Group Decision and Negotiation **13**: 173–189, 2004

The Influence of Task Load and Automation Trust on Deception Detection

DAVID P. BIROS
Adjunct Professor of IRM, Department of Systems and Engineering Management, Air Force Institute of Technology, 2950 Hobson Blvd., Wright-Patterson AFB, OH 45433, USA (E-mail: David.Biros@pentagon.af.mil)

MARK DALY
33rd Information Operation Squadron, Lackland AFB, TX 78243, USA (E-mail: Mark.Daly@lackland.af.mil)

GREGG GUNSCH
Assistant Professor of Electrical Engineering, Department of Engineering and Computer Science, 2950 Hobson Blvd., Wright-Patterson AFB, OH 45433, USA (E-mail: Gregg.Gunsch@afit.edu)

Abstract

The purpose of this research was to investigate the effects that user task load level has on the relationship between an individual's trust in and subsequent use of a system's automation. Military decision-makers trust and use information system automation to make many tactical judgments and decisions. In situations of information uncertainty (information warfare environments), decision-makers must remain aware of information reliability issues and temperate their use of system automation if necessary. An individual's task load may have an effect on his use of a system's automation in environments of information uncertainty.

It was hypothesized that user task load will have a moderating effect on the positive relationship between system automation trust and use of system automation. Specifically, in situations of information uncertainty (low trust), high task load will have a negative effect on the relationship. To test this hypothesis, an experiment in a simulated command and control micro-world was conducted in which system automation trust and individual task load were manipulated. The findings from the experiment support the positive relationship between automation trust and automation use found in previous research and suggest that task load does have a negative effect on the positive relationship between automation trust and automation use. Experiment participant who incurred a higher task load exhibited an over-reliance on their automated information systems to assist them in their decision-making activities. Such an over-reliance can lead to vulnerabilities of deception and suggests the need for automated deception detection capabilities.

As the world becomes more technologically advanced, society becomes more accustomed to technology and automation as it becomes routine and integral in our lives. We tend to trust the technology, or as Barber (1983) suggests, we gain expectations of technical competence in the technology. The increasing reliance and use in automation and technology has led researchers to examine the many aspects of human-computer interaction (Dillion and Morris 1996; Murray and Caldwell 1999; Parasuraman 1987; Wickens 1999). Trust in automation is one area that continues to generate interest among researchers (Fields 2001; Jian, Bisantz, Drury, and Llins 1998; Lee and Moray 1992; Muir 1994; Muir and Moray 1996; Sheridan 1988; Tseng and Fogg 1999).

Research has suggested that trust can affect how people accept and rely on automated systems to support their decision making processes (Sheridan 1988). For example, researchers have studied issues of human trust in simulated automated environments in which they found that an operator's decision to use automatic or manual control of a process depended on the trust he had in the system's automation and his confidence in his own abilities to control the system (Lee and Moray 1994; Muir and Moray 1996). Others have suggested that people become vulnerable to negative consequences because of their trust in information systems (Bonoma 1976; Giffin 1967). These vulnerabilities will likely increase as society continues its growing use of and reliance on information systems and automation. As individuals place great trust in the information systems they use, they will be exposing themselves to greater opportunities for deception by their adversaries.

The purpose of this study was to begin to answer the following two questions: First, is there a relationship between users' perceived trust in an automated system (i.e., automated information systems) and their subsequent use of the automation? Second, does task-load moderate the relationship between trust and automation use in a deceptive environment? For example, in an information warfare environment (Denning 1999), ones opponent might choose to manipulate information feeding a decision support system. Both questions were framed in an environment in which trust level was degraded due to reports of errors in system reliability and data was intentionally manipulated in order to deceive the user.

Human trust in automated systems

Nass, Fogg and Moon (1996) found that the trust relationship between humans and computers is similar to the trust relationship that humans have between each other. Many social rules and dynamics, which guide behavior in human interactions, also apply to human-computer interactions. Given this finding, it may be assumed that antecedents of trust in human-computer interaction are likely to be the same as antecedents of trust between humans as information systems may be considered conduits of human interaction. Definitions of trust, as they apply in human-computer interactions, are drawn from definitions of trust developed to apply to human relationships (Muir 1996). The word trust is so frequently used in our everyday language that most sources assume the audience knows what it means and thus it is rarely defined. When it is defined within scholarly literature, however, definitions form a wide range of meanings (McKnight and Chervany 1996). Before talking specifically about research regarding trust in automation, a brief review of the meaning and concepts of trust is presented.

Zuboff (1988) studied how people trust automated systems in the workplace. Her research found that people tended to distrust the technology of the automated system and thus used the system less, or that they tended to over-trust the system, which resulted in other problems when the system failed. Other empirical studies, consistent with Zuboff's, have shown that people's strategies with regards to the use or non-use of automated aids may be affected by their trust in the system. Muir (1987) developed a hypothetical model of human–machine trust, which consisted of the linear combination of characteristics that Barber (1983) outlined. Muir's model depicted human trust as the combination of persistence

of predictable behavior, technically competent performance, fiduciary responsibility, and the interaction between these characteristics.

By conducting a series of experiments on a continuous chemical process control simulation, Lee and Morray (1992) and Muir and Morray (1996) extended earlier work by Muir and developed a dynamic model of trust. The model depicts that an individual's current level of trust is affected by his previous level of trust as well as system factors such as the existence of automation faults (i.e., inaccurate readings by automated sensors) and the degradation system performance. Further, it incorporates the additional characteristics of predictability, dependability, and faith but relates them to the human–computer trust relationship developed by Muir. The studies found that workers monitoring automated systems became complacent when the system was perceived to perform correctly, and that workers who perceived an automated system to be prone to errors spent more time monitoring the system. In addition, the studies demonstrated that an operator's decision to use the automation or manual controls depended on his perceived reliability of the automated system (trust in the system), as well as his perceived reliability of manual control (trust in self) to manage the system. There was a very high correlation between an individual's trust in the system and the use of automation. These studies also produced evidence that once an individual perceives an error in the automation, his trust in the system will degrade for a period of time and then gradually rebound over time.

Automation bias

More recent studies in the use of automation are consistent with Zuboff's (1988) study. The work of Muir and others describes the phenomenon called automation bias (e.g., Mosier, Skitka, Heers, and Burdick 1997, 1998; Skitka, Mosier, and Burdick 1999). Automation bias is the " tendency to use automation as a heuristic replacement for vigilant information seeking and processing" (Skitka, Mosier, and Burdick 1999, p 704). In other words, the tendency for a decision-maker to over-rely on automation to perform tasks and make decisions rather than using the automated system as an aid is one component of the decision-making process. Theses studies identified two classes of errors that routinely emerge in highly automated decision environments: omission errors and commission errors. Omission errors are defined as, "failures to respond to system irregularities or events when automated devices fail to detect or indicate them," and commission errors as "errors which occur when people incorrectly follow an automated directive or recommendation, without verifying it against other available information" (Skitka and Moiser 1999, p. 344). A Conejo and Wickens (1997) study involving an Army threat target recognition tool, provides a good example of a commission error. They found that on occasion when an automated cue was unreliable (i.e., directing attention to something that was not the designated target), pilots were still very likely to choose the non-target as the target, despite the fact that the true target was known to the pilot and visible on the system display. These studies provide evidence that automation bias exists and may be due to excessive reliance on trusted automated systems.

Situational awareness and task load

Situational Awareness (SA) is the decision-maker's moment-by-moment ability to monitor and understand the state of a complex system and its environment. The completeness and accuracy of decision-makers situational awareness' is crucial to the ability to make decisions during emergencies (Wickens 1998). To maintain an accurate SA, the decision-maker should take into account both information that is available and that which can be activated from memory (Lyons 2000). In a high task environment, when a decision-maker is confronted with several issues, memory load can quickly become overwhelmed (Lyons 2000). This memory overload can cause an individual to begin to dismiss important cues, existing and past, from the environment (Weick 1995). The presence of an uncertain environment, say an information attack, may be one such critical cue overlooked by an individual in a high task environment. The dismissal of which may result in a vulnerability to deceptive information and an undesirable decisions being made.

Task load may also play a role in the negative effects of automation bias by individuals committing automation commission errors, i.e., errors made when an individual takes an inappropriate action due to over reliance on automated information or direction. As task load increases, individuals may rely more on automation, even in situations in which the automation may not be reliable. Skitka, Mosier and Burdick (1999) studied the use of automated monitoring aids in situations where information presented may be unreliable and found that task load had an effect on whether individuals used system automation or not. As the demands of verifying the information increased, individuals decreased their verification efforts and used the system's automation more. In general, when given a choice, individuals tend to prefer options that require lower investments in terms of attention and effort. Based on a summary of the social facilitation and arousal literature, Weick (1995, p. 102) contends that as "arousal (i.e., workload) increases, people tend to abandon recently learned responses and categories and fall back on earlier, over learned, often simpler responses." When individuals come to trust system automation, and become accustomed to using it, a high task load environment may cause them to overuse the automation even though it may not be reliable.

Research model, construct definition, and hypotheses development

The literature on human-human trust relationships; human-computer relationships and human-computer trust relationships in adversarial relationships all provide possible trust models from which hypotheses may be proposed. The hypotheses proposed are based on a model composed of relationships taken from Muir and Morray's (1996) dynamic model of trust, Llinas et al.'s (1998) lens model, Fields' (2001) adapted model of trust . They are depicted in Figure 1.

In this model, dependability and predictability contribute to an individual's overall trust in a system's automation. System trust then leads to use of system automation. User task load acts as a moderating variable between system automation trust and automation use and was used to test the effect this construct has on a decision-makers use of available

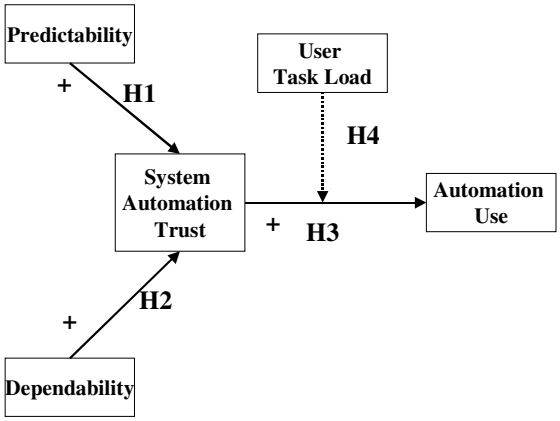


Figure 1. Developed Trust and Use Model.

automation. With this model, *dependability* is defined as having the trait of being dependable or reliable; yielding the same or compatible results in different experiments or statistical trials. *Predictability* is defined as the ability to foretell (declare or indicate in advance) on the basis of observation or experience. *System automation trust* is defined as having confidence in and entrusting the system automation to do the appropriate action. *Task load* is defined as the level of workload of an individual throughout the course of the experimental simulation (i.e., the number of resources an individual is responsible for and the number of enemy resources that need to be accounted for during the experiment simulation.) For the purpose of this study, *automation use* is defined as using system automation features in lieu of manual system techniques (i.e., using a system feature to accomplish many similar tasks simultaneously rather than completing each task individually in a more deliberate manner).

Trust is a multidimensional and dynamic construct that changes over time (Muir 1987, 1994; Remple, Holmes, and Zanna 1985). It starts with the foundation of predictability built on observable factors, which, demonstrated over time, lead to a perception of dependability. In the latter stages, because one cannot determine dependability of future behavior that has not been exhibited, trust depends on a leap of faith. In human-computer relationships, faith is based on the past perceived predictability and dependability (Lee and Morray 1992; Muir 1987; Muir and Morray 1996). In an information attack situation, it appears intuitive that observable and situational indications of information warfare (IW) activities would decrease the perception of predictability and dependability.

Hypothesis 1: Perceptions of predictability of system automation are positively related to trust.

Hypothesis 2: Perceptions of dependability of an automated information system are positively related to trust.

Individual users who trust technology are more inclined to utilize it for the purpose and in the manner in which it was designed (Lee and Moray 1994; Muir 1987; Seong, Llinas, Dury,

and Bisantz 1999). The introduction of automated technology has changed the roles of human operators from that of direct computer control to management of differing levels of computer control. Individuals must know how to interact with system automation, know when to rely on it and know when to intervene in the process when it is suspect (Seong and Bisantz 1998). Sheridan (1980) emphasizes that an individual's trust in automation plays a key role in determining the level of reliance a user places on automation. It has been demonstrated that low trust in automation delays its use (Riley 1996). This study therefore, proposes the following:

Hypothesis 3: Trust in system automation is positively correlated with use of system automation.

Even if Hypothesis 3 is supported, there may be circumstances when automation use will not decline even when trust in the system automation is suspect. One of these circumstances may be that of user task load. As the task load increases and more environmental cues are being interjected into the environment, a decision-maker may resort to using the automation as a means of keeping up with the environment. The increased task load may be causing a decreased state of situational awareness and thus environmental cues, such as information attack indicators, may be forgotten or ignored (Weick 1995). Thus, it is expected that:

Hypothesis 4: User task load has a moderating effect on the relationship between system automation trust and use of system automation.

The model, construct definitions and hypothesis development set the foundation of this study. They are the basis of a study to determine the effects task load and information system reliability have on an individual's use of automated information systems for decision-making support. The hypotheses were test in a laboratory setting as described below.

Experiment design

In order to investigate an individual's trust and subsequent use of automation, an experiment was designed around a military command and control (C2) scenario that was used with a high-fidelity computer simulation, the AWACS Weapons Director Trainer (AWDT). The objectives of command and control are to identify targets in an area of operations and coordinate the attack of those target. This system allows subjects to be immersed into a military command and control (C2) micro-world. Computer simulations provide a conduit between laboratory and field experiments by providing a more realistic and natural environment. The use of micro-world simulations provides for greater experimental control. Despite being conducted in a laboratory setting, the AWDT system simulates a real-world decision-making environment that may be experienced by weapons directors on board E-3 Airborne Warning and Control System (AWACS) aircraft. Several crews aboard the AWACS coordinate their efforts to provide airborne surveillance, and command, control and communications functions for tactical and air defense forces. Weapons directors are responsible for directing airborne assets, detecting, identifying, tracking and intercepting

airborne threats, as well as conducting search and rescue missions, should the need arise. Thus, effective decision-making is of paramount importance. The AWDT system allows for the collection of quantitative measures over the course of each experimental session as well as providing a mechanism for collecting measurable attitudes an beliefs through survey questionnaires. The system also has the capability to capture and record actions of participants as well as measure an individual's task load throughout a simulation session.

Participants consisted of 40 military graduate students with military ranks of 2nd Lieutenant through Major along with Junior and Senior members of an Reserve Officer Training Corps (ROTC) detachment at university in the mid-western US, with education levels ranging from, undergraduate to graduate. They were randomly assigned to four treatment groups to be discussed later. Both female (approximately 25 percent) and male participants participated. A majority of participants (93 percent) liked using computers and were comfortable with their use in the Air Force. All participants indicated in a post simulation evaluation sheet, that the training provided was sufficient to use the system and that the simulation was easy to understand. In addition, over half the participants indicated verbally that they would like to "play the game again." These comments were in line with the experimenter's observation that all subjects appeared engaged and focused on the task during the treatment scenarios. All participants arrived in military uniform to help in portraying a military environment. Since the AWDT system is able to score and record participant performance, the participant with the highest score was given extra credit by one of his/her course instructors.

Experiment manipulations

The first experimental manipulation was the construct called Information Warfare (IW). It has been shown that indications of IW may reduce the level of trust individuals have in the automated system they are using (Bisantz et al. 2000; Fields 2001). IW was operationalized by planting the idea of IW (i.e., successful computer attack) in the minds of participants via the scenario description. The scenario revealed that the system had been down for a time due to an IW attack, but presently the system was up and working but the reliability of systems recommendations were in question. Treatment groups two and four were subjected to the IW manipulation during the simulation. Groups one and three read a similar scenario, but no IW or equipment problems were indicated. The second manipulation, User Task Load, was operationalized by increasing the number of resources a participant was responsible for, along with the number of resources used by the attacking force. Task load was increased by a factor of approximately 2.5. Treatment groups one and two were subjected to a low task load, while treatment groups three and four were subjected to high task loads. The load measure file, automatically generated during the simulation, verified the task load of the individuals. It allow the participants to indicate how much of a task load they felt they were incurring. Thus it served as a manipulation check. Of the 20 individuals in the low-task load groups, 90 percent indicated on the post simulation evaluation sheet that they felt the scenario was low-task load while of the 20 individuals in the high-task load groups, 85 percent indicated they felt the scenario was high task load.

In order to test the hypotheses, an experiment was designed that manipulated two independent variables, Information Warfare (IW) based on Denning's (1999) theory of IW and User Task Load (described below). Implied IW provided an environment in which system automation can be in question, thus providing a basis for potential decreased trust in the system automation. These variables were crossed in a 2 × 2 configuration.

Each participant was given training on the simulator concept and computer interface. After completion of training, each participant was tasked to perform a hidden-profile, decision-making task that involved controlling multiple aircraft types to defend an area of operation and attack when possible in a simulated battle space. Control of the aircraft types was performed through various user actions on the AWDT system. The AWDT system was described to the subjects as a new AWACS component undergoing initial testing by Air Force Research Laboratories. A screen shot of the AWDT interface is depicted in Figure 2.

Participants were tasked to direct air assets against enemy assets with the aim of eliminating all threats in the area of operation. In addition, participants had the opportunity to attack enemy positions as resources allowed. They had the option to direct their assets using manual controls or to accept system recommendations that then automatically directed their assets. Manual direction of assets requires a minimum of three mouse clicks and manual positioning using the mouse. Automatic direction of assets could be accomplished by either clicking each visual representation of the current recommendations within an allotted time frame or by clicking a menu item that will accept all current active recommendations. (i.e., more that one recommendation can be accepted at a time.)

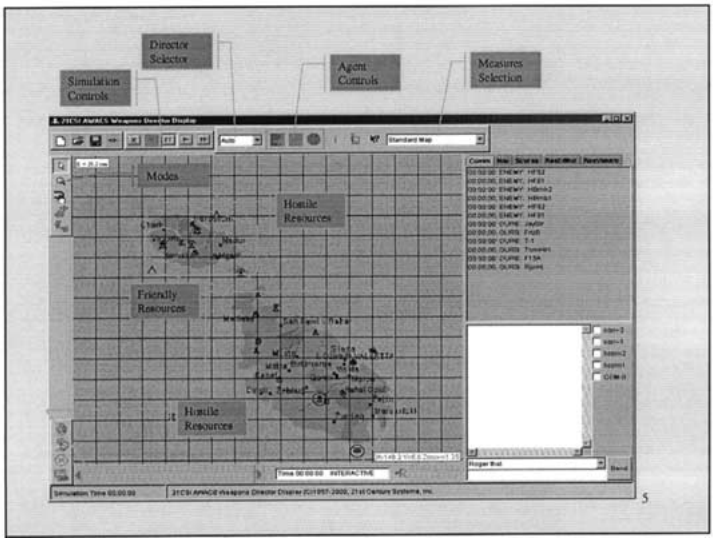


Figure 2. AWDT Interface.

The participants were told that system recommendations are derived from sophisticated algorithms designed with the aid of experienced AWACS Weapon Directors. Participants can assume that recommendations are designed to best utilize existing resources while attempting to eliminate the highest occurring threat. They were told that they could view their scores at anytime during the scenario. Scores were generated automatically by the system by taking the weighted sum of all eliminated enemy assets minus the weighted sum of the subject's assets destroyed. A positive score indicates that the value of enemy resources destroyed is greater than the value of the participant's resources destroyed while a negative score indicates the opposite.

Dependent measures

As suggested earlier, system predictability and dependability will affect users' perceived level of trust in a system's automation, and perceived level of trust should impact their subsequent use of the system's automation. In addition, it was hypothesized that an increase in user task load would have a moderating effect on the relationship between automation trust and user use of system automation.

Because cognitive phenomena like attitudes, motivations, expectations, intentions, and preferences are difficult to observe, a questionnaire was used to measure the specific constructs of interest including trust, predictability, and dependability. This questionnaire is a derivative of one that was developed specifically for the use of measuring the above mentioned constructs, as well as others, by actual weapons directors using the AWDT system. Credibility of the original questionnaire was established using a Q-Sort analysis using six subject matter experts. The reliability analysis produced an alpha = 0.75 for predictability and 0.85 for trust. (Hoffman 2000). All constructs were measured on a 6-point Likert-type scale in an attempt to force agreement or disagreement with each item and avoid neutrality. The scale ranged from 1 (Very Strongly Disagree) on the left to 6 (Very Strongly Agree) on the right. Questions one through seven dealt with the construct of trust, questions eight through ten dealt with the construct of predictability, and question eleven dealt with the construct of dependability. In addition to the Cronbach's Alpha, a factor analysis was completed on the questionnaire data and verified two constructs were measured; trust and predictability, as was desired. The survey was administered after the second training session and again after the final experimental simulation. The survey administered after the second training session served as a baseline to measure the overall trust individuals placed in the system's automation before any treatment was applied.

Automation use was measured by determining the number of times a subject accepted system recommendations versus the number of recommendations issued. The act of depending largely on the system may be an indicator and measure of trusting behavior. The AWDT system allows for the collection of this measure through automated data capture. The system can provide information on the number of recommendations given, whether the recommendations were accepted, and the manner in which they were accepted. (i.e., an individual recommendation was accepted or a group of recommendations were accepted.)

Analysis and results

Pearson product-moment correlations tested hypotheses 1, 2, and 3. Analysis of variance (ANOVA) was used to test the difference in automation use between the environmental conditions of IW and Non-IW (high and low trust environments) as well as between the various participant task loads. In addition, ANOVA was used to determine if user task load has a moderating effect on the relationship between system automation trust and use of system automation.

Predictability and dependability as related to trust (H1, H2)

Hypothesis H1 predicts a positive correlation between perceptions of predictability of system automation with perceptions of trust in system automation, while hypothesis H2 predicts a positive correlation between perceptions of dependability of system automation with perceptions of trust in system automation. A review of the correlation analysis in Table 1 shows a statistically significant and strong positive correlation between perceived predictability and trust, and between perceived dependability and trust, both at a significance level of $p < 0.01$. Values are shown for both pre-and post-treatment measures with the pre-treatment measure based on the training session. The Cronbach's alpha for the Trust items was alpha = 0.7967 for the pre-treatment and 0.9244 for post-treatment. Alphas for pre- and post-treatment Predictability items were 0.7528 and 0.7421, respectively. A Cronbach's analysis was not accomplished for dependability due to the single question asked regarding this construct.

The findings in Table 1 support both Hypothesis 1 and 2 and suggest that as ratings of perceived predictability and dependability in system automation rise, so too, do ratings of trust in system automation.

Relationship between system automation trust and system automation use (H3)

Hypothesis 3 predicts that an individual's perception of trust in system automation will be positively correlated with his use of system automation. Automation use was measured as

Table 1. Pearson Product-Moment Correlations of Trust with Predictability and Dependability

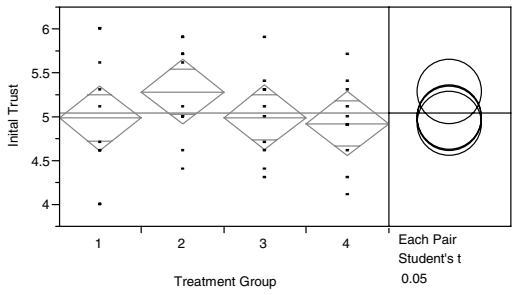
Constructs	Time Frame	r
Predictability	Pre-treatment	0.4759*
	Post-treatment	0.6756*
Dependability	Pre-treatment	0.5734*
	Post-treatment	0.6878*

*Significant at $p < 0.01$.

a ratio between the number of system recommendations given and the number of recommendations accepted by an individual. The higher the ratio, the greater the automation was used. A correlation analysis indicated a statistically significant and strong positive correlation ($r = 0.6839$) between ratings of trust in system automation and automation use at a significance level $p < 0.01$ using post-treatment trust and automation measures. This finding supports Hypothesis 3 and suggests that as a user's perception of trust in system automation increases so will his use of that system's automation.

Moderating effect of user task load on the relationship between system automation trust and system automation use (H4)

Hypothesis 4 predicts a user's task load will have a moderating effect on the relationship between system trust and use of system automation. The parametric results shown below in Figures 4, 5, and 6 provide evidence that task load has a moderating effect on the relationship between trust and automation use at a significant level of $a < 0.05$. Figures 4 and 5 are ANOVA plots with associated Tukey's Multiple Comparison results showing the mean trust levels of each treatment group before and after treatments were applied. Figure 3 indicates all participants show a high level of system automation trust pre-treatment with no statistically significant difference between treatment groups.



Oneway Anova

Summary of Fit

Rsquare	0.061285
Adj Rsquare	-0.01694
Root Mean Square Error	0.576363
Mean of Response	5.0525
Observations (or Sum Wgts)	40

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment Group	3	0.780750	0.260250	0.7834	0.5110
Error	36	11.959000	0.332194		
C. Total	39	12.739750			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	10	4.99000	0.18226	4.6204	5.3596
2	10	5.29000	0.18226	4.9204	5.6596
3	10	5.00000	0.18226	4.6304	5.3696
4	10	4.93000	0.18226	4.5604	5.2996

Std Error uses a pooled estimate of error variance

Means Comparisons

Dif=Mean[i]-Mean[j]

	2	3	1	4
2	0.000000	0.290000	0.300000	0.360000
3	-0.29	0.000000	0.010000	0.070000
1	-0.3	-0.01	0.000000	0.060000
4	-0.36	-0.07	-0.06	0.000000

Alpha= 0.05

Comparisons for each pair using Student's t

t

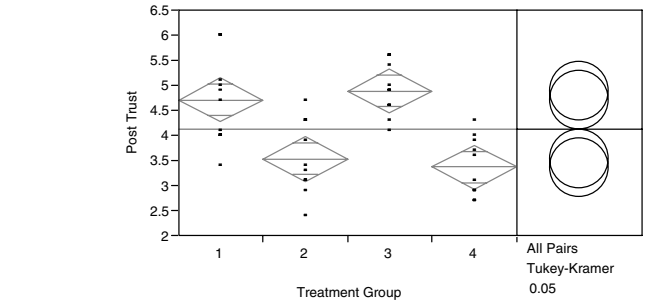
2.02809

Abs(Dif)-LSD

	2	3	1	4
2	-0.52276	-0.23276	-0.22276	-0.16276
3	-0.23276	-0.52276	-0.51276	-0.45276
1	-0.22276	-0.51276	-0.52276	-0.46276
4	-0.16276	-0.45276	-0.46276	-0.52276

Positive values show pairs of means that are significantly different.

Figure 3. Descriptive Statistics of Pre-Treatment Trust Measures.



Oneway Anova

Summary of Fit

Rsquare	0.520801
Adj Rsquare	0.480868
Root Mean Square Error	0.687911
Mean of Response	4.135
Observations (or Sum Wgts)	40

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment Group	3	18.515000	6.17167	13.0418	<.0001
Error	36	17.036000	0.47322		
C. Total	39	35.551000			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	10	4.72000	0.21754	4.2788	5.1612
2	10	3.54000	0.21754	3.0988	3.9812
3	10	4.90000	0.21754	4.4588	5.3412
4	10	3.38000	0.21754	2.9388	3.8212

Std Error uses a pooled estimate of error variance

Means Comparisons

Dif=Mean[i]-Mean[j]

	3	1	2	4
3	0.00000	0.18000	1.36000	1.52000
1	-0.18000	0.00000	1.18000	1.34000
2	-1.36000	-1.18000	0.00000	0.16000
4	-1.52000	-1.34000	-0.16000	0.00000

Alpha= 0.05

Comparisons for all pairs using Tukey-Kramer HSD

q*
2.69323

Abs(Dif)-LSD

	3	1	2	4
3	-0.82855	-0.64855	0.53145	0.69145
1	-0.64855	-0.82855	0.35145	0.51145
2	0.53145	0.35145	-0.82855	-0.66855
4	0.69145	0.51145	-0.66855	-0.82855

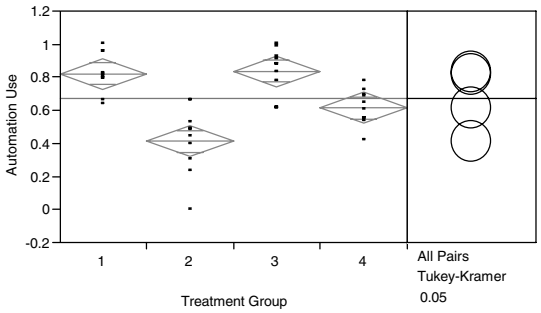
Positive values show pairs of means that are significantly different.

Figure 4. Descriptive Statistics of Post-Treatment Trust Measure.

However, the evidence shown in Figure 4, does allow for the claim that treatments 1 and 3 are statistically significantly different from treatments 2 and 4 in their associated mean values of system automation trust. That is to say, the treatments in which information provided was not in question (i.e., non-IW) are significantly different from the treatments in which information was in question (i.e., IW) in terms of overall system automation trust.

This information taken together with the results of hypothesis 3, which suggests that trust and automation use are positively correlated, could lead one to conclude that the levels of automation use between treatment groups 1 and 3 and between treatment groups 2 and 4 would show no statistically significant difference. Figure 5 below, shows that there is indeed no statistically significant difference in automation use between treatment groups 1 and 3, (high trust groups) but does show a statistically significant difference, although minor, between groups 2 and 4 (low trust groups). In should be noted that the confidence level for the entire set of comparison means is 95% (alpha of 0.05), but the confidence level for any particular comparison (i.e., treatment 2 and 4) is larger than 95% as the Tukey method uses an experiment-wise error rate rather than a pre-comparison error rate (Devore, 2000).

These findings support Hypothesis 4 and suggest that despite perceptions of low system automation trust, individuals tend to use automation more when task loads increased. This is an issue of concern especially in conditions where the deceptive information is present.



Oneway Anova

Summary of Fit

Rsquare	0.609655
Adj Rsquare	0.577126
Root Mean Square Error	0.145543
Mean of Response	0.675258
Observations (or Sum Wgts)	40

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Treatment Group	3	1.1910235	0.397008	18.7420	<.0001
Error	36	0.7625792	0.021183		
C. Total	39	1.9536027			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1	10	0.824020	0.04602	0.73068	0.91736
2	10	0.418150	0.04602	0.32481	0.51149
3	10	0.841230	0.04602	0.74789	0.93457
4	10	0.617630	0.04602	0.52429	0.71097

Means Comparisons

Dif=Mean[i]-Mean[j]

	3	1	4	2
3	0.000000	0.017210	0.223600	0.423080
1	-0.01721	0.000000	0.206390	0.405870
4	-0.2236	-0.20639	0.000000	0.199480
2	-0.42308	-0.40587	-0.19948	0.000000

Alpha= 0.05

Comparisons for all pairs using Tukey-Kramer HSD

q*
2.69323

Abs(Dif)-LSD

	3	1	4	2
3	-0.1753	-0.15809	0.048301	0.247781
1	-0.15809	-0.1753	0.031091	0.230571
4	0.048301	0.031091	-0.1753	0.024181
2	0.247781	0.230571	0.024181	-0.1753

Positive values show pairs of means that are significantly different.

Figure 5. Descriptive Statistics of Post-Treatment Automation Use.

Discussion

This study examined the interactions between humans and computer automation and the effects the humans' perceptions of trust played in the way they utilized this automation. It also analyzed how workload affected the relationship between trust and system automation use. As predicted, there is evidence to suggest the factors of perceived predictability and dependability in a system's automation are positively correlated with ratings of trust in the system's automation, and that this associated level of trust is positively correlated with the level of system automation use by an individual. Therefore, in order to encourage the use of system automation and provide the potential for decreased decision time for decision-makers, system designers and developers must provide systems that provide a high level of system predictability and dependability.

There also appears to be evidence to suggest that task load may play an important role in modifying the level of use of system automation when trust in the automation is low, such as in conditions in which information warfare activities are or have been occurring. This finding may be significant in settings in which commission errors caused by overuse of automation in unreliable situations may cause sub-optimal decision-making.

Implications

The implications of findings are considered in two areas: research and application. In regards to research, the findings in this study are encouraging enough to continue this stream of research. When organizations reduce their workforces, automation may be used to offset the decrease in personnel and possibly put an increasing workload on the remaining individuals. It would be beneficial to determine a suitable range within which an individual can maintain an appropriate level of situational awareness so factors such as unreliability of system automation remain part of the decision maker's environmental cues used to make critical decisions. This suitable-range theory could be accomplished in future research by varying the workload over time and seeing how the same individual reacts in the different reliability scenarios. Because this study dealt with inexperienced individuals in regards to weapons directors duties, it would be important to determine if these results hold true for actual weapons directors in a more realistic environment. Experience level and confidence in one's own abilities have been shown to affect the level of trust an individual places in a system and plays a role in the type and amount of automation an individual uses (Lee and Morray 1994). Past research has shown that trust is a dynamic construct and that once lost can be regained over time (Seong et al. 1998). Because environmental conditions can reduce the level of perceived system trust, and therefore, system automation use, further research is necessary in this area to investigate what actions can facilitate the regaining of an individuals trust in the system once perceived trust is reduced.

This research may be useful in developing better information security measures. From a defensive standpoint, operational protocols could be developed such that in times of uncertainty, increasing personnel to limit the effects of over-use of system automation could reduce individual workload. Also, it may be possible to maintain an individual's situational awareness regarding uncertainty in a system's performance by providing on-screen indications of the uncertainty and system reliability. These cues may help reduce the amount of automation use by individuals and thus, fewer commission errors would be committed. Finally, over-reliance on information technology could lead to opportunities for system users to be deceived and result in an undesirable decision making environment. As such, the incorporation of an automated system to detect potential deception certainly appears to be a worthy endeavor.

Implications for automating deception detection

This research demonstrated that individuals rely on information technology even when its veracity is in question and thus, opportunities for deception are apparent. Biros et al. (2003), demonstrated the how systems users were deceived while continuing to place their trust in compromised information systems. Therefore, it stands to reason that systems to aid decision-makers in determining the veracity of the information provided by their decision support technology would be quite valuable. An automated deception detection system based on some form of cue or heuristics technology that would aid the decision-making at judging the accuracy of the information present in times of high task load appears to be the next

logical step. This research demonstrated just how trusting people are in the information systems they use, thereby underscoring the need for a deception detection capability.

Research limitations and advantages

Despite the evidence in support of these results, certain limitations must be acknowledged when considering these findings. With regard to design and internal validity issues, repeated testing and instrumentation may be a threat to the internal validity. The same trust measurement questionnaire was given immediately before and immediately after the treatment scenario. This may have led at least some of the participants answer trust questions based on their previous answers realizing that this was a test measure and they wanted to be consistent.

It is important to note that, while the respondents would typically not conduct these tasks in a typical daily setting, they were military members stationed on a military installation. Thus, the military culture of getting the job done and strong ethic toward winning the battle, aid in the realism of the study. The respondents took the tasks of the experiment quite seriously and when the task load increased, they tried even harder in spite of the potential for deception. This is a clear indicator that a deception detection system would be most useful in this setting.

Concluding remarks

The evidence derived from this research suggests an individual's use of a system's automation capability is directly and positively related to the level of perceived trust the individual places in that system's automation. In addition, an individual's task load may have a moderating affect on the relationship between user trust and automation such that, during times of increased task load, a decision-maker may resort to using the system automation despite a lower level of perceived trust in the system's automation. These results have important implications for the organization dependent on automated decision support systems. Over-dependence may cause an increased occurrence of automation commission errors causing undesired and possibly catastrophic effects. Further research is warranted and should draw upon the results presented in this study.

References

Barber, B. *The Logic and Limits of Trust.* New Brunswick NJ: Rutgers University Press, 1983.

Biros, D., G. Fields, and G. Gunsch, G. "The Effect of External Safeguards on Human-Information System Trust in an Information Warfare Environment," *Proceedings of the 36th Hawaii International Conference on Systems Sciences (HICSS)*, 2003.

Bonoma, T. V. (1976). "Conflict, Cooperation, and Trust in Three Power System," *Behavioral Science* 21 (6), 499–514.

Conjeo, R. and C. D. Wickens. (1997). "The Effects of Highlighting Validity and Feature Type on Air-to-Ground Target Acquisition Performance," University of Illinois Institute of Aviation Technical Report ARL-97-11/ NAWC-ONR-97-1.

Denning, D. E. (1999). *Information Warfare and Security*. Reading MA: Addison-Wesley.

Devore, J. L. (2000). *Probability and Statistics for Engineering and the Sciences*. Pacific Grove CA: Duxbury.

Dillon, A. and M. G. Morris. (1996). "User Acceptance of Information Technology: Theories and Models" *in* M. E. Williams (Ed.), *Annual Review of Information Science and Technology (ARIST)*. Medford NJ: *Information Today* 31, 3–32.

Fields, G. S. (2001). *The Effect of External Safeguards on Human-Information System Trust in an Information Warfare Environment.* MS thesis, AFIT/GIR/ENV/01M-07. School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March.

Giffin, K. "The Contribution of Studies of Source Credibility to a Theory of Interpersonal Trust in the Communication Process," *Psychological Bulletin* 68 (2), 104–120.

Hoffman, K A. (2000). *Trust and Performance with Intelligent Agent Technology: Implications for Human Interactions.* MS thesis, Department of Psychology, University of South Florida.

Jina, J.-Y. et al. (1998). *Foundations for an Empirically Determined Scale of Trust in Automated Systems*. Center for Multisource Information Fusion no. CMIF-1-98, State University of New York at Buffalo.

Lee, J. and N. Moray. (1994). "Trust, Self-Confidence, and Operators Adaptation to Automation," *International Journal of Human-Computer Studies* 40 (1), 153–184.

Lee, J. and N. Moray. (1992). "Trust, Control Strategies and Allocation of Function in Human-Machine Systems," *Ergonomics* 35 (10), 1243–1270.

Llinas, J. et al. (1998). *Studies and Analyses of Vulnerabilities in Aided Adversarial Decision Making: Final Report, 1 April 1996- 1 February 1997.* Contract AFRL/HE-WP-TR-1998-0099. Buffalo NY: State University of New York at Buffalo, February.

Lyons, M. D. (2000). "A Test Paradigm for War game 2000," unpublished paper. n. pag. http://www.dodccrp.org/Proceedings/DOCS/wcd00000/wcd000e2.htm August.

McKnight, D. Harrison, and N. L. Chervany. "The Meanings of Trust," Research working paper, n. pag. http://www.misrc.umn.edu/wpaper/wp96-04.htm. 26 October.

Mosier, K. L., L. J. Skitka, S. Heers, and M. D. Burdick. (1997). "Patterns in the Use of Cockpit Automation," *in* M. Mouloua and J. Koonce (Eds.), *Human-Automation Interaction*: *Research and Practice.* Hilsdale NJ: Lawrence Erlbaum Assoc. Inc., 167–173.

Muir, B. M. (1987). "Trust Between Humans and Machines, and the Design of Decision Aids," *International Journal of Man-Machine Studies* 27, 527–539.

Muir, B. M. (1994). "Trust in Automation: Part I. Theoretical Issues in the Study of Trust and Human Intervention in Automated Systems," *Ergonomics* 39 (3), 1905–1922.

Muir, B. and N. Moray. (1996). "Trust in Automation: Part II. Experimental Studies of Trust and Human Intervention in a Process Control Simulation," *Ergonomics* 37 (11), 429–460.

Murray, S. A. and B. S. Caldwell. (1999). "Operator Alertness and Human-Machine System Performance During Supervisory Control Tasks," *in* M. W. Scerbo and M. Mouloua (Eds.), *Automation Technology and Human Performance*. Mahwah NJ: Lawrence Erlbaum Associates.

Nass, C., B. J. Fogg, and Y. Moon. (1996). "Can Computers Be Teammates?" *International Journal of Human-Computer Studies* 45 (6), 669–678

Parasuraman, R. "Human-Computer Monitoring," *Human Factors* 29 (6), 695–706, December.

Rempel, J. K., J. G. Holmes, and M. P. Zanna. (1985). "Trust in Close Relationships," *Journal of Personality and Social Psychology* 49, 95–112.

Riley, V. (1996). "Operator reliance on Automation: Theory and Data," *in* R. Parasuraman and M. Mouloua (Eds.), *Automation and Human Performance*. Mahwah NJ: Lawrence Erlbaum Associates, 19–35.

Sall, J., A. Lehman, and L. Creighton. (2001). *JMP Start Statistics: A Guide to Statistics and Data Analysis*. Pacific Grove, CA: Duxbury.

Seong, Y., J. Llinas, C. G. Drury, and A. M. Bisantz. (1999). "Human Trust in Aided Adversarial Decision-Making Systems," *in* M. W. Scerbo and M. Mouloua (Eds.), *Automation Technology and Human Performance*. Mahwah, NJ: Lawrence Erlbaum Associates.

Seong, Y. and A. M. Bisantz. (1998). "Modeling Human Trust in Complex, Automated Systems Using a Lens Model Approach," Studies *and Analyses of Aided Adversarial Decision Making* (abstract) 95–100.

Sheridan, T. B. (1988). "Trustworthiness of Command and Control Systems," *IFAC Man Machine Systems* 427–431.

Skitka, L. J. and K. L. Mosier. (1999). "Automation Use and Automation Bias," *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*.

Skitka, L. J., K. L. Mosier, and M. Burdick. (2000). "Accountability and Automation Bias," *International Journal Human-Computer Studies* 52, 701–717.

Tseng, S. and B. J. Fogg. (1999). "Credibility and Computing Technology," *Communications of The ACM* 42 (5) 39–45, May.

Weick, K. E. (1995). *Sensemaking in Organizations*. Thousand Oaks: SAGE Publications Inc.

Wickens, C. D. (1999). "Automation in Air Traffic Control: The Human Performance Issue," *in* M. W. Scerbo and M. Mouloua (Eds.), *Automation Technology and Human Performance*. Mahwah NJ: Lawrence Erlbaum Associates.

Zuboff, S. (1988). *In the Age of the Smart Machine: The Future of Work and Power*. Oxford: Heinemann Professional.