

JATE

Journal of Aviation Technology and Engineering 6:1 (2016) 1–8

What Are the Predictors of System-Wide Trust Loss in Transportation Automation?

Stephen Rice, Scott R. Winter, John E. Deaton and Ismael Cremer

(Florida Institute of Technology–Melbourne)

Abstract

Prior research has examined how individuals place trust in single (e.g., Meyer, 2001, 2004) and multiple (e.g., Geels-Blair, Rice, & Schwark, 2013) automated devices when one fails. This has shown that participants are influenced by system-wide trust (SWT). What has been missing is an investigation into what types of people succumb to SWT effects. The current study attempts to replicate SWT findings and identify possible predictors of individuals likely to be influenced by SWT. The findings did demonstrate a replication of SWT. The study found that “*feelings of negativity when automated devices fail*” was a significant predictor of SWT theory.

Keywords: Automation, Trust, Trust in Automation, System-Wide Trust, Aviation

Introduction

Previous research in terms of system-wide trust (SWT) has studied how operator’s trust in an automated single device may be affected when it fails (Dixon & Wickens, 2006; Dixon, Wickens, & Chang, 2005; Dixon, Wickens, & McCarley, 2007; Geels-Blair et al., 2011; Keller & Rice, 2010; Lee & Moray, 1994; Parasuraman, Sheridan, & Wickens, 2000; Rice, 2009; Rice & Geels, 2010; Rice et al., 2008; Wickens & Colcombe, 2007; Wiegmann, Rich, & Zhang, 2001). The SWT effect is also seen when multiple devices are included in the study (Geels-Blair, Rice, & Schwark, 2013; Keller & Rice, 2010; Rice & Geels, 2010). Operators tended to group multiple devices together, and if one device failed, trust in the entire system was reduced and treated as unreliable.

The following sections will define SWT and discuss the concepts of trust in automation. This will support the need for the study and the implications that SWT theory has on consumer perceptions.

Automation and Trust

Wickens and Hollands (2000) define automation as a mechanical or electrical task of work that would aid or replace the human operator. There are four stages of automation that have been identified: synthesis, diagnosis, response selection, and response execution (Parasuraman, Sheridan, & Wickens, 2000). Automation can be used to assist a human operator with tasks that may be unsuitable for a human to accomplish. However, it has been identified by several researchers

(e.g., Meyer, 2001, 2004; Parasuraman & Riley, 1997; Rice, 2009) that automation errors can cause distrust and the system being ignored completely.

Social psychology defines trust as a construct that involves the predictability of another person (Deutsch, 1958; Eckel & Wilson, 2004; Ergeneli, Saglam, & Metin, 2007). Research has shown that this construct can be related in human interaction with automated systems (Reeves & Nass, 1996; Rice, 2009; Parasuraman & Riley, 1997). When a human places a high level of trust in another person, it is believed that they will accomplish what is expected of them, and this is viewed as a positive interaction (Lee & See, 2004). These interactions can be extrapolated to understand how humans would interact with machines, thus allowing the concept of trust to be applied when using automated systems and demonstrating how it varies when components do not operate as expected (Geels-Blair, Rice, & Schwark, 2013).

The increase in the use of automation over the last few decades also adds a factor of complexity. Operators can overcome the complexity of a system by augmenting supervision and using trust in automation. Trust and reliance in automation allow for adaptive behavior when complete understanding in a situation is impractical (Lee & See, 2004). The failure in automation when multiple aids are used is an important aspect that researchers are interested in studying (Keller & Rice, 2010; Rice & Geels, 2010).

Previous research has identified a positive relationship with respect to the shift in operator trust levels and the reliability of the automated aid. An operator will rely on the aid if a high level of trust exists (Parasuraman & Riley, 1997; Rice, 2009). Moreover, recent studies have shown that operators group multiple aids as one system and that if one aid is deemed unreliable, the operator is seen to treat all other aids as unreliable (Geels-Blair, Rice, & Schwark, 2013; Keller & Rice, 2010; Rice & Geels, 2010). This supports the SWT theory.

System-Wide Trust Theory

Prior research focused on single-aid devices and the relationship of automation trust (Maltz & Shinar, 2003; Meyer, 2001, 2004; Rice, 2009). The research involved pairing operators with only single automated devices. Keller and Rice (2010) examined whether participants had system-wide or component-specific trust when multiple aids were used in a system. Component-specific perception infers that an operator would place his or her trust separately between the different aids and would act according to the perceived reliability of the specific aid itself. If one aid failed, the operator would still continue to trust the other aid. SWT occurs when the operator disregards the reliability of an aid as an independent entity and treats the reliability based on the system as a whole. In other words, when one aid is

unreliable, the level of trust that the operator has toward the other aids would be adversely affected.

Keller and Rice (2010) tested these two theories by using a simulated flight task. In the experiment, the operator was assigned to check for system errors in two gauges while flying an unmanned aerial system simulator. The percentage of reliability varied in each condition by either being 100%, 85%, or 75% reliable. The choice to ignore or agree with the aid was up to the participant. In this study, the data supported the SWT theory and that one unreliable aid negatively affects the trust in the other reliable aids. The limitations of this study involved sample size and number of gauges. The small sample size was from the same university, thus limiting generalizability. Furthermore, the use of only two gauges meant that the unreliable aid accounted for 50% of the system.

A later study conducted by Rice and Geels (2010) included four gauges in the experiment to increase the ratio of perfect to imperfect aids from 1:1 to 3:1. A single-task paradigm was used. Moreover, misses were used instead of false alarms to investigate whether the system-wide affect still occurred. Similar results to the previous experiments were obtained, demonstrating that the participants were still influenced by the effects of SWT.

Another study involved the use of eight gauges to investigate SWT (Geels-Blair, Rice, & Schwark, 2013). The adverse effect of false alarms and misses on operator trust was examined. In this study, one of the eight gauges was either 100% or 70% reliable and included either false alarms or misses. The data depicted that an operator would still group the reliability of the system regardless of the number of automated aids, thus supporting SWT. Furthermore, this study revealed that false alarms correlated with a stronger SWT effect than misses. This study used a larger sample size, allowing for a more accurate effect size.

A recent study examined the level of trust in automation when a group of European participants were given a hypothetical scenario (Winter, Rice, & Reid, 2014). In the first scenario, a group of participants were asked to imagine that they were on a flight where the oxygen masks falsely deployed. They were then asked to rate their trust in the automation that operates the oxygen masks, autopilot system, airplane flaps, landing gear, and seat video monitor on a Likert-type scale from -3 , indicating extreme distrust, to $+3$, indicating extreme trust. A neutral option of zero indicated neither trust nor distrust. A separate control group was asked to rate their trust in the automation system given the same scenario without the failure.

The results showed that participants rated their trust high when the system did not fail; however, there was a clear drop in trust in all the automated devices for the condition where the oxygen mask deployed incorrectly. That is, trust in all the other automated devices was pulled down when compared to the control group, indicating that the SWT

effect was influencing participants' trust in the automated devices that were unrelated to the failed device.

Other research has also indicated that a level of trust and familiarity has a role in pilots' belief in automation reliability and its functioning capacity (Mehta, Rice, & Winter, 2014). While it is difficult to prove that SWT is a major contributor to particular accidents, there have been many incidents and accidents involving interaction with GPS tools in both the aviation and maritime industries (Johnson, Shea, & Holloway, 2008). A prime example of this was in November 2004 when a C-212-CC twin-engine aircraft flew into terrain due to lack of hazard monitoring and overreliance on GPS alarms.

Current study. Previous research has shown a strong effect of SWT across several paradigms. However, there is a missing gap in the literature that we hope to fill. Specifically, no study that we know of has looked at what types of people tend to succumb to SWT effects. The goal of this study is twofold. First, we would like to replicate the previous findings to ensure the robustness of the SWT effect in hypothetical situations using samples from different countries. Second, we wish to determine if there are any possible predictors of SWT and if those predictors replicate across two paradigms. In two studies, participants were given scenarios about an airplane flight or a rental car trip. In the experimental conditions, participants were told about an automated device failure. They were then asked to rate their trust in both the failed device and other unrelated devices located on the same airplane/car. We predicted that the SWT effects seen in previous studies would replicate to this paradigm and that we would find a significant predictor of SWT effects.

Study 1 Methods

Participants. Ninety-four participants (29 females) from the United States took part in the study via an online survey. The mean age was 33.93 ($SD = 10.11$).

Materials and recruitment. The study was presented online using FluidSurveys[®]. Participants were recruited via Amazon's[®] Mechanical Turk[®] (MTurk). MTurk is a global online service that enables participants (Turkers) to participate in human intelligence tasks (HITs) in exchange for monetary compensation. Participation in any HIT is voluntary and anonymous.

Procedure. Participants first signed an electronic consent form. They were then presented with the following scenario: "Imagine that you are flying on a 4-hour commercial airplane flight from one major city to another. Sometime during the flight, an alarm goes off throughout the cabin and oxygen masks fall from the compartments above passenger seats. Following this, the pilot comes on the intercom and says that there was a mistake and the automation that operated the oxygen masks failed. He says that there is no actual emergency and not to worry. The pilot then tells you the altitude of flight and how long it will be before you land." In a separate baseline condition,

participants were given the same scenario without the failed oxygen mask situation.

Participants were then asked to rate their trust in the automation that operates the oxygen masks, autopilot system, airplane flaps, landing gear, and seat video monitor on a Likert-type scale from -3 (extremely distrust) to $+3$ (extremely trust) with a neutral option of zero (neither trust nor distrust). Following this, participants were asked to provide the following demographic information: gender, ethnicity, age, income, number of flights per year, flight mileage per year, proficiency at interacting with automated devices, frequency of interaction with automated devices, knowledgeability about automated devices, feelings of negativity when automated devices fail, strength of fear of flying, and highest education level.

Design. Different participants were randomly placed in the experimental and control conditions, and all participants provided ratings for all the different automated devices.

Study 1 Results

Compliance with statistical assumptions. The data were analyzed for compliance with applicable regression and analysis of variance (ANOVA) assumptions, which included correct specification of the form of the relationship between the independent variables and dependent variable, correct specification of the independent variables in the regression model, reliability, constant variance of the residuals, independence of the residuals, and normality of the residuals. The data was found to have met all statistical assumption requirements.

Figure 1 presents these data. A two-way ANOVA using Condition as a between-participants factor and DeviceType as a within-participants factor revealed a main effect of Condition, $F(1, 92) = 53.10, p < .001, \eta p^2 = .37$; a main effect of DeviceType, $F(4, 368) = 40.13, p < .001, \eta p^2 = .30$; and a significant interaction between the two factors, $F(4, 368) = 38.16, p < .001, \eta p^2 = .29$. These data reveal that there was an SWT effect that replicates findings from previous studies.

Next, each participant's score in the failure condition was subtracted from the average score for all participants in the baseline condition for each automated device. These scores indicated how much trust had dropped from a situation where there was no failure to a situation where there was a failure. These difference scores were then used as the criterion variables in two regression analyses. In the first analysis, the difference score for the unreliable device was the criterion variable, with gender, ethnicity, age, income, number of flights, flight mileage, proficiency, frequency, knowledgeability, feelings of negativity, fear of flying, and highest education level used as possible predictors. We used stepwise regression to eliminate ineffective predictors. The resulting model included one of the

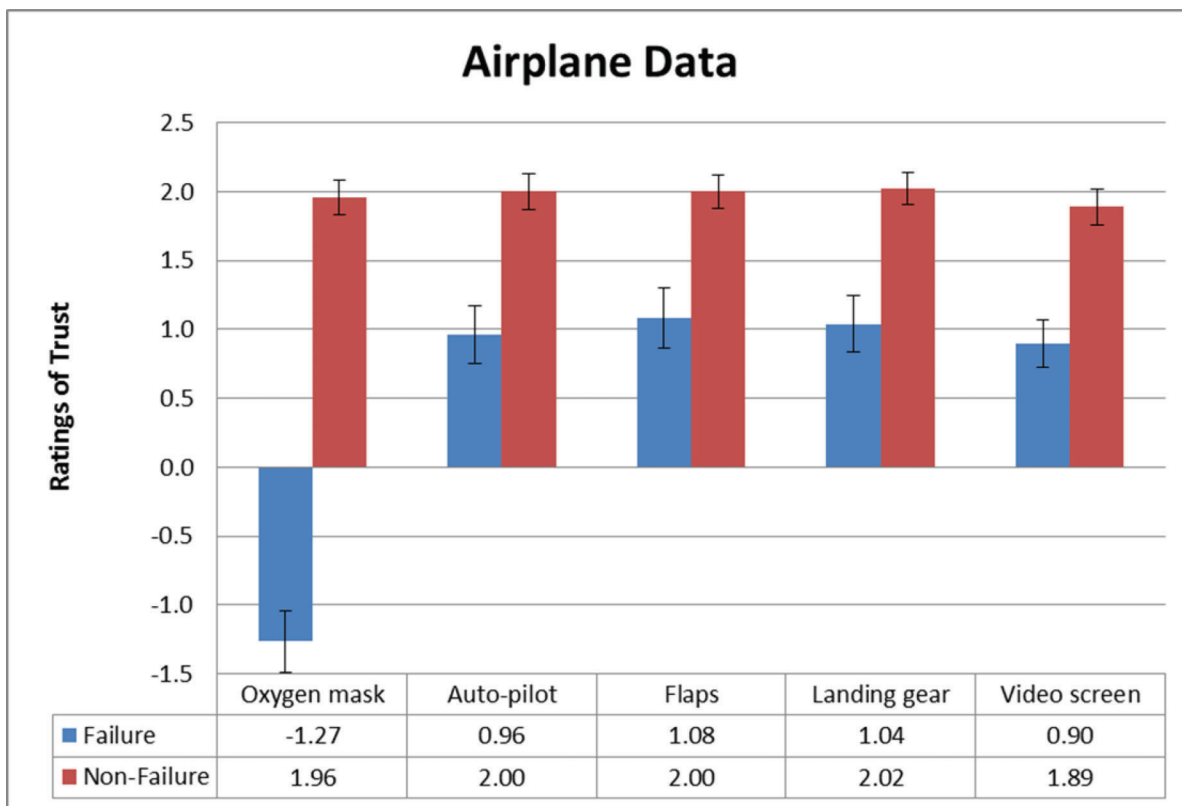


Figure 1. Data from Study 1. SE bars are included.

original predictors. *Feelings of negativity when automated devices fail* significantly predicted loss of trust scores, $\beta = .403$, $t(47) = 3.02$, $p = .004$, in the device that failed. This model accounted for 16% of the variance in the criterion, $F(1, 47) = 9.13$, $p = .004$. These results indicate that the stronger the feelings of negativity when automated devices fail, the more there is loss of trust when an automated device fails.

A second regression analysis was conducted using the average difference scores for the remaining devices (i.e., the devices that did not fail) with the same predictors. The resulting model again included one of the original predictors. *Feelings of negativity when automated devices fail* significantly predicted loss of trust scores, $\beta = .453$, $t(47) = 3.48$, $p = .001$, in the devices that did not fail. This model accounted for 21% of the variance in the criterion, $F(1, 47) = 12.14$, $p = .001$. These results indicate that the stronger the feelings of negativity when automated devices fail, the more there is a loss of trust during SWT loss.

Study 1 Discussion

The purpose of the study was twofold. First, we wanted to replicate findings from previous studies in a design using hypothetical scenarios. Our hypothesis was that the SWT effect would present itself even though the participants were only told hypothetical information about the devices. This prediction was supported by the

data. There was clearly a loss of trust in the unreliable device, and this loss of trust spread to the other devices despite participants not having any information about their reliability.

The second purpose of the study was to determine if any demographic predictors could help explain why participants tend to succumb to SWT effects. The data revealed one significant predictor. Participants who tended to feel more negatively about automation that fails also tended to lose more trust in the devices described in the scenario. We discuss the implications of this finding in more detail in the general discussion.

Study 2 Introduction

Study 1 presented both a replication of the previous findings and a possible predictor for SWT effects. The goal of Study 2 was to replicate this finding in a different paradigm. In this study, participants were given a scenario about a rental car.

Study 2 Methods

Participants. One hundred twenty participants (49 females) from the United States took part in the study via an online survey. The mean age was 35.11 ($SD = 11.14$).

Materials and recruitment. Study 2 was identical to Study 1 with the exception that the scenario was in regard

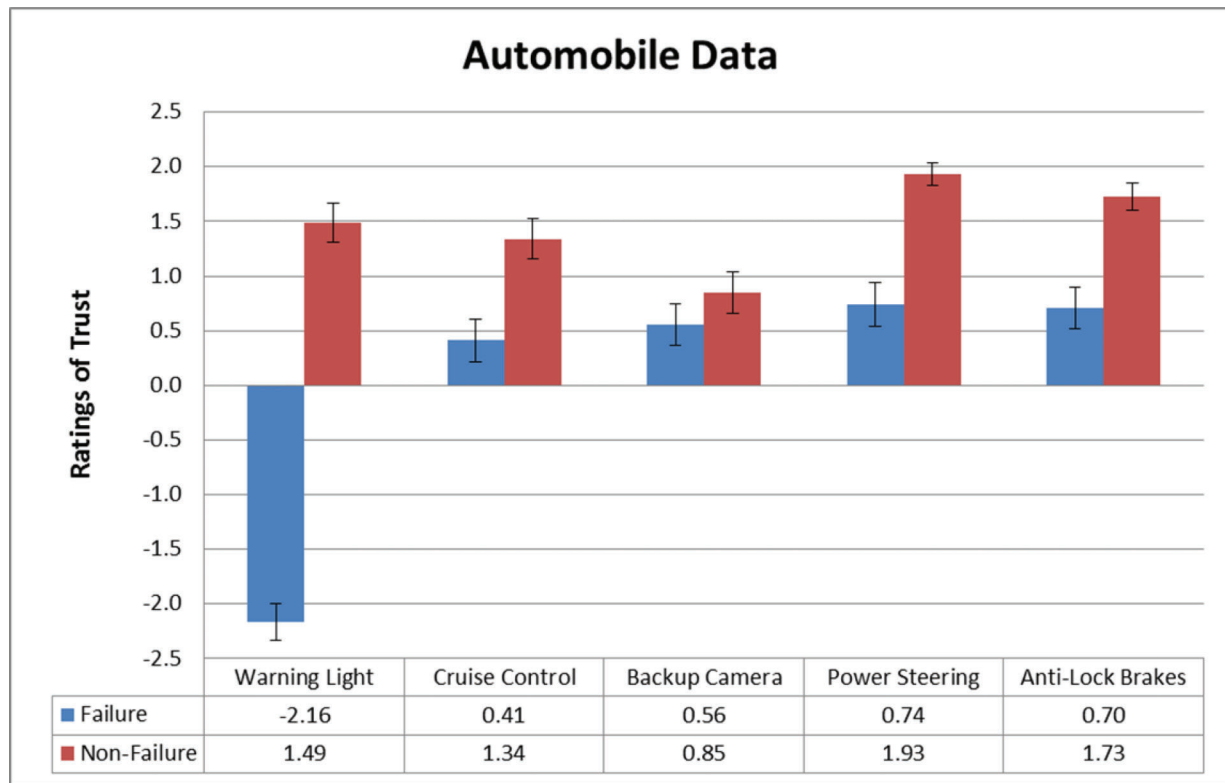


Figure 2. Data from Study 2. SE bars are included.

to a rental car. Participants were presented with the following storyline: “Imagine that you are driving on a 4-hour trip from one major city to another in a rental car. Sometime during the drive, the low-fuel warning light comes on. However, when you stop and check the fuel level, the gas tank is almost full. You realize that the low-fuel warning light is not working properly.” In a separate baseline condition, participants were given the same scenario without the failed low-fuel warning light situation.

Participants were then asked to rate their trust in the automation that operates the low-fuel warning light, the cruise control system, the backup camera, the power steering system, and the antilock brakes. Following this, participants were asked to provide the same demographic information as provided in Study 1 except that two driving-specific questions were added: how many times and how many miles a year the participant drives (on average).

Study 2 Results

Compliance with statistical assumptions. The data were analyzed for compliance with applicable regression and ANOVA assumptions, which included correct specification of the form of the relationship between the independent variables and the dependent variable, correct specification of the independent variables in the regression model, reliability, constant variance of the residuals, independence of the residuals, and normality of the residuals. The data was found to have met all statistical assumption requirements.

Figure 2 presents these data. A two-way ANOVA using Condition as a between-participants factor and DeviceType as a within-participants factor revealed a main effect of Condition, $F(1, 118) = 56.95, p < .001, \eta^2 = .33$; a main effect of DeviceType, $F(4, 472) = 53.58, p < .001, \eta^2 = .31$; and a significant interaction between the two factors, $F(4, 472) = 51.24, p < .001, \eta^2 = .30$. These data reveal that there was a general SWT effect that replicates findings from Study 1 and previous literature.

The resulting regression model for the unreliable device criterion variable included one of the original predictors. *Feelings of negativity when automated devices fail* significantly predicted loss of trust scores, $\beta = .381, t(59) = 3.16, p = .002$, in the device that failed. This model accounted for 15% of the variance in the criterion, $F(1, 59) = 14.55, p = .002$. These results indicate that the stronger the feelings of negativity when automated devices fail, the more there is a loss of trust that occurs when an automated device fails.

A second regression analysis was conducted using the average difference scores for the remaining devices (i.e., the devices that did not fail) with the same predictors. The resulting model again included one of the original predictors. *Feelings of negativity when automated devices fail* significantly predicted loss of trust scores, $\beta = .428, t(59) = 3.64, p = .001$, in the devices that did not fail. This model accounted for 18% of the variance in the criterion, $F(1, 59) = 19.80, p = .001$. These results indicate that the stronger the feelings of negativity when automated devices fail, the more there is a loss of trust during SWT loss.

Study 2 Discussion

The data from Study 2 replicated that of Study 1. First, the SWT effect again remained strong. Second, the data revealed one significant predictor: participants who tended to feel more negatively about automation that fails also tended to lose more trust in the devices described in the scenario.

General Discussion

In the current study, we examined the SWT effect in two settings (commercial flight and rental car) to examine whether this effect remains strong even in hypothetical situations and what might predict loss of trust in similar automated devices simply due to a failure in one presumably unrelated device. Clearly, there are two important findings stemming from the current set of studies that shed further light on the research conducted in SWT. First, SWT is a robust phenomenon, and not just when tested in simulated behavioral settings (Keller & Rice, 2010; Rice & Geels, 2010; Geels-Blair, Rice, & Schwark, 2013). The data reveal that a loss of trust in the unreliable device spread to other devices despite participants not having any information about their reliability. Moreover, SWT appears to show up in a variety of environments. In this case, SWT manifests itself in both airplane and automobile modes. There is no evidence at this time to indicate that similar results would not be anticipated using other automated contexts, regardless of platform; further research will need to explore this possibility.

Second and perhaps related to the first point, SWT appears in participants regardless of gender, ethnicity, age, income, etc. The only significant predictor insofar as the demographic factors included in this study was *feelings of negativity when automated devices fail*. This latter finding is not altogether surprising; however, what is surprising is that *none* of the other demographic factors explained why participants tend to succumb to SWT effects. In this case, what we didn't find is quite revealing. This indicates that SWT is a widespread phenomenon and undoubtedly a powerful influence on how participants perceive their automated systems, regardless of a number of personal characteristics.

Few findings in behavioral research are as compelling as what we found here in SWT, particularly given our large sample size. Of course, other demographic factors or personal characteristics may indeed predict a participant's response to automation unreliability. While this is exciting data in a nascent area (this theory was just developed a few years ago), we recognize that this is still a starting point, and future research will be conducted by the current authors to examine additional demographic or personal variables and their relationship to SWT.

Given that the current findings were derived from survey data only (subjective opinion), an important question that

needs to be investigated is whether these findings will remain intact and manifest themselves in participant behavior with actual operators of the system under investigation. Thus, in the first example used in this study (airplanes), would we see the same results, either in subjective opinion or more interestingly in actual behavior, if our participants were trained operators of aircraft (i.e., pilots). It is always possible that there is a dissociation between what people say and what they would actually do in a specific circumstance; however, it should be noted that the original SWT effect was developed using behavioral data, so we do not expect any major deviations from what we have found here.

Future research needs to test this assumption and determine if the SWT predictor found here also appears in specific operator behaviors, such as decision making in the cockpit. For example, would pilots who tend to react negatively when automated devices fail be less likely to trust unrelated automation if one system is unreliable and, as a result, be less likely to use reliable and independent automated systems, such as in an emergency situation? If this were the case, then we would have the potential for a serious safety of flight issue. Pilots would not use other automated systems when the use of such systems could mitigate an evolving emergency situation. How best to deal with such an eventuality is a critical research issue. Most likely, we would have to either redesign displays to make it apparent that they refer to systems that are in reality independent and/or develop modifications to existing training programs to make pilots aware of this possibility.

More to the point, what do the current study's results mean for operational personnel today? At the very least, we need to inform operators that the phenomenon exists in the first place. Once so informed, operators can take into account the state of affairs at any given moment in time and determine whether there are any significant links between automated systems.

The question here is not so much how do we reduce or eliminate feelings of negativity when an automated device fails. This is probably a realistic assessment and one that is not easily trained out of operators. The real issue is the ability to assess when there is system dependency or whether automated systems function independently. The best method for instilling this is with proper training and knowledge of how these various automated systems function together. We don't want pilots, for example, blindly accepting what an automated system is doing. Rather, we want pilots to have sufficient knowledge of how the automation functions so they can make informed decisions and make assessment as to the reliability of the automation.

Limitations

There are several limitations of this study that we wish to point out. First, our data was collected using a hypothetical

situation rather than a real-world scenario. However, for obvious ethical and logistical reasons, it would be virtually impossible to collect real-world data for these types of situations. Second, our data is limited to online participants from the United States, specifically those participants who share the MTurk environment. While several research studies have shown that MTurk data is as valid and reliable as university participant pools (e.g., Buhrmester, Kwang, & Gosling, 2011; Germine et al., in press), we still acknowledge the limitations of the sampling procedure and the issues it generates with external validity. It would be prudent to caution generalizability due to the nature of the data collected. Third, a regression analysis does not necessarily show causal relationships. We acknowledge that future research should be conducted in order to test the relationships noted here in a more controlled experimental environment. Fourth, it is well known that automation errors have differential effects on operator trust (e.g., Meyer, 2001, 2004; Parasuraman & Riley, 1997; Rice, 2009). In the current study, we only presented scenarios where the automated devices false-alarmed. Further research should investigate situations where the automated devices miss an event.

Practical Implications

We wish to note at least two practical implications of this study. First, it is critical for designers of systems to understand why people succumb to the ubiquitous SWT effect (Keller & Rice, 2010; Rice & Geels, 2010; Geels-Blair, Rice, & Schwark, 2013; Winter, Rice, & Reid, 2014). While these data are preliminary and can only represent a starting point for understanding the predictors of SWT, they do point to an interesting possibility that much of the SWT effect might be due to emotional factors rather than cognitive ones. When designing a new system with complex multiple displays and automated devices, it is important to be aware of how a failure in one device might trigger an SWT effect and subsequent loss of trust in other devices within the system.

Second, as mentioned above, it is important for operators and those who train them to understand what is happening. It is said that opaque automation systems, meaning systems that are not clear in terms of usability to the user, degrade trust (Rice, 2009); it might also be argued that opaque explanations for phenomena such as the SWT effect might also have negative effects on operator trust and training methodologies. If operators are trained to become aware of the SWT effect and how easy it is to succumb to it, they might be more resistant to its effects and act in a more rational manner. This conclusion, however, is only marginally supported, as the samples did not include any pilots. Their attitudes are divergent than those from the general population, and SWT is an essential necessity of their successful work.

Conclusions

The purpose of this study was to determine if there are any significant predictors of the SWT effect. Via a regression analysis across two studies, the data clearly show that there is at least one significant predictor and that this predictor has a very strong effect on SWT. We conclude that *feelings of negativity when automated devices fail* appears to drive the SWT effect to a large degree, accounting for up to 21% of the variance in the data. This study has strong practical implications for designers of complex automated systems.

References

- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science*, 6(3), 3–5.
- 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, 474–486.
- Dixon, S. R., Wickens, C. D., & Chang, D. (2005). Mission control of multiple unmanned aerial vehicles: A workload analysis. *Human Factors*, 47, 479–487.
- Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse than misses? *Human Factors*, 49, 564–572.
- Deutsch, M. (1958). Trust and suspicion. *The Journal of Conflict Resolution*, 2, 265–279.
- Eckel, C. C., & Wilson, R. K. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization*, 55, 447–465.
- Ergeneli, A., Saglam, G., & Metin, S. (2007). Psychological empowerment and its relationship to trust in immediate managers. *Journal of Business Research*, 60, 41–49.
- Geels-Blair, K., Rice, S., & Schwark, J. (2013). Using system-wide trust theory to reveal the contagion effects of automation false alarms and misses on compliance and reliance in a simulated aviation task. *International Journal of Aviation Psychology*, 23(3), 245–266, DOI: 10.1080/10508414.2013.799355
- Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B., . (in press). Is the web as good as the lab? Comparable performance from web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*.
- Johnson, C. W., Shea, C., & Holloway, C. M. (2008). The role of trust and interaction in GPS related accidents: A human factors safety assessment of the Global Positioning System (GPS). Proceedings of the 26th International Conference on Systems Safety, Vancouver, Canada, 2008.
- Keller, D., & Rice, S. (2010). System-wide versus component-specific trust using multiple aids. *Journal of General Psychology*, 137(1), 114–128.
- Lee, J. D., & Moray, N. (1994). Trust, self-confidence and operators' adaptation to automation. *International Journal of Human-Computer Studies*, 40, 153–184.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46, 50–80.
- Maltz, M., & Shinar, D. (2003). New alternative methods in analyzing human behavior in cued target acquisition. *Human Factors*, 45(2), 281–295.
- Mehta, R., Rice, S., Winter, S. R. (2014). Examining the relationship between familiarity and reliability of automation in the cockpit. *Collegiate Aviation Review*, 32(2), 1–13.

- Meyer, J. (2001). Effects of warning validity and proximity on responses to warnings. *Human Factors*, 43, 563–572.
- Meyer, J. (2004). Conceptual issues in the study of dynamic hazard warnings. *Human Factors*, 46(2), 196–204.
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, abuse. *Human Factors*, 39, 230–297.
- Parasuraman, R., Sheridan, T., & Wickens, D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30, 286–297.
- Reeves, B., & Nass, C. (1996). The media equation: How people treat computers, television, and new media like real people and places. *International Journal of Instructional Media*, 33, 19–36.
- Rice, S. (2009). Examining single and multiple-process theories of trust in automation. *Journal of General Psychology*, 136(3), 303–319.
- Rice, S., & Geels, K. (2010). Using system-wide trust theory to make predictions about dependence on four diagnostic aids. *Journal of General Psychology*, 137(4), 362–375.
- Rice, S., Trafimow, D., Clayton, K., & Hunt, G. (2008). Impact of the contrast effect on trust ratings and behavior with automated systems. *Cognitive Technology Journal*, 13(2), 30–41.
- Wickens, C. D., & Colcombe, A. (2007). Dual-task performance consequences of imperfect alerting associated with a cockpit display of traffic information. *Human Factors*, 49, 839–850.
- Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.), Upper Saddle River, NJ: Pearson Prentice Hall.
- Wiegmann, D. A., Rich, A., & Zhang, H. (2001). Automated diagnostic aids: The effects of aid reliability on users' trust and reliance. *Theoretical Issues in Ergonomics Science*, 2(4), 352–367.
- Winter, S. R., Rice, S., & Reid, K. (2014). Using system-wide trust theory to analyze passenger loss of trust in aircraft automation. 5th International Conference on Applied Human Factors and Ergonomics.
-

Dr. Stephen Rice is an associate professor at Embry-Riddle Aeronautical University, Daytona Beach. He received his PhD in human factors from the University of Illinois at Urbana-Champaign in 2006. His research interests include aviation psychology, automation, trust, stigmas, and human performance.

Dr. Scott R. Winter is an assistant professor at the Florida Institute of Technology. He completed his PhD in aviation technology from Purdue University in 2013. His research interests have centered on decision making, pilot training, and aviation human factors.

Dr. John Deaton is a professor at the Florida Institute of Technology. He completed his PhD in applied experimental psychology from the Catholic University of America in 1988. His research interests include cross-cultural training, knowledge management, and content development.

Dr. Ismael Cremer is an assistant professor at the Florida Institute of Technology. He completed his MS in environmental science and engineering from the King Abdullah University of Science and Technology in 2010 and completed his PhD at the Florida Institute of Technology in 2015. His research interests include environmental science and aviation sustainability.