased use in the GPS because of real-life inexperience of using a GPS in those hazards. This points to the need to examine the effect experience also has when looking at trust in a risky situation, which is the focus of Experiment 2.

EXPERIMENT 2: METHOD

A second aim of this dissertation was to investigate how experiential factors can drive a difference in trust with different levels of risk. One potential factor may be past experience and/or training with the automation. Yuliver-Gavish and Gopher (2011) found that participants relied on a decision support system more after they gained experience using the system. In contrast, Sanchez et al. (2011) found that experienced farmers relied less on automation than participants with no farming experience. At face value, these two studies seem to offer conflicting results, but it is important to note that the participants in the Sanchez et al. (2011) study had never used that particular automated system. Therefore, the type of experience in the Sanchez et al. (2011) study is better classified as subject matter expertise and a situational factor. It could be argued that prior experience with system reliability can reduce uncertainty. The type of experience in the Yuliver-Gavish and Gopher (2011) study is better classified as a learned factor. However, if operators have interacted with a flawed system, experience may not always lead to increased trust. Manzey et al. (2012) demonstrated that negative past experiences led to less trust in an automated system. Similarly, Bahner et al. (2008) demonstrated that when participants experience an automation failure during training, they are less complacent during the experimental trial. In other words, participants who experience automation failures in training are less trusting during the experimental trial. This effect may also interact with how participants trust in a high risk scenario, further exacerbating any effects of trust within a risky scenario.

Design

The design for Experiment 2 was 2 (Experience/Information) x 2 (Risk) between subjects design. The Experience/Information variable was defined by whether or not the participant experienced a failure during training. The two levels of Risk were defined by the amount of money a participant will stood to lose in the scenario, similarly to Experiment 1.

Like Experiment 1, it was hypothesized that subjective trust would be lower in the high risk condition compared to the low risk condition. Similarly, behavioral trust would be lower in the high risk condition compared to the low risk condition. Specifically, participants would destroy more enemies in the automation's zone in the high risk condition compared to the low risk condition.

It was also hypothesized that subjective trust would be lower in the Experience condition compared to the Information condition. Similarly, behavioral trust would be lower in the Experience condition compared to the Information condition. Participants would destroy more enemies in the automation's zone in the Experience condition compared to the Information condition. Furthermore, a significant Experience/Info x Risk Interaction was hypothesized where subjective trust would be lowest in the High Risk/Experience condition. Behavioral trust would also be lowest in the High Risk/Experience group.

Experience/Information

Participants in the Experience condition experienced automation failures in training. During the seven minute practice scenario, the automation was 70% reliable, exactly like the experimental trial. Participants in the Information condition did not experience automation failures in training. During the seven minute practice scenario, the automation was 100% reliable, however, participants were told that the automation may or may not be perfectly reliable. Even though the participants in the Information condition were told that the automation may not be perfectly reliable, in actuality they experienced 100% reliable automation. In both conditions, the automation was 70% reliable.

Risk

The two levels of risk were defined in the same way as Experiment 1. All participants were given $50 at the beginning of the experimental trial. In the low risk condition, participants stood to lose $10 for scoring below 1225 points. In the high risk condition, participants stood to lose $40 for scoring below 1225 points.

Apparatus

The seven minute DDD scenario from Experiment 1 was also used in Experiment 2.

Procedure

60 participants (F = 26, *M* = 21.94Participants first completed the Balloon Analogue Risk Task (BART) to measure risk-taking behavior. Following the BART, participants viewed a self-paced PowerPoint presentation explaining DDD, the goals of the game, and how to specifically play the game. After viewing the PowerPoint,

participants completed a five minute practice scenario to make sure they were competent in how to generally play DDD. Participants needed to demonstrate that they can move throughout the scenario as well as attack enemies. Participants then completed a seven minute practice scenario with the same parameters of the experiment. Participants in the Experience condition experienced 70% reliable automation in training. Participants in the Information condition experienced 100% reliable automation in training, but were told that the automation may or may not be perfectly reliable. Following these two practice scenarios, participants completed the seven minute experimental trial. Half of the participants experienced the low stakes condition and the other half of participants experienced the high stakes condition.

Following the experimental condition, participants completed the two subjective trust questionnaires (Merritt, 2011; Jian et al., 2000) and the same individual differences measures as Experiment 1. This concluded the experimental portion of the session and participants were debriefed on the experiment and paid for their participation.

EXPERIMENT 2: RESULTS

Performance

Similarly to Experiment 1, the percentage of enemy incursions into the red-zone was used to characterize operator performance, thus, lower numbers indicate better performance. A 2 (Experience/Info) x 2 (Risk) ANOVA was conducted on the percentage of enemies allowed into the red-zone. Results revealed a significant difference for Risk, $F(1, 56) = 5.86$, $p < .05$, $\eta_p^2 = .10$, such that there were more incursions in the Low Risk condition ($M = 16.4\%$, $SD = 11.9\%$) compared to the High Risk condition ($M = 10.6\%$, $SD = 5.9\%$). Results revealed no significant difference in incursions for the Experience condition ($M = 12.2\%$, $SD = 11.0\%$) compared to the Information condition ($M = 14.8\%$, $SD = 8.3\%$), $F(1, 56) = 1.10$, $p = .30$. There was also no significant Experience x Risk interaction, $F(1, 56) = .17$, $p = .68$.

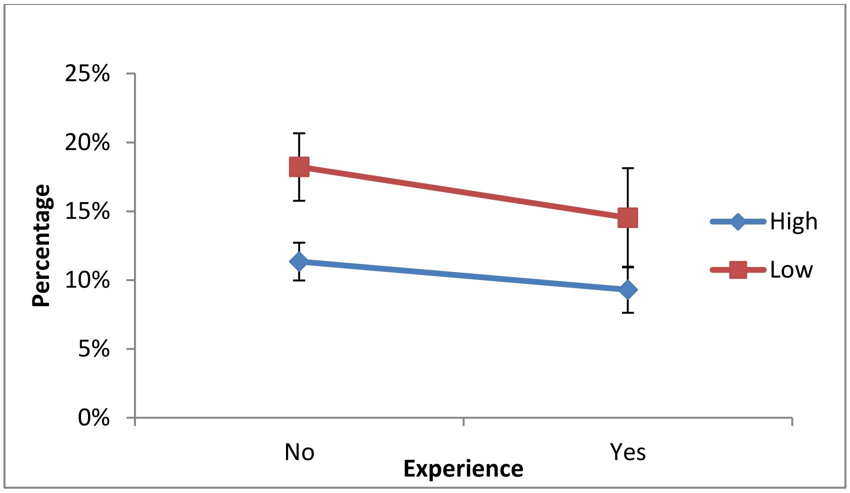

Figure 10: Percent Incursions

A 2 (Experience/Info) x 2 (Risk) ANOVA was conducted on the percentage of incursions in the participant's zone of responsibility. Results revealed a significant difference for Risk, $F(1, 56) = 7.64$, $p < .05$, $\eta_p^2 = .12$, such participants allowed more incursions in their zone of responsibility in the Low Risk condition (M = 7.1%, SD = 10.4%) compared to the High Risk condition (M =1.5%, SD = 3.4%. Results revealed no significant difference in incursions in the participant's zone of responsibility for the Experience condition (M = 4.9%, SD = 9.0%) compared to the Information condition (M = 3.6%, SD = 7.4%), $F(1, 56) = .43$, $p = .51$. There was also no significant Experience x Risk interaction, $F(1, 56) = .28$, $p = .60$.

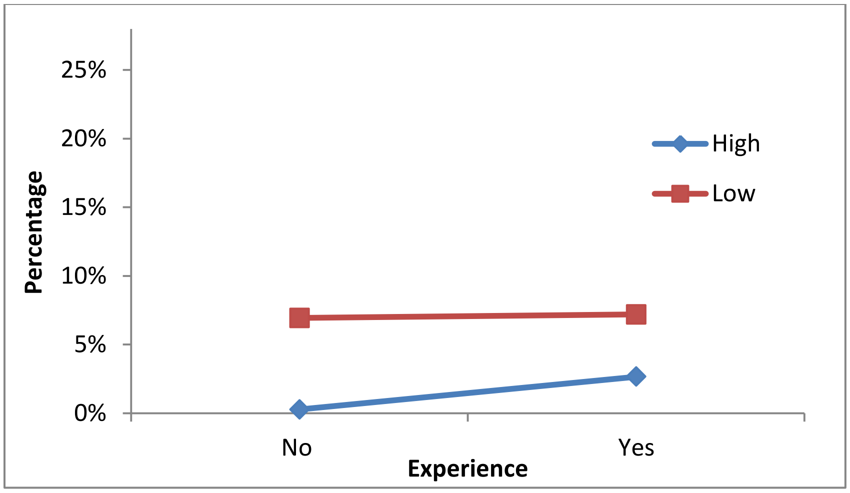

Figure 11: Percentage of Incursions in Participant's Zone of Responsibility

A 2 (Experience/Info) x 2 (Risk) ANOVA was conducted on score. Results revealed a significant difference for Risk, $F(1, 56) = 7.67$, $p < .05$, $\eta_p^2 = .12$, such participants scored significantly fewer points in the Low Risk condition ($M = 1211.33$, $SD = 132.2$) compared to the High Risk condition ($M = 1283.67$, $SD = 52.4$). Results revealed no significant difference in score for the Experience condition ($M = 1240$, $SD = 110.4$) compared to the Information condition ($M = 1255$, $SD = 103.0$), $F(1, 56) = .33$, $p = .57$. There was also no significant Experience x Risk interaction, $F(1, 56) = .97$, $p = .33$.

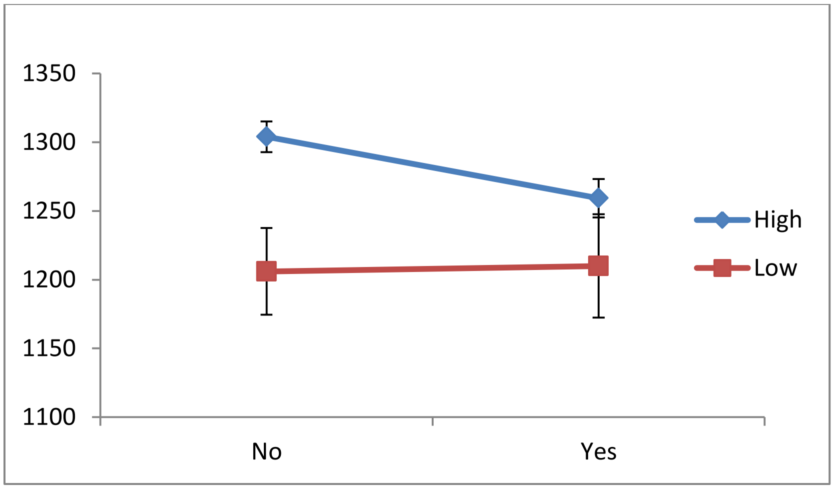

Figure 12: Score

Behavioral Trust

Similarly to Experiment 1, the number of enemies a participant killed in the automation's green zone was used to characterize behavioral trust. Therefore, a high percentage of green zone interventions indicates low trust. A 2 (Experience/Info) x 2 (Risk) ANOVA conducted on the percentage of interventions revealed a significant difference for Experience/Info, $F(1, 56) = 4.91$, $p < .05$, $\eta_p^2 = .08$, such that the percentage of interventions was higher in the Experience condition (M = 51.3%, SD = 37.3%) compared to the Information condition (M =30.1%, SD = 36.2%). Results revealed no significant difference in interventions between the High Risk condition (M = 45.1%, SD = 36.2%) and the Low Risk condition (M = 36.4%, SD = 39.8%), $F(1, 56)$ = .82, p = .37. Furthermore, there was no significant Experience x Risk interaction, $F(1, 56)$ = .16, p = .69.

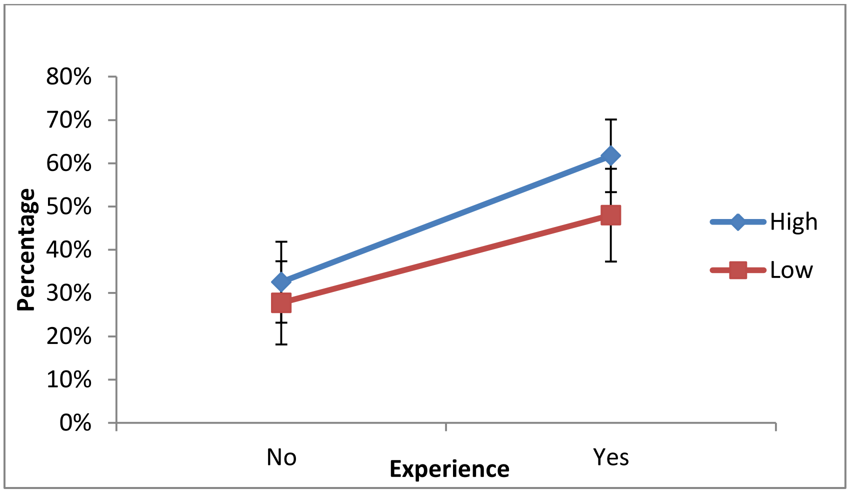

Figure 13: Percent Interventions

Subjective Trust

The Merrit (2011) Self Report Trust scale was one questionnaire used to measure subjective trust. A 2 (Experience/Info) x 2 (Risk) ANOVA revealed no significant difference in trust ratings between the High Risk (*M* = 2.91, *SD* = .85) and the Low Risk (*M* = 2.86, *SD* = .91) conditions, *F*(1, 56) = .06, *p* = .81. There was also no significant difference in trust ratings between the Experience condition (*M* = 2.83, *SD* = .86) and the Information condition (*M* = 2.94, *SD* = .90), *F* (1, 56) = .24, *p* = .63, and no significant Risk x Experience interaction, *F* (1, 56) = 1.60, *p* = .21.

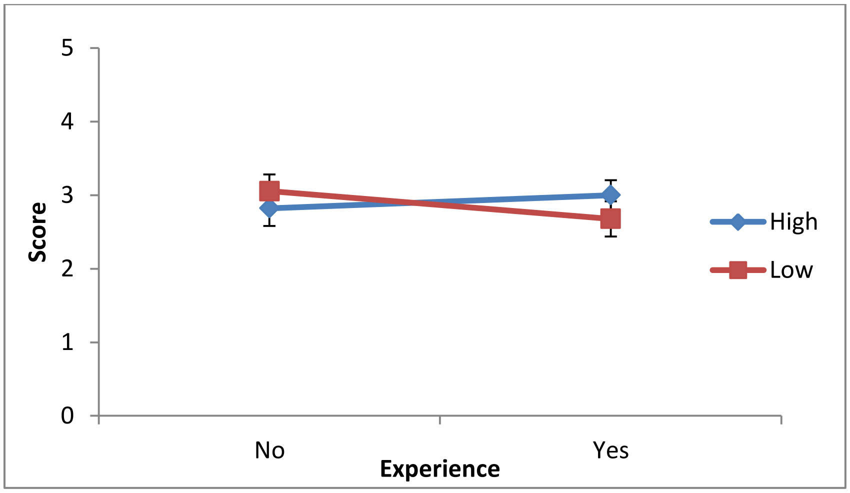

Figure 14: Subjective Trust on Self-Report Scale

The Jian et al. (2000) Checklist for Trust was also used to measure subjective trust. A 2 (Experience/Info) x 2 (Risk) ANOVA revealed no significant difference in trust ratings between the High Risk (*M* = 4.15, *SD* = 1.7) and the Low Risk (*M* = 4.21, *SD* = 1.1) conditions, $F(1, 56) = .04$, $p = .85$. There was also no significant difference in trust ratings between the Experience condition (*M* = 4.13, *SD* = 1.1) and the Information condition (*M* = 4.23, *SD* = 1.1), $F(1, 56) = .13$, $p = .72$, and no significant Risk x Experience interaction, $F(1, 56) = .13$, $p = .72$.

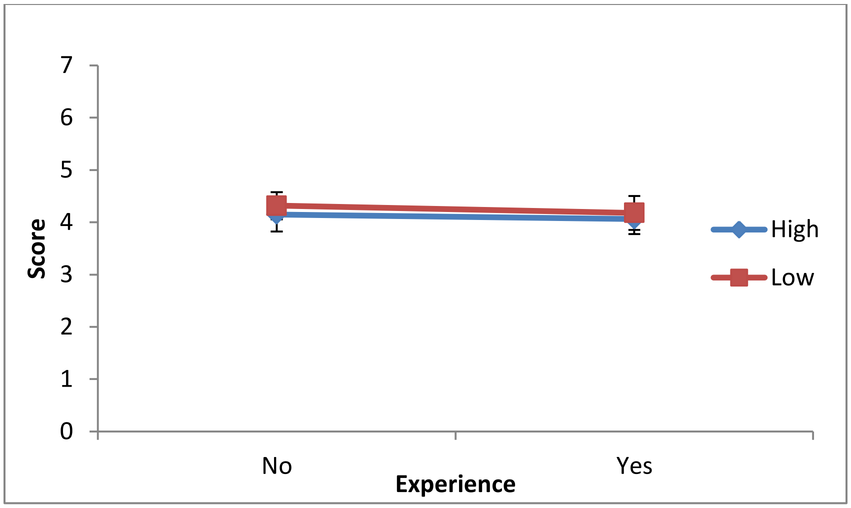

Figure 15: Subjective Trust on Checklist for Trust

Balloon Analogue Risk Task (BART)

A correlation analysis was conducted to examine the relationship between the scores on the Balloon Analogue Risk Task (BART) and the number of times a participant intervened in the automation's zone of responsibility. Results found no significant correlation, *r* = -.02, n = 30, *p* = .91.

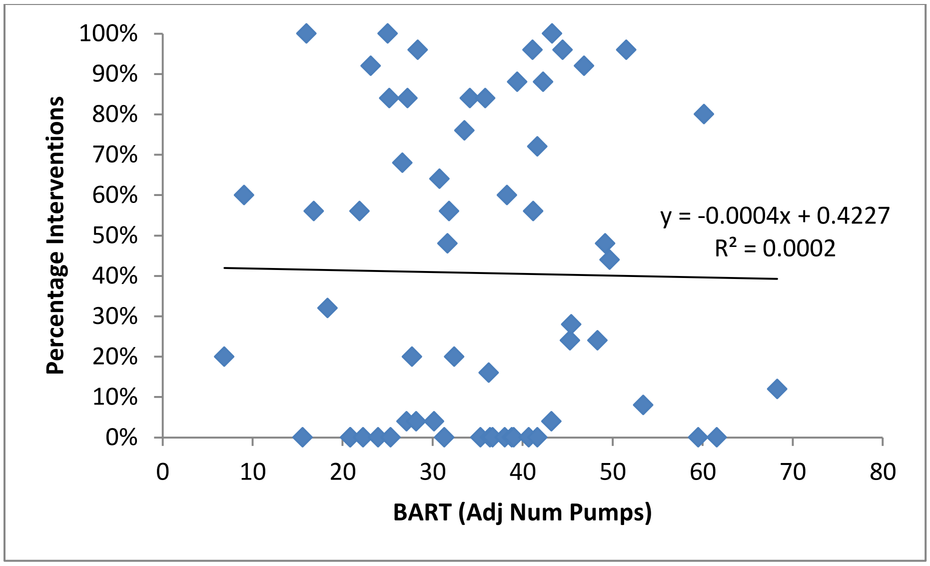

Figure 16: Relationship between BART and interventions

Individual Differences

Complacency Potential Scale

A correlational analysis was conducted to examine if overall complacency or if any of the subscales of the Complacency Potential Scale correlate with performance, subjective trust, or behavioral trust. Results demonstrated that Total Complacency had a significant positive relationship with both the Self-Report Scale of Trust, $r = .53$, n = 60, $p < .05$, and the Checklist for Trust, $r = .51$, n = 60, $p < .05$. Results also demonstrated that the Confidence subscale had a significant positive correlation with the Self-Report Scale of Trust, $r = .46$, n = 60, $p < .05$ and the Checklist for Trust, $r = .40$, n = 60, $p < .05$. The Reliance subscale had significant positive correlations with the Self-Report Scale of Trust, r = .49, n = 60, p < .05 and the Checklist for Trust, r = .58, n = 60, p < .05. The

Trust subscale has significant positive correlations with incursions, $r = .28$, $n = 60$, $p < .05$, Self-Report Scale of Trust, $r = .31$, $n = 60$, $p < .05$, and the Checklist for Trust, $r = .24$, $r = 60$, $p < .05$. The Trust subscale had a significant negative correlation with interventions, $r = -.35$, $n = 60$, $p < .05$. The Safety subscale had significant positive correlations with the Self-Report Scale of Trust, $r = .34$, $n = 60$, $p < .05$ and the Checklist for Trust, $r = .31$, $n = 60$, $p < .05$.

Mini IPIP

A correlational analysis was conducted to examine if any of the personality subscales on the Mini-IPIP correlate with performance, subjective trust, or behavioral trust. Results demonstrated that Agreeableness had significant positive correlations with the Self-Report Scale of Trust, $r = .34$, $n = 60$, $p < .05$ and the Checklist for Trust, $r = .25$, $n = 60$, $p < .05$. The Neuroticism subscale had a significant positive correlation with the Checklist for Trust, $r = -.27$, $n = 60$, $p < .05$. The Intellect subscale had a significant positive correlation with the Checklist for Trust, $r = .35$, $n = 60$, $p < .05$ and Score, $r = .26$, $n = 60$, $p < .05$.

Computer Attitudes Scale

A correlational analysis was conducted to examine if the scores on the Computer Attitudes Scale (CAS) correlate with performance, subjective trust, or behavioral trust. Results demonstrated that the (CAS) had a significant positive correlation with score, $r = .25$, $n = 60$, $p < .05$.

National Technology Readiness Scale

A correlational analysis was conducted to examine if overall technology readiness of if any of the National Technology Readiness subscales correlate with performance, subjective trust, or behavioral trust. Results demonstrated that the Discomfort subscale had a significant positive correlation with the Checklist for Trust, $r = .41$, n = 60, $p < .05$. Overall Technology Readiness had a significant positive correlation with the Checklist for Trust, $r = .30$, n = 60, $p < .05$.

Exploratory Analyses

Exploratory correlations were analyzed in order to further understand the relationships between the dependent measures. Results found a significant negative correlation between percent interventions and percent incursions, $r = -.69$, n = 60, $p < .05$, and both the Self-Report Scale of Trust, $r = -.39$, n = 60, $p < .05$, and the Checklist for Trust, $r = -.31$, n = 60, $p < .05$. Results also demonstrated a significant negative correlation between percent incursions and score, $r = -.67$, n = 60, $p < .05$.

These correlational relationships were explored further and results demonstrated that within the High Risk/Experience conditions, percent interventions has a significant negative correlation with both the Self-Report Scale of Trust, $r = -.52$, n = 15, $p < .05$, and the Checklist for Trust $r = -.71$, n = 15, $p < .05$. Within the Low Risk/Experience condition, percent interventions also has a significant negative correlation with the Self-Report Scale of Trust, $r = -.61$, n = 15, $p < .05$. Within the Low Risk/Experience condition, scores on the Self-Report Scale of Trust are positively correlated with Score, $r = .50$, n =15, $p < .05$. This is also true for the High Risk/Experience condition, scores on the Self-Report Scale of Trust are positively correlated with score, $r = .58$, n = 15, $p < .05$. Within the High Risk/Experience condition, score on the Checklist for Trust are positively correlation with score, $r = .60$, n = 15, $p < .05$.

EXPERIMENT 2: DISCUSSION

The purpose of Experiment 2 was to investigate whether prior experience with automation failures in training exacerbates the effect risk level has on trust in automation. It was hypothesized that in Experiment 2, trust would be lower in the high risk condition compared to the low risk similarly to Experiment 1. It was also hypothesized that trust would be lower when participants experienced automation failures in training compared to when they were only told that a failure may occur. Furthermore, it was expected that trust would be lowest in a high risk condition where participants experienced automation failures in training.

Results partially supported these hypotheses. As predicted, behavioral trust was lower in the Experience conditions compared to the No Experience conditions. This is consistent with previous studies that have found that participants relied on a decision support system less after they gained experience using an imperfect system (Manzey et al., 2012). More interventions can also be thought of as being less complacent; therefore these results are consistent with the findings of Bahner et al. (2008) which demonstrated that when participants experience automation failures in training, those participants are less complacent during the experimental trial. However, there was no difference in subjective trust ratings between experience conditions. Furthermore, when exploring correlational relationships, results found that interventions negatively correlated with subjective trust, as predicted, but only within the experience conditions.

GENERAL DISCUSSION

In replication of Experiment 1, it was predicted that behavioral trust would be lower in a high risk condition compared to a low risk condition. However, the results of Experiment 2 failed to replicate this finding. And similarly to Experiment 1, subjective trust ratings did not differ between risk conditions. It is worth noting that variability for interventions in all conditions for Experiment 2 was quite high, perhaps indicating a difference in strategy between participants. Some participants intervened often, while many participants did not intervene at all. What is driving this difference in strategy is worth investigating in future studies. One possibility is that when risk interacts with experience, what a risky behavior means varies. When designing these experiments, it was predicted that completely relying on potentially imperfect automation would be seen as risky, but the argument could be made that sending one's UAVs into the automation's AOR is risky since this leaves vulnerabilities in one's own AOR, particularly when a participant has not seen any automation failures. This difference could result in two different strategies. In future studies, taking care to evaluate how participants are defining risk within the scenario is important.

Results indicated that participants had more incursions and scored fewer points in the low risk conditions of both Experiment 1 and Experiment 2, which could perhaps be a result of lower motivation. Future studies should also measure if participants differ in engagement or motivation during situations of low risk. Perhaps, participants who were less engaged or unmotivated because of the low risk involved scored less.

With regards to a theoretical understanding of how risk affects trust, some unresolved issues still remain. Lee and See (2004) propose that beliefs underlie trust which then in turn affects behavior. Where risk falls in this relationship is still undetermined. Hoff and Bashir (2014) would propose that risk is a situational factor that affects trust which should in turn affect behavior. However, neither of these experiments demonstrated this specific relationship. The percentage of interventions was positively correlated with subjective trust, however, the fact that you had experience with failures in training or not drove this correlation. Results from Experiment 1 would imply that risk is a factor that affects behavior and not trust. This would be more similar to the Mayer et al. (1995) model of trust which proposes that a level of trust is developed based on a system's ability and integrity, and risk affects actual behavior, but not trust. However, this finding was not replicated in Experiment 2. As mentioned above, perhaps the definition of risk within the scenario was too ambiguous, leading to two different strategies. Or perhaps an individual difference not investigated such as self-confidence or video game experience is driving the different strategies. Further research is needed to establish how risk fits into a theoretical model of trust.

The results of these two experiments have implications for how operators interact with and trust automation in high stakes environments, particularly in regard to appropriate calibration of trust. Results of Experiment 1 demonstrate that in a high risk situation, operators may under-rely on automation and unnecessarily increase their own workload. In Experiment 1 participants in the high risk condition intervened more; however, they did not score significantly higher compared to the low risk condition. This implies that participants in the high risk condition were doing more work for no additional benefit or payout. It therefore becomes necessary to make sure that operators are appropriately calibrating behavior to maximize the benefits provided by automation. Results from Experiment 2 suggest that exposure to automation failures during practice may beneficial in preventing complacency and this has important implications for training protocols.

Results of these experiments also demonstrate the challenge of examining the effect of risk on trust in automation. These results imply that subjective ratings of trust are not always indicative of how operators will actually behave and rely on automation in a situation involving substantial risk. Individual strategies and definitions of risk make it difficult to separate general trends in behavior. Furthermore, it is difficult to simulate real risk in the laboratory environment. However, this dissertation provide results that are applicable to training and making sure operators are using automation appropriately to maximize benefits while also minimizing consequences.

Limitations

One potential limitation of the experiment is that risk is difficult to simulate in the laboratory. The different risk levels was carefully based on previous literature which titrated payments and risk aversion in a controlled experiment (Harinck et al., 2007) to see at what value risk aversion was strongest, however, there is still the possibility that participants in this study did not consider losing $40 as a high risk. This could possibly be because any payment would have been a net gain. Also, this set of experiments utilized the undergraduate student population at George Mason University which may not generalize to the population.

Future Directions

In the future, research should focus on examining the complex relationship between experience and risk. Experiment 2 found results that suggest participants were employing different strategist and future studies should investigate further what is driving the strategy difference.

APPENDIX

Table 1: Experiment 1: Complacency Potential Scale Correlations

	Total Complacency	Confidence	Reliance	Trust	Safety
Incursions	-.13	-.26	-.09	.24	-.21
Score	.33	.16	.41*	.02	.29
Interventions	-.04	.20	-.19	-.26	.11
Self-Report Scale of Trust	.31	.36*	.30	.06	-.03
Checklist for Trust	.18	.25	.19	-.04	-.01
M	44.0	14.6	10.8	11.9	6.7
SD	5.09	2.56	1.88	1.97	1.47

Table 2: Experiment 1: Mini-IPIP Correlations

	Extroversion	Agreeableness	Conscientiousness	Neuroticism	Intellect
Incursions	.18	.03	.03	-.11	-.16
Score	-.12	.12	.34	-.19	-.01
Interventions	-.13	-.17	-.34	.25	.15
Self-Report Scale of Trust	.09	-.05	.15	-.17	.03
Checklist for Trust	.05	.09	.10	.03	.02
M	11.6	12.9	15.1	10.6	13.5
SD	3.39	3.03	2.80	2.91	1.87

Table 3: Experiment 1: Technology Readiness Scale Correlations

	Overall	Optimism	Innovativeness	Discomfort	Insecurity
Incursions	.02	.13	-.12	-.10	.15
Score	-.17	-.04	.37*	.17	.10
Interventions	-.09	-.02	.22	-.19	-.32
Self-Report Scale of Trust	.06	.08	.01	.04	.03
Checklist for Trust	.13	.05	.05	.15	.06
M	3.32	4.24	3.41	3.0	2.68
SD	.29	.45	.69	.36	.55

Table 4: Experiment 2: Complacency Potential Scale Correlations

	Total Complacency	Confidence	Reliance	Trust	Safety
Incursions	.11	.10	-.03	.28*	-.09
Score	.05	-.07	.10	-.03	.23
Interventions	-.19	-.06	-.07	-.35*	-.06
Self-Report Scale of Trust	.53*	.46*	.49*	.31*	.34*
Checklist for Trust	.51*	.40*	.58*	.25	.31*
M	45.8	15.2	10.5	11.0	6.2
SD	6.30	2.23	2.16	2.35	1.50

Table 5: Experiment 2: Mini-IPIP Correlations

	Extroversion	Agreeableness	Conscientiousness	Neuroticism	Intellect
Incursions	.12	.16	-.07	.08	-.06
Score	-.12	-.04	.09	.01	.26*
Interventions	-.01	-.12	.00	-.11	-.20
Self-Report Scale of Trust	.15	.34*	.23	-.09	.21
Checklist for Trust	.20	.25	.20	-.27*	.35*
M	12.4	15.9	15.4	10.7	14.0
SD	3.38	3.01	2.78	2.89	1.9

Table 6: Experiment 2: Technology Readiness Scale Correlations

	Overall	Optimism	Innovativeness	Discomfort	Insecurity
Incursions	-.01	.05	.08	-.01	-.16
Score	.14	.11	.01	.16	.14
Interventions	-.09	-.13	-.08	-.14	.07
Self-Report Scale of Trust	.19	.24	.10	.23	-.01
Checklist for Trust	.30*	.18	.09	.41*	.20
M	3.32	4.09	3.29	2.82	2.25
SD	.45	.60	.97	.47	.74

Table 7: Exploratory Correlations: High Risk/No Experience

	Interventions	Incursions	Self-Report Scale of Trust	Checklist for Trust	Score
Interventions					
Incursions	-.95*				
Self-Report Scale of Trust	-.24	.12			
Checklist for Trust	-.14	.01	.8*		
Score	-.73*	.50	.44	.40	

Table 8: Exploratory Correlations: Low Risk/No Experience

	Interventions	Incursions	Self-Report Scale of Trust	Checklist for Trust	Score
Interventions					
Incursions	-.63*				
Self-Report Scale of Trust	-.17	-.18			
Checklist for Trust	-.29	.19	.59*		
Score	-.08	-.70*	.50	.07	

Table 9: Exploratory Correlations: High Risk/Experience

	Interventions	Incursions	Self-Report Scale of Trust	Checklist for Trust	Score
Interventions					
Incursions	-.80*				
Self-Report Scale of Trust	-.52*	.10			
Checklist for Trust	-.61*	.21	.65*		
Score	-.23	-.39	.58*	.60*	

Table 10: Exploratory Correlations: Low Risk/Experience

	Interventions	Incursions	Self-Report Scale of Trust	Checklist for Trust	Score
Interventions					
Incursions	-.67*				
Self-Report Scale of Trust	-.71*	.65*			
Checklist for Trust	-.29	.08	.70*		
Score	.14	-.81*	-.33	.12	

REFERENCES

Ajzen, I., & Fishbein, M. (1980). Understanding attitudes and predicting social behaviour.

Bagheri, N., & Jamieson, G. A. (2004). Considering subjective trust and monitoring behavior in assessing automation-induced "complacency.". *Human performance, situation awareness, and automation: Current research and trends*, 54-59.

Bahner, J. E., Hüper, A. D., & Manzey, D. (2008). Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies*, *66*(9), 688-699.

Bailey, N. R., & Scerbo, M. W. (2007). Automation-induced complacency for monitoring highly reliable systems: The role of task complexity, system experience, and operator trust. *Theoretical Issues in Ergonomics Science*,*8*(4), 321-348.

Barber, B. (1983). The logic and limits of trust.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, *10*(1), 122-142.

Biros, D. P., Daly, M., & Gunsch, G. (2004). The influence of task load and automation trust on deception detection. *Group Decision and Negotiation*,*13*(2), 173-189.

Bottom, W. P., Holloway, J., Miller, G. J., Mislin, A., & Whitford, A. (2006). Building a pathway to cooperation: Negotiation and social exchange between principal and agent. *Administrative Science Quarterly*, *51*(1), 29-58.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Cameron, L. A. (1999). Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry*, *37*(1), 47-59.

Carpenter, J., Verhoogen, E., & Burks, S. (2005). The effect of stakes in distribution experiments. *Economics Letters*, *86*(3), 393-398.

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.

Deutsch, M. (1958). Trust and suspicion. *Journal of conflict resolution*, 265-279.

Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *Mis Quarterly*, 373-396.

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*(3), 601-620.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological assessment*, *18*(2), 192.

Driscoll, J. W. (1978). Trust and participation in organizational decision making as predictors of satisfaction. *Academy of management journal*,*21*(1), 44-56.

Ferrary, M. (2003). Trust and social capital in the regulation of lending activities. *The Journal of Socio-Economics*, *31*(6), 673-699.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior*, *6*(3), 347-369.

Ganesan, S. (1994). Determinants of long-term orientation in buyer-seller relationships. *the Journal of Marketing*, 1-19.

Gill, J., & Butler, R. (1996). Cycles of trust and distrust in joint-ventures.*European Management Journal*, *14*(1), 81-89.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models.*Personality psychology in Europe*, *7*(1), 7-28.

Güth, W. (2002). Do banks crowd in business ethics? An indirect evolutionary analysis. *International Review of Economics & Finance*, *10*(1), 1-17.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization*, *3*(4), 367-388.

Harinck, F., Van Dijk, E., Van Beest, I., & Mersmann, P. (2007). When gains loom larger than losses reversed loss aversion for small amounts of money.*Psychological science*, *18*(12), 1099-1105.

Ho, G., Wheatley, D., & Scialfa, C. T. (2005). Age differences in trust and reliance of a medication management system. *Interacting with Computers*,*17*(6), 690-710.

Hoff, K. A., & Bashir, M. (2015). Trust in automation integrating empirical evidence on factors that influence trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(3), 407-434.

Hoffman, E., McCabe, K., & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review*,*86*(3), 653-660.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects.*American economic review*, *92*(5), 1644-1655.Inkpen, A. C., & Currall, S. C. (1998). The nature, antecedents, and consequences of joint venture trust. *Journal of International Management*,*4*(1), 1-20.

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.

Johansson-Stenman, O., Mahmud, M., & Martinsson, P. (2005). Does stake size matter in trust games?. *Economics Letters*, *88*(3), 365-369.

Johns, J. L. (1996). A concept analysis of trust. *Journal of Advanced Nursing*, *24*(1), 76-83.

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis.*Journal of Economic Psychology*, *32*(5), 865-889.

Karlan, D. S. (2005). Using experimental economics to measure social capital and predict financial decisions. *American Economic Review*, 1688-1699.

Lewandowsky, S., Mundy, M., & Tan, G. (2000). The dynamics of trust: comparing humans to automation. *Journal of Experimental Psychology: Applied*, *6*(2), 104.

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243-1270.

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: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50-80.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., ... & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75.

Lewicki, R. J., & Bunker, B. B. (1995). *Trust in relationships: A model of development and decline*. Jossey-Bass.

Li, D., Rau, P. L. P., & Li, Y. (2010). A cross-cultural study: effect of robot appearance and task. *International Journal of Social Robotics*, *2*(2), 175-186.

Liaw, S. S. (2002). An Internet survey for perceptions of computers and the World Wide Web: relationship, prediction, and difference. *Computers in human behavior*, *18*(1), 17-35.

Lyons, J. B., & Stokes, C. K. (2011). Human–human reliance in the context of automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 0018720811427034.

Madhavan, P., & Wiegmann, D. A. (2005). Cognitive anchoring on self-generated decisions reduces operator reliance on automated diagnostic aids. *Human Factors: The Journal of the Human Factors and Ergonomics Society*,*47*(2), 332-341.

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*, *8*(4), 277-301.

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*, 1555343411433844.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of management review*, *20*(3), 709-734.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management review*, *23*(3), 473-490.

Merritt, S. M. (2011). Affective Processes in Human–Automation Interactions. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *53*(4), 356-370.

Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(2), 194-210.

Moorman, C., Deshpande, R., & Zaltman, G. (1993). Factors affecting trust in market research relationships. *the Journal of Marketing*, 81-101.

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *The journal of marketing*, 20-38.

Munier, B., & Zaharia, C. (2002). High stakes and acceptance behavior in ultimatum bargaining. *Theory and Decision*, *53*(3), 187-207.

Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, *7*(3), 171.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues*, *56*(1), 81-103.

Nass, C., Steuer, J., & Tauber, E. R. (1994, April). Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 72-78). ACM.

Nyhan, R. C. (2000). Changing the paradigm trust and its role in public sector organizations. *The American Review of Public Administration*, *30*(1), 87-109.

Parasuraman, A. (2000). Technology Readiness Index (TRI) a multiple-item scale to measure readiness to embrace new technologies. *Journal of service research*, *2*(4), 307-320.

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

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230-253.

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

Perkins, L., Miller, J. E., Hashemi, A., & Burns, G. (2010, September). Designing for human-centered systems: Situational risk as a factor of trust in automation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, No. 25, pp. 2130-2134). SAGE Publications.

Rasmussen, J., Pejterson, A. M., & Goodstein, L. P. (1994). Cognitive systems engineering. New York: John Wiley.

Rau, P. P., Li, Y., & Li, D. (2009). Effects of communication style and culture on ability to accept recommendations from robots. *Computers in Human Behavior*, *25*(2), 587-595.

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

Riedl, R., Hubert, M., & Kenning, P. (2010). Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. *Mis Quarterly*, *34*(2), 397-428.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust1. *Journal of personality*, *35*(4), 651-665.

Sanchez, J., Rogers, W. A., Fisk, A. D., & Rovira, E. (2014). Understanding reliance on automation: effects of error type, error distribution, age and experience. *Theoretical Issues in Ergonomics Science*, *15*(2), 134-160.

Sanchez, J., Fisk, A. D., & Rogers, W. A. (2004, September). Reliability and age-related effects on trust and reliance of a decision support aid. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*(Vol. 48, No. 3, pp. 586-589). SAGE Publications.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*(5626), 1755-1758.

Singh, I. L., Molloy, R., & Parasuraman, R. (1993). Automation-induced" complacency": Development of the complacency-potential rating scale. *The International Journal of Aviation Psychology*, *3*(2), 111-122.

Stokes, C. K., Lyons, J. B., Littlejohn, K., Natarian, J., Case, E., & Speranza, N. (2010, May). Accounting for the human in cyberspace: Effects of mood on trust in automation. In *Collaborative Technologies and Systems (CTS), 2010 International Symposium on* (pp. 180-187). IEEE.

Togia, A., Malliari, A., Korobili, S., & Tsigilis, N. Selected Papers Presented at the 2nd Qualitative and Quantitative Methods in Libraries.

Wicks, A. C., Berman, S. L., & Jones, T. M. (1999). The structure of optimal trust: Moral and strategic implications. *Academy of Management review*,*24*(1), 99-116.

Williamson, O. E. (1993). Calculativeness, trust, and economic organization.*The Journal of Law & Economics*, *36*(1), 453-486.

Yuviler-Gavish, N., & Gopher, D. (2011). Effect of descriptive information and experience on automation reliance. *Human Factors*, *53*, 230–244.

BIOGRAPHY

Kelly Satterfield received her Bachelor of Science from the University of Dayton in 2009 and her Master of Arts from George Mason University in 2013.