Quantifying Compliance and Reliance Trust Behaviors to Influence Trust in Human-Automation Teams

Jayson G. Boubin, Christina F. Rusnock, and Jason M. Bindewald
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433

Abstract. Automation is utilized heavily in many domains to increase productivity. With new, more complex automation, like the self-driving car, humans will be required to forego direct task performance in favor of maintaining a supervisory role over automation systems. While the use of these systems generally results in greater performance than humans performing alone, humans are reluctant to adopt these superior systems due to a lack of trust. The United States Department of Defense is investigating trust in automation in order to influence the rate of adoption of automation technology. Studying trust in automation systems requires a mechanism for quantifying and measuring trust. This paper proposes a method for measuring human trust behaviors with regard to human-automation systems through response rates of compliance and reliance. Using behavioral data from a human-subjects experiment involving automated agents, we create a system dynamics model which relates trust to other system level variables. Using this trust model, engineers will be able to study trust in human-automation team scenarios in order to design automation systems with higher rates of adoption.

Keywords. Trust, System Dynamics, Modeling, Compliance, Reliance, Automation

INTRODUCTION

Automation is used in nearly every industry to improve productivity, increase efficiency, and prevent or limit human error. Increasing the quantity of automated tasks not only enables individual human operators to accomplish more tasks, but also allows them to accomplish tasks of greater complexity. Thus automation, throughout many domains, allows the human operator to supervise systems which make low level decisions, allowing operators to make more complex, system level choices. With the advent of self-driving cars (Google, 2015), automation systems are now, more than ever, able to make independent decisions. These low level decisions are supervised by human operators, and often affect not only operators but other stakeholders such as passengers or pedestrians as well. Fortunately, these decision-making automated systems are high performing because they are able to quickly analyze large amounts of information about their specific context which often includes the outside world. They do not experience fatigue, distraction, or other human qualities which limit decision making ability. Combined with speed and processing power, the large amount of information available to these systems and their programmatic nature allow them to perform better than systems controlled entirely by the decisions of human operators. For example, between 2014 and 2015, Google's self-driving car fleet experienced no collisions which were the fault of the automation after traveling a combined 424,331 miles (Google, 2015). The Virginia Tech Transportation Institute found that self-driving cars were three to four times less likely to crash compared to collisions involving human operated cars (Blanco et al., 2016).

Despite superior performance, most automated systems do not have the ability to replace the human operator. While Google's self-driving car automation was active for the majority of the driving process, between 2014 and 2015, human operators were required to take control of their self-driving cars over 300 times for automation failures (Google, 2015). Human operators provide value because they are capable of making judgments in situations not accounted for in the automation's design. The human and automation must become a team to safely accomplish tasks as complex as driving. In this team, the human maintains a supervisory role, while the automation makes the majority of the decisions.

In addition to the transportation sector, the United States Department of Defense is also interested in the use of human-automation teams. The Defense Science Board Autonomy Task Force sought to understand and take advantage of advances in automation research by reviewing current work. The Task Force found a number of defense focused uses for automation technology, such as Unmanned Aerial Vehicles (UAVs), Unmanned Ground Systems, Unmanned Maritime Vehicles, and Unmanned Space Systems. While performance benefits are apparent from current automation use in the DoD, there is still resistance to further adoption. The misconception that automation systems are self-governing and make decisions without the possibility of human intervention has dampened the rate of adoption of automation technology. These systems are perceived as ineffective and are not trusted by users and leadership. The Task Force made a series of suggestions for the DoD in order to increase the adoption and efficacy of automated systems. For example, they recommended that the Under Secretary of Defense for Acquisition, Technology, and Logistics create operational training techniques that focus explicitly on building trust in automation systems (Kaminski, 2012).

The United States Air Force has also expressed great interest in automation technology. The Air Force is interested in the use of automation technology in UAV's, manned vehicles such as the F-35, satellite communication, and cyber-security. In these contexts, the Air Force seeks to implement automation in such a way that it will aid airmen in performing the increasingly demanding tasks required of them (Endsley, 2015). While automation would most certainly assist in the performance of these tasks, these performance gains can only be realized if users accept the system, which requires trust. By understanding operator and stakeholder trust in an automated system, we can influence the design process to create more adoptable automation. In order to study and influence operator trust in human-automation teams, we must first be able to quantify trust in automation systems that have some degree of autonomy. With the ability to study trust in a system, we will be able to influence user acceptance and create more trustworthy automation.

Compliance and reliance are two behaviors that are indicative of trust or mistrust. In the alarm automation domain, compliance describes the operator's response when an alarm sounds, whether true or false. A compliant operator will rapidly switch attention from concurrent activities to the alarm domain (and possibly immediately initiate an alarm-appropriate response) (Dixon & Wickens, 2006). Reliance, in the alarm automation domain, refers to the operator's response when the alarm is silent. Reliant operators have ample resources to allocate to concurrent tasks because they rely on the automation to let them know when a problem occurs (Dixon & Wickens, 2006). In our research, we broaden these original definitions, which focus on alarms to include more general interactions between the human and automation. For the purpose of this research we will define *compliance* as the acceptance of an automation's actions by the human. We will define *reliance* to mean the acceptance of an automation's non-action by the human.

Reliance and compliance have previously been studied in order to limit automation failure due to misuse and disuse. Misuse refers to the failures that occur when operators inadvertently violate critical assumptions and rely on automation improperly, whereas disuse signifies failures that occur when people reject the capabilities of automation and dismiss a correct indication (Lee & See, 2004). Automation misuse, disuse, and abuse have been studied to determine their causes for preventative purposes. Research shows that misuse stems from over-reliance in automation due to factors such as high workload, little self-confidence, or poor heuristics due to lack of training or decision bias. Disuse stems from automation failure and the subsequent decrease in trust which is then exhibited by a decrease in operator reliance or compliance (Parasuraman & Riley, 1997).

Studies on reliance have shown it to be the product of emotion, which was deemed "critical for the appropriate direction of attention since it provides an automated signal about the organism's past experience" (Lee & See, 2004). While compliance and reliance were introduced in separate contexts in the 1980's, in the early 2000's researchers began studying the elements together as trust components. Researchers in fields dealing with heavy automation, particularly unmanned aerial vehicle control, began using compliance and reliance as behavioral indicators of trust (Dixon & Wickens, 2006; Dixon, Wickens, & Chang, 2005; Meyer, Gurion, & Sheva, 2004; Wickens, Dixon, Goh, & Hammer, 2005). UAS studies determined compliance and reliance to be somewhat independent from each other, meaning that they are not entirely dependent nor independent (Dixon, Wickens, & McCarley, 2006).

Purpose

While compliance and reliance have been studied in the context of systems specific to automation alarms, we believe it is possible to generalize the definitions of compliance and reliance to fit any system which includes trust. The purpose of this paper is to 1) expand the definition of reliance and compliance beyond automation alarms, 2) demonstrate a method for inferring trust by quantifying compliance and reliance in the context of a specific system, and 3) use these values to create a model for compliance and reliance for automated systems in general. Additionally, we used these measurements to evaluate how automation strategy and taskload influence reliance and compliance.

METHODOLGY

To accomplish our research goals, we performed human-in-the-loop experiments using Space Navigator, an air-traffic control style game specifically designed to evaluate the effectiveness of human-automation teams. It includes multiple types of automated agents, the ability to capture and time events, and the ability to impose easily distinguishable taskload levels, all within a basic computer game environment that is easy to learn and enjoyable to play.

Experimental Environment and Task

Space navigator has four components: ships, planets, no-fly zones, and bonuses. Ships in space navigator are spawned off screen every two seconds and come onto the screen moving in a random direction at a fixed velocity. Each ship has its own color corresponding to its destination planet. The human is able to direct ships by pressing the ship on the touchscreen and drawing a line from the ship to any point. Planets in Space Navigator signify the ship destinations. If a ship reaches its planet, the player is awarded 100 points. If two ships crash before either reaches its planet, the player loses 100 points per ship. The player may also route ships through no-fly zones and bonuses. A ship routed through a no-fly zone, a discolored square on the screen, will lose 10 points per second it is within that region. A ship routed through a bonus will gain 50 points per bonus. Bonuses are spawned at a static rate and do not disappear if they are not collected. There will always be at least two no-fly zones on screen at any given time, they change position on the map at a static rate, and may overlap.

Automated Agents

In some games of Space Navigator, simple reflex agents (referred to as automated agents) can specify ship routes. In automation games, if a ship has not been routed within two seconds, the ship will be routed by a predesignated automation strategy. Space Navigator has three automation strategies which accomplish the route-creation task of directing ships to their corresponding planet. The first strategy, referred to as similar automation, uses the Nearest Neighbor algorithm to interpret player data and provide a path similar to the player's previous draws. The second strategy, dissimilar automation, uses a trigonometric function to create a sinusoidal curved path which may or may not reach the planet. The player's play style has no bearing on dissimilar automation routes. The final strategy, line automation, draws a straight line directly to the planet. These automation strategies all ignore other ships, bonuses, and no-fly zones. They only consider the ship and its calculated path to the planet (Bindewald, Peterson, & Miller, 2016).

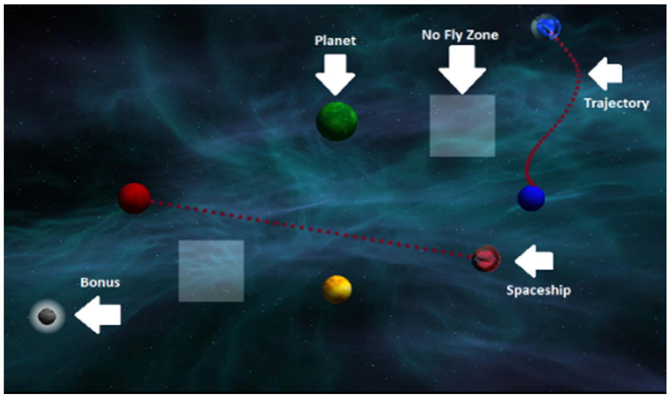

Figure 1: Image of Space Navigator Gameplay

Participant Study

The experiment included 36 participants, 6 females and 30 males, with a mean age of 32.5 and a standard deviation of 5.87. Each participant played 17 games of Space Navigator, five training games using no automation, and 12 experimental trials with three games for each automation strategy (no automation, line, similar, and dissimilar) over the course of 96 hours. The data collected included bonus collection, crashes, destination reaches, no-fly zone entries and exits, ship spawns, ship movement off screen, and automation and manual draw start and end times. The remaining method sections define the dependent variables—compliance and reliance—and the independent and endogenous variables— automation strategy and taskload.

Dependent Variables: Compliance and Reliance

As discussed above, this research extends the definitions of compliance and reliance beyond the alarms literature. For this research, we define *compliance* to mean the acceptance of an automation's actions by the human, and we define *reliance* to mean the acceptance of an automation's non-action by the human. In the context of Space Navigator, an operator's compliance rate is measured as the rate at which ships routed by automation are not redrawn (i.e. accepted by the operator). The player, at any time, has the ability to draw a route for any ship on the screen. Redrawing of automated routes represent a lack of compliance, whereas the lack of redraws represents the willingness of the person to comply with the automation's suggested route. Thus, redrawing 20% (keeping 80%) of the routes represents a compliance rate of 80%.

In the context of Space Navigator, an operator's reliance rate can be defined as the rate at which *initial* routes are allowed to be drawn by the automation as opposed to drawn by the human (recall there is a 2 sec time delay for automation route draws). A reliant operator in this situation would allow the automation to draw paths for as many ships as possible, trusting the automation to handle the route creation task. An unreliant operator would not allow the automation to draw routes, and instead of waiting for the automation to draw routes, the operator would draw the initial route.

Recall that the Space Navigator game consists of more than just route creation; operators are also charged with collision avoidance, no-fly zone avoidance, and picking up bonuses. These tasks are not fully accounted for by any of the automation types and even if they were, the automation cannot predict random movements of the no-fly zones and placement of new bonuses. Thus the automation will not play perfectly, and a player trying to maximize their score should not be 100% compliant with or 100% reliant on the automation. However, the automated agents are helpful enough that it would also be disadvantageous to be totally unreliant or non-compliant with the automation under all conditions.

Independent Variables

This research examines how automation strategy influences a player's reliance and compliance rates. Automation strategy refers to the method the automation uses to generate routes—similar, dissimilar, or line. Automation strategies may affect compliance and reliance by exhibiting different performance, or the appearance of different performance, than other strategies. Players can compare and contrast automation strategies over the course of the study, allowing them to form opinions about the superiority of certain automation strategies or the predictability of the automation strategy, which may influence their compliance and/or reliance rates.

Endogenous Variable

This research also examines how taskload influences a player's reliance and compliance rates. Taskload is the current demand being experienced by the operator and is measured as the number of ships on screen at a given time. Ship spawns, arrivals at planets, movements off-screen, and crashes all impact the quantity of ships on screen, thus creating a variable, taskload, throughout game play. Taskload may affect compliance and reliance by forcing participants to, in periods of higher taskload, rely on automation more to draw initial routes and comply with automation by accepting automation-drawn routes. Increased taskload can also affect the complexity of the game, requiring players to re-draw routes to avoid crashes.

ANALYSIS AND RESULTS

Independent Variable: Automation Strategy

We expected to see increased compliance and reliance rates with automation which was more understandable and higher performing. We hypothesized that similar and line automation would experience higher compliance and reliance rates than dissimilar automation because the similar strategy is consistent with the user's own behavior and the line strategy is predictable, and thus relatively easy to accommodate into an overarching strategy. Consistent with this hypothesis, the similar automation strategy produced the highest compliance and reliance rates, with means of 79.3% and 69.5%, respectively. Line automation had the next highest compliance and reliance rates of 76.3% and 66.3%, respectively, whereas dissimilar had the lowest compliance and reliance rates with 63.8% and 59.5%, respectively. The one-way between subjects ANOVAs reveal that there is a significant effect of automation strategy on compliance at the p<.05 level [F (2,105) = 13.54, p = 5.86E-06], as well as for reliance at the p<.05 level [F (2,105) = 4.045, p = 0.02].

Post hoc comparisons for the reliance ANOVA using the Tukey HSD test indicated that the mean reliance for similar automation (m = 0.6952, SD = 0.162) was significantly higher than reliance for the dissimilar automation (m =0.5953, SD = 0.1633). Reliance for line automation (m = 0.663, SD = 0.1274) was not significantly different from the reliance for either similar or dissimilar automation. Comparisons for the compliance ANOVA using the Tukey HSD test indicated that the mean compliance for similar automation (m = 0.7926, SD = 0.1288) was significantly higher than the compliance for the dissimilar automation (m = 0.6383, SD = 0.1622). The mean compliance for line automation (m = 0.7628, SD = 0.1027) was significantly higher than the mean compliance for dissimilar automation, but not similar automation.

These results show that, automation strategy has an effect on compliance and reliance, and that users are more likely to rely upon automation that performs the task similar to the way the user performs the task. However, simply performing significance tests on mean compliance and reliance by automation strategy may not be sufficient to determine true differences. Because there is a fair amount of variability in the mean values, it is likely that the compliance and reliance rates vary according to additional factors, such as the taskload.

Endogenous Variable: Taskload

For each automation strategy, we hypothesized that compliance and reliance would increase as taskload increases, due to the participant's inability to fully manage the task at hand. To analyze compliance and reliance rates with respect to taskload, we examined these rates for each discrete level of taskload (between 5 and 11 ships on screen). For each level of taskload, we calculated an average compliance and reliance based on behavioral data for those states across each of the participant's twelve experimental trials. Taskload was found to be a significant factor affecting compliance and reliance rates for every automation strategy. Qualitative observation of trends seen in Figure 2 suggest that compliance and reliance increase with taskload for the similar and dissimilar automation strategies, as expected, but decreases with increasing taskload for the line automation. This decrease in reliance and compliance for the line automation strategy indicates that this particular strategy decreases in effectiveness as taskload increases.

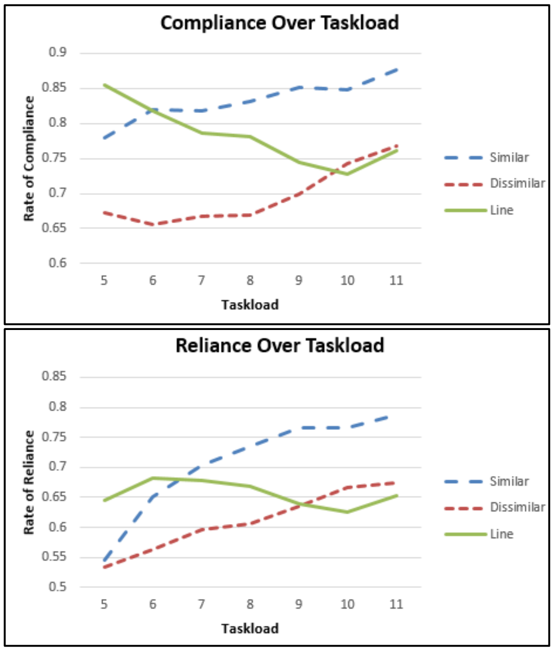

Figure 2: Graphs of Compliance and Reliance by Taskload

Compliance and Reliance: System Level Model

The Causal Loop Diagram in Figure 3 captures the relationship between attributes such as user stress, trust, compliance, and reliance, along with automation performance, predictability and taskload.

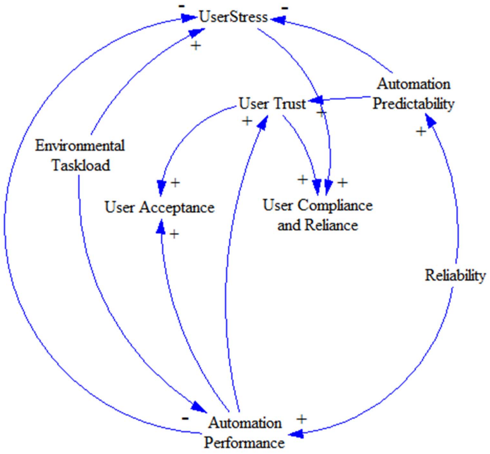

Figure 3: Causal Loop Diagram of Human-Automation Team Trust

In the Causal Loop Diagram system attributes are denoted by text, and the relationships between those attributes are denoted by arrows. The causal loop diagram contains variables empirically studied by our experiments (reliability, performance, environmental taskload, compliance, and reliance) and subjective human behaviors (user acceptance, stress, automation predictability, and user trust)

Each arrow, or causal link, is directed and signed. Variables with positive causal links have positive relationships (thus an increase in the start node, results in an increase in the end node), while variables with negative causal links have negative relationships (thus an increase in the start node, results in a decrease in the end node).

In Figure 3, compliance and reliance correlate positively with both user stress and user trust. When operators experience higher stress–due to high taskload or confusing automation—their compliance and reliance behaviors increase. Likewise, when participants trust in automation increases–due to predictable automation or increased system performance—participants tend to increase trusting behaviors. In Figure 2, there is an increase in both compliance and reliance over taskload for dissimilar and similar automation. However, there is a decrease in compliance and reliance over taskload for line automation. When taskload increases, the causal loop diagram reveals that stress increases, which increases trusting behavior in a positive reinforcement loop—as seen with similar and dissimilar automation strategies.

This chain of events is in contrast to the opposite effect taskload has of decreasing system performance, which is a negative reinforcement loop. When system performance decreases, trust decreases which decreases trusting behavior. These two loops, both controlled by taskload, cause either an increase or decrease in trust behavior depending on automation design. In similar and dissimilar automation, the positive reinforcing loop overcomes the negative reinforcing loop. In line automation, the opposite occurs, with decreased performance playing a larger role in determining trusting behaviors. Thus, the effects taskload has on trust behaviors are dependent on automation design.

DISCUSSION

Mean compliance rates ranged from 63.8-79.3% and mean reliance rates ranged from 59.5-69.5%, with similar automation experiencing the highest rates and dissimilar automation experiencing the lowest rates. However, by examining only the mean compliance and reliance by automation strategy, we may not get an accurate picture of true behavioral differences in the trust users place in each automation strategy. For example, when examining these rates across taskload, we see that for both similar and dissimilar automation, compliance and reliance rates increase as taskload increases. On the other hand, the compliance rates start off high for line automation when taskload is low, but decline as taskload increases. Line automation reliance rates stay fairly constant across taskload. This is likely a result of the interaction between effective game play strategy and taskload. When few ships are on screen, straight line automation performs well. It simply draws the shortest path from the ship to the planet, which may be enticing to participants as it is the only automation strategy in early game that performs better than they do. In late game, as the environment becomes more crowded with ships and routes, the shortest path method may not be as viable. These paths are often drawn through the center of the screen and are thus prone to intersection and crashes. Unlike similar automation which mimics the participants play, line automation may be seen as harmful to the player’s route management strategy when taskload has increased. The ability of the similar and dissimilar automation to avoid the pitfalls of the shortest path strategy may make these automation strategies more enticing as taskload (and the need for automated assistance) increases. Thus, interpretations of compliance and reliance need to be made in context, as the preferred system design may vary depending on context. In this case, automation becomes more of a necessity at higher taskload, thus line automation would not be the preferred automation design, despite its predictability.

CONCLUSION

Our analysis has established a process that quantifies reliance and compliance based trust. We were able to elicit factors that affect compliance and reliance from a system, quantify them, and use participant data to create a general model for human trust in a human-automation team scenario. The ability to identify human trust in automated systems and design functions and models to predict trust will enable system designers to account for the effects of degraded trust on human-automation team performance. By quantifying trust, and accounting for factors that influence reliance and compliance rates, we can create more accurate human performance models and other developmental and operational test and evaluation platforms which will help build and appropriately calibrate operator trust in automated systems.

DISCLAIMER

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Air Force, Department of Defense, nor the U.S. Government.

REFERENCES

Bindewald, J. M., Peterson, G. L., & Miller, M. E. (2016). Clustering-Based Online Player Modeling. In T. Cazenave, S. Edelkamp, & M. Winands (Eds.), *International Joint Conference on Artificial Intelligence (IJCAI) - Computer Games Workshop*. New York, New York, USA.

Blanco, M., Atwood, J., Russell, S., Trimble, T., McClafferty, J., & Perez, M. (2016). *Automated Vehicle Crash Rate Comparison Using Naturalistic Data. Virginia Tech. Transportation Institute*.

Dixon, S. R., & Wickens, C. D. (2006). Automation reliability in unmanned aerial vehicle control: a reliance-compliance model of automation dependence in high workload. *Human Factors*, *48*(3), 474–486.

Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2006). On The Independence of Compliance and Reliance : Are Automation False Alarms Worse Than Misses? *Human Factors*, *49*(4), 564–572.

Dixon, S., Wickens, C., & Chang, D. (2005). Mission Control of Multiple Unmanned Aerial Vehicles: A Workload Analysis. *Human Factors*, *47*(3), 479–487.

Endsley, M. R. (2015). *Autonomous Horizons: System Autonomy in the Air Force - A Path to the Future* (Vol. 1).

Google. (2015). *Google Self-Driving Car Testing Report on Disengagements of Autonomous Mode*.

Kaminski, P. (2012). Role of Autonomy in DoD Systems, (July).

Lee, J. D., See, K. A., & City, I. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80.

Maynard, P. W., & Rantanen, E. M. (2005). Pilot Dependance on Imperfect Diagnostic Automation in Simulated UAV Flights: An Attentional Visual Scanning Analysis. *13th International Symposium on Aviation Psychology, Daytona, OH.*, 1–6.

Meyer, J., Gurion, B., & Sheva, B. (2004). Conceptual Issues in the Study of Dynamic Hazard Warnings, (December).

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society.*