Engineering Trust in Complex Automated Systems

Transparency of tools used in emergency operations can increase trust among commercial pilots interacting with an automated tool for emergency operations.

By Joseph B. Lyons, Kolina S. Koltai, Nhut T. Ho, Walter B. Johnson, David E. Smith, & R. Jay Shively

FEATURE AT A GLANCE:
We studied the transparency of automated tools used during emergency operations in commercial aviation. Transparency (operationalized as increasing levels of explanation associated with an automated tool recommendation) was manipulated to evaluate how transparent interfaces influence pilot trust of an emergency landing planning aid. We conducted a low-fidelity study in which commercial pilots interacted with simulated recommendations from NASA's Emergency Landing Planner (ELP) that varied in their associated levels of transparency. Results indicated that trust in the ELP was influenced by the level of transparency within the human–machine interface of the ELP. Design recommendations for automated systems are discussed.

KEYWORDS:
trust in automation, transparency, commercial aviation, human–machine interface, commercial airline pilot, automated tools, emergency operations, pilot error, shared awareness, NASA, automated aids

Trust is believed to play a critical role in shaping how humans interact with technology (Chen & Barnes, 2014; Lee & See, 2004; Lyons & Stokes, 2012). It is relevant to the use of technology because it shapes how humans rely on and accept novel technology. However, as humans, we tend toward suboptimal reliance strategies involving technology (Geels-Blair, Rice, & Schwark, 2013; Lee & See, 2004; Lyons & Stokes, 2012), resulting in potential errors when we choose to use an error-prone tool, fail to use a potentially beneficial technology, or attribute unreliable facets of one system component to another part of the system.

Researchers have recently relied on the construct of transparency to address, in part, the challenge of trust calibration. Broadly, transparency represents a process for establishing shared awareness and shared intent between a human and a machine (see Lyons, 2013). Real-life incidents, such as the Asiana 214 San Francisco crash and the US Airways 1549 Hudson River landing, highlight the importance of transparency of automated systems in calibrating trust (National Transportation Safety Board, 2009, 2013).

Transparency may support more optimal trust calibration by providing humans with insight into the analytical, intentional, and awareness-based parameters of an intelligent system, such as a robot (Lyons, 2013). In fact, prior studies have consistently demonstrated that added transparency supports trust development (Chen, Barnes, & Harper-Sciarini, 2011; Wang, Jamieson, & Hollands, 2009).

In this brief report, we share the results of a low-fidelity study of transparency of automated tools for commercial aviation.

EMERGENCY LANDING TECHNOLOGIES

The National Aeronautics and Space Administration (NASA) has developed tools to support emergency procedures in commercial aviation. Automated aids in aircraft can be useful, given that aviation is a highly complex and cognitively demanding task (Wickens, 2007); yet care is required to avoid adding further complexity to an already daunting task domain. One such automated tool is the Emergency Landing Planner (ELP; Meuleau, Plaunt, Smith, & Smith, 2008). The ELP was designed to support rapid analysis of complex situations, including damage to the aircraft, adverse weather, and status of possible landing sites to recommend a safe route and desired approach.

Recommendations from the ELP are important because the automated tool conducts extensive analysis of many elements that would be too challenging for a pilot to consider quickly during an emergency situation. For instance, when an aircraft sustains damage or experiences equipment failure, it might change the flight dynamics of the aircraft, requiring higher airspeed, gentler turns, or climb/descent limitations. Further, weather patterns can be complex, forcing pilots to rapidly choose new flight paths and landing possibilities. However, as with any automated tool, commercial pilots may not rely optimally on the ELP (Meuleau, Neukom, Plaunt, Smith, & Smith, 2011), and its use may be associated with other unintended consequences (see Parasuraman & Riley, 1997). Thus the ELP was a good domain to investigate the effects of added transparency on user trust for commercial airline pilots.

STUDY METHOD

The current study used a repeated-measures design to present the ELP interface and output in a series of vignette-based scenarios to examine the role of transparency on trust among commercial pilots ($N = 12$; despite the small sample size, the repeated-measures design allowed for sufficient statistical power for our analyses). Our consultation with two subject-matter experts and the engineers of the ELP revealed that an important factor that influences pilot trust in the ELP is the pilot's understanding of the rationale behind the ELP's recommendations. This fact motivated us to develop the following three transparency conditions (see Figure 1): control (the current baseline for the ELP output), risk-based transparency, and logic-based transparency. The control condition provided the same information that the ELP uses to make its recommendation (e.g., weather; runway characteristics, such as length; and terrain). Because this condition was the extant status of the ELP, it was considered to be an adequate baseline condition.

The second condition, risk-based transparency (referred to as the *value* condition), offered all the information in the control condition, but in addition, pilots were given a risk statement about the probability of success for that landing maneuver (e.g., "There is a 34% chance that you will be able to successfully complete the approach and landing under current conditions"). Note, however, that in this study, a 34% success rate does not translate into a 66% chance of crashing; it means only that the probability of landing successfully on the first attempt is 34%. This important nuance was explained to the pilots during their training with the ELP and its interface.

Finally, the logic-based transparency condition (referred to as the *logic* condition) included everything within the control and value conditions, but additionally, a statement was included to provide the logic behind the risk statement (e.g., runway unacceptable – the landing crosswind is too high for a safe landing). The idea was that the pilot would receive increased levels of transparency into the ELP output as he or she progressed from the control through the logic conditions. The order of the conditions was counterbalanced, and order had no impact on trust.

We developed a series of six scenarios with emergencies ranging from weather challenges at the planned airport or on the current trajectory to aircraft maintenance issues. The scenarios are summarized in Figure 2. Each scenario involved a situation in which the pilot had to divert from the current flight path and select a new airport to land the plane. The pilots were not given a time limit but were told to make the best decision as quickly as possible.

For each scenario, pilots were presented with a list of diversion options. The list order was randomized, and the pilots were required to consider the evaluation of each option the ELP provided. Note that outside of the current experiment, the ELP's top selection is currently the first one listed, but the pilots in this study had no prior experience with the ELP.

Pilots were initially trained on the ELP through a brief presentation that explained its purpose and design, how to use it, and how it integrates information to make a recommendation. All the pilots were aware that the ELP was a NASA technology. Pilots were then shown a series of six vignettes using all six scenarios described earlier, within three trials. Each trial contained two vignettes with the same transparency condition. Trust was assessed following each trial. Pilot trust was measured through a seven-item scale that gauged reliance intentions on the ELP.

Example items included "I would feel comfortable relying on the recommendations of the ELP in the future"; "When the task was hard, I felt like I could depend on the ELP"; and "I would be comfortable allowing the system to make a diversion decision for me." The seven items were rated on a Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The pilots were also asked to rate how helpful several aspects of the ELP were. These factors included the Automatic Terminal Information Service (i.e., real-time weather information), approach information, the logic and value statements provided in the respective logic and value conditions, airport ratings, the list of diversions, individual diversion or runway information, airport facilities descriptions (e.g., emergency vehicle information), and the diversion path.

RESULTS

The trust scale showed acceptable reliability (alpha > .85). As shown in Figure 3, trust varied by transparency condition, $F(2, 22) = 4.92$, $p < .05$. Specifically, trust was highest in the logic condition and lowest in the control condition. Contrasts revealed that trust in the control condition was significantly lower than trust in the logic condition, $p < .05$. Pilots rated aspects of the transparency-based interface as being helpful during the scenarios.

As shown in Figure 4, pilots rated the transparency aspects of the value statement and the logic as very useful. Further, qualitative reports from the pilots indicated a strong preference for the logic condition. The pilots reported that this condition was useful for the following reasons:

- "[It] gave me the most useful information at a glance."
- "I realized I missed the reasoning when I didn't have it."
- "It helped me catch something that I would have otherwise missed."

ENGINEERING TRANSPARENCY INTO COMPLEX AUTOMATION

The study demonstrated that added transparency can facilitate higher trust and utility among operators of complex systems. Establishing optimal reliance strategies is critical to human–machine interactions (Lee & See, 2004; Lyons & Stokes, 2012); thus researchers and designers should continue to seek out and understand relevant factors that
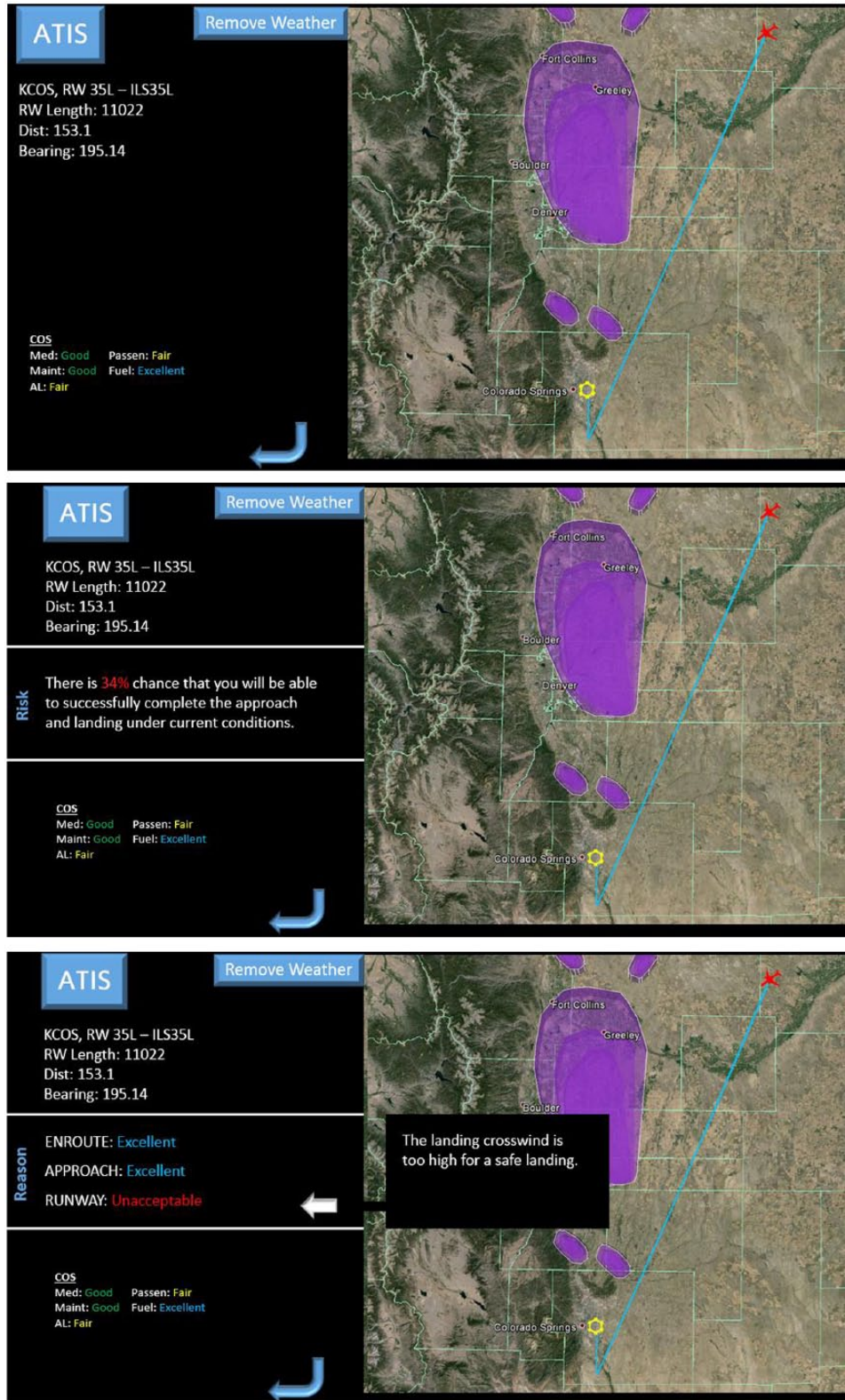
Figure 1. Examples of a diversion recommendation in the control (top), value (middle), and logic (bottom) conditions.

Figure 2. Emergency scenarios for the Emergency Landing Planner experiment. ORD = Chicago O'Hare; DEN = Denver; PHX = Phoenix; COS = Colorado Springs; SFO = San Francisco; EGE = Eagle County; SEA = Seattle; DFW = Dallas/Fort Worth; NTD = need to divert.

influence the trust development process between humans and machines.

Although transparency, as defined by Lyons (2013), is broader than just analytic components, our results suggest that these are often critical for systems that offer recommendations for action. In particular, it suggests that these recommendations should be accompanied by the logic from which they were derived, or the pilots, who must balance the dynamic demands of aviating, navigating, communicating, and managing systems, may ignore them (Wickens, 2007). We believe that the importance of transparency will increase as technology – notably, technology with autonomous capabilities – proliferates throughout industry and government. As such proliferation occurs, it will be crucial for operators to adequately understand the logic behind autonomous systems.
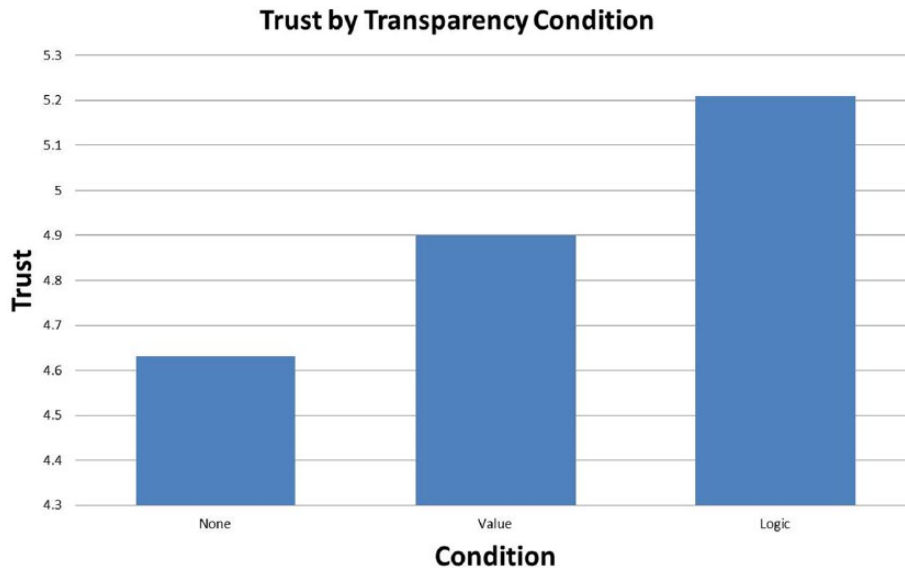

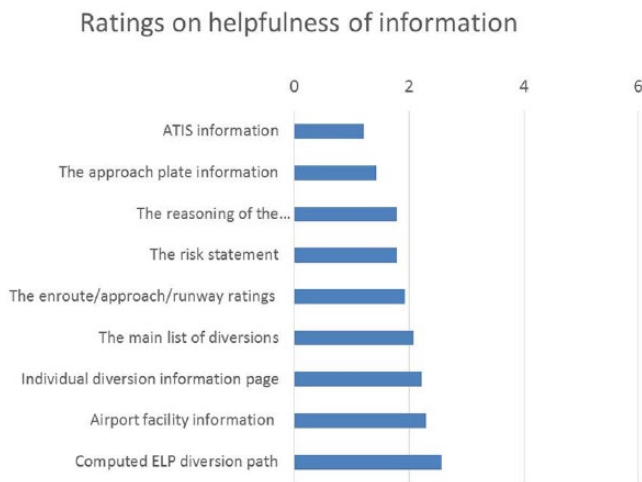Figure 3. Trust by transparency condition.

Figure 4. Helpful ratings of the Emergency Landing Planner information (1 = *extremely helpful*, 6 = *not at all helpful*).

REFERENCES

Chen, J. Y. C., & Barnes, M. J. (2014). Human–agent teaming for multirobot control: A review of the human factors issues. *IEEE Transactions on Human–Machine Systems*, *44*, 13–29.

Chen, J. Y. C., Barnes, M. J., & Harper-Sciarini, M. (2011). Supervisory control of multiple robots: Human-performance issues and user-interface design. *IEEE Transactions on System, Man, and Cybernetics, Part C: Applications and Reviews*, *41*, 435–454.

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*, 245–266.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*, 50–80.

Lyons, J. B. (2013). Being transparent about transparency: A model for human–robot interaction. In D. Sofge, G. J. Kruijff, & W. F. Lawless (Eds.), *Trust and autonomous systems: Papers from the AAAI Spring Symposium* (Tech. Rep. SS-13-07). Menlo Park, CA: AAAI Press.

Lyons, J. B., & Stokes, C. K. (2012). Human–human reliance in the context of automation. *Human Factors*, *54*, 111–120.

Meuleau, N., Neukom, C., Plaunt, C., Smith, D., & Smith, T. (2011). The emergency landing planner experiment. In *ICAPS-11: Proceedings of the Scheduling and Planning Applications Workshop* (pp. 60–67). Freiburg, Germany: AAAI.

Meuleau, N., Plaunt, C., Smith, D., & Smith, C. (2008). Emergency landing planning for damaged aircraft. In *ICAPS-08: Proceedings of the Scheduling and Planning Applications Workshop*. Sydney, Australia: AAAI.

National Transportation Safety Board. (2009). *Loss of trust in both engines after encountering a flock of birds and subsequent ditching on the Hudson River US Airways Flight 1549* (AAR-10/03 PB2010-910403). Washington, DC: Government Printing Office.

National Transportation Safety Board. (2013). *Descent below visual glidepath and impact with seawall Asiana Airlines Flight 214* (AAR-14/01 PB2014-105984). Washington, DC: Government Printing Office.

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

Wang, L., Jamieson, G. A., & Hollands, J. G. (2009). Trust and reliance on an automated combat identification system. *Human Factors*, *51*, 281–291.

Wickens, C. (2007). Aviation. In F. T. Durso, R. S. Nickerson, S. T. Dumais, S. Lewandosky, & T. J. Perfect (Eds.), *Handbook of applied cognition* (pp. 361–389). Hoboken, NJ: Wiley.

Joseph B. Lyons is technical adviser for the Human Trust and Interaction Branch within the 711 Human Performance Wing at Wright-Patterson AFB, Ohio. He received his PhD in industrial/organizational psychology from Wright State University in Dayton, Ohio, in 2005. His interests include human–machine trust, interpersonal trust, leadership, and social influence. From 2011 to 2013, he served as a program officer for the Air Force Office of Scientific Research.
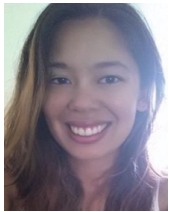
Kolina S. Koltai is a research associate at NASA Ames in the Flight Deck Display Research Laboratory under the tutelage of Nhut Ho and Walter Johnson. Her focus is applying social-psychological principles to human–computer interaction issues. She is studying the relationship between trust and automation transparency. Her previous research examined how culture, organizations, and automation capability affect trust and reliance in automation.

Nhut T. Ho is a professor of mechanical engineering at California State University, Northridge, where he is director of the Systems Engineering Research Lab. His research interests and projects include human–automation trust, modeling and creation of principles for designing transparency in human–automation teamwork systems, and case and field studies to determine how fighter pilots interact with and accept automatic collision avoidance systems. He obtained his MS and PhD from MIT.

Walter B. Johnson is a research psychologist with NASA's Human Systems Integration Division at the Ames Research Center and leads the Flight Deck Display Research Laboratory there. While at NASA, his research has focused on the development of new operational concepts for the U.S. air traffic management system and on the design of airside and groundside interactive displays and decision support tools that support those concepts.

David E. Smith is a senior researcher in the Intelligent Systems Division at NASA Ames Research Center, where he conducts research on automated planning and scheduling. He received his PhD in 1985 from Stanford University. He served as an associate editor and guest editor for the Journal of Artificial Intelligence Research *and is a Fellow of the* Association for the Advancement of Artificial Intelligence.

R. Jay Shively is the human–systems integration lead for NASA's UAS Integration into the NAS project. In this role, he leads a multidisciplinary team across four NASA research centers to address barriers impeding the integration of unmanned aircraft systems (UAS). To serve in this role, he transitioned to NASA after 25 years with the U.S. Army. During his tenure with the Army, he was the Human Systems Integration group leader for the Aeroflightdynamics Directorate. There he coordinated efforts on helicopter brown-out symbology, UAS ground station design, and cockpit design issues.

eid