# Influence of Cultural, Organizational, and Automation Capability on Human Automation Trust: A Case Study of Auto-GCAS Experimental Test Pilots

**Kolina Koltai, Dr. Nhut Ho, Dr. Gina Masequesmay, David Niedober**
California State University, Northridge
18111 Nordhoff St. JD1624C
Northridge, CA 91330
nhut.ho.51@csun.edu
**Mark Skoog**
NASA Dryden Flight Research Center
Mark.A.Skoog@nasa.gov

**Artemio Cacanindin**
Air Force Flight Test Center
Artemio.Cacanindin@edwards.af.mil
**Dr. Walter Johnson**
NASA Ames Research Center
Walter.Johnson@nasa.gov
**Dr. Joseph Lyons**
Air Force Research Lab
Joseph.Lyons.6@us.af.mil

## ABSTRACT
This paper discusses a case study that examined the influence of cultural, organizational and automation capability upon human trust in, and reliance on, automation. In particular, this paper focuses on the design and application of an extended case study methodology, and on the foundational lessons revealed by it. Experimental test pilots involved in the research and development of the US Air Force's newly developed Automatic Ground Collision Avoidance System served as the context for this examination. An eclectic, multi-pronged approach was designed to conduct this case study, and proved effective in addressing the challenges associated with the case's politically sensitive and military environment. Key results indicate that the system design was in alignment with pilot culture and organizational mission, indicating the potential for appropriate trust development in operational pilots. These include the low-vulnerability/high risk nature of the pilot profession, automation transparency and suspicion, system reputation, and the setup of and communications among organizations involved in the system development.

## Keywords
Trust, automation suspicion, reliance, F-16, military organizational, pilot culture, extended case study methodology, automatic ground collision avoidance

## INTRODUCTION
With the increase of automation in aviation, inappropriate reliance on automation becomes an increasingly relevant issue [11]. Misuse and disuse of automation characterize inappropriate reliance [6]. Misuse is described as the reliance on an automated system for something other than its intended purpose and disuse is the disengaging of the automation [6, 11]. Research has shown that reliance on automation depends on many factors, including the trust that the operator places in the automation and issues related to the capability and complexity of the automation [9]. In particular, trust has been shown to be a meaningful concept guiding human-automation reliance similar to the way trust mediates the relationships between individuals, between organizations, and between individuals and organizations [3, 5, 9, 12]. The fundamental premise of trust is the willingness of an individual to be vulnerable to the actions of another entity [9], may be it a person or a machine (i.e., automation). Lee and See [6] conducted a comprehensive review of the studies in the area of trust in automation, using the review as the basis for an integrative view linking organizational, sociological, interpersonal, psychological, and neurological perspectives on inter-personal trust to the issue of human-automation trust. In particular, they found that there is a general lack of data on, or research that examines, how cultural and organizational factors influence human-automation trust and reliance. Furthermore, while some studies have shown that these factors can influence human-automation interaction in unexpected ways, they have been mostly confined to experiments that examine the effects of a limited set of independent variables in a well-controlled environment. In actual operations, interactions between humans and automation usually take place in settings where there are many more variables of interest than data points available, and where the investigators do not have control over the events. Thus, to build on existing experimental data, and to lay the foundation for future research, it is essential to capture the richness of the phenomenon and the extensiveness of the real-life context in which human interacts with automation. The work presented in this paper sought to address this in the context of a case study of an actual military automated system.


*HCI-Aero 2014,* Jul 30-August 1, 2014, Silicon Valley, California, USA.

Practical lessons learned and real-world perspectives of the appropriateness of reliance (i.e., where trust and use of automation matches system capabilities) are identified. Moreover, in addition to a standard approach examining how intrinsic properties of automation influences reliance, our emphasis on cultural and organizational factors allowed us to examine how human-automation reliance can be influenced in ways that are less directly related to the characteristics of the automation. Finally, the case study method presented here facilitates the collection of evidence from multiple sources, thus utilizing converging operations to identify the presence of latent variables or constructs (e.g., trust, distrust) in a triangulating fashion.

To identify foundational lessons and best practices about appropriate reliance on automation, our research approach involved conducting a case study of key Department of Defense (DOD) personnel with experience in operating and developing a military automated system. The Automatic Ground Collision Avoidance System (Auto-GCAS) was selected as the context for this case study which, in turn, employs a multi-case design. Cultural, organizational, and automation capability factors are studied through three interrelated cases: 1) Auto-GCAS experimental test pilots; 2) Auto-GCAS engineers; and 3) management who lead and oversee the Auto-GCAS program. This paper reports data on the experimental test pilots (ETPs), and future papers will address engineers and managers.

The remainder of the paper first discusses the methodology and methods utilized, and then presents lessons learned from the methodological research design and implementation processes, and key findings on the influence of automation capability, cultural, and organizational factors on trust development. In the last section, conclusions and implications for using the results of this research to add to the body of knowledge on human automation trust are discussed.

## METHODOLOGY AND METHODS

### Extended Case Study Methodology

Our methodology was inspired by Michael Burawoy, who proposed in his 1991 book, Ethnography Unbound [2], to combine the advantages of two traditions, qualitative and quantitative approaches, of empirical evidence and theories. Burawoy emphasized the contribution of qualitative research in gaining a “Verstehen” understanding (i.e., understanding the meaning of actions from the actor's point of view) of the meanings and values of the people we study. By immersing themselves into the culture and world of their subjects, ethnographers can gain a deeper understanding of the norms and values of that group, which provides a subjective worldview of the research that cannot be gained from detached surveys or interviews. Qualitative methodology that utilizes methods like ethnographic field notes, participant observation, and open-ended questions provide a richer set of data, which makes it possible to introduce a subject’s interpretation into conclusions. This subjective, complex, and in-depth process provides context and completeness to information gathered from the population. On the other hand, Burawoy notes, quantitative methodologies, that include using surveys and closed question interviews, can provide a larger N and a basis for comparison between different populations or individuals. Studies utilizing quantitative methodologies are often quick to complete, and very effective in gathering data to proposed questions, but they often create many more unanswered questions at the end. Using these methodologies together, theories concluded from the collected data can then be tested to contribute to the larger scientific enterprise of knowledge seeking. Burawoy also advocated the meticulous examination of local societal trends and institutional policies, and connection to the macro levels of these trends and policies, in order to understand how everyday life challenges or reinforces the hegemonic order [2]. By doing so, he advocated that researchers should be both empirically and theorically driven, and that they study cases in the context of macro trends. In the spirit of this advocacy, we attempted to examine the development of pilots’ trust on Auto-GCAS in context of the larger organizational and cultural systems of the Air Force, NASA and Lockheed/Martin. Taking from Burawoy, the theories generated from the data collection process are expanded to include macro ideologies; in this case study, for example, the military industrial complex is a macro idea that was used to aid in examining and explaining the organizational factors that influenced trust development of Auto-GCAS. Following Burawoy’s suggestions, the project’s research design expanded upon his theory to capture the complex and intricate world of ETPs within the organizational context of this project.

### Methods

Figure 1 depicts the project’s extended case study methodology, which expands Burawoy’s extended case methodology to include not just participant-observation data, but also interview and survey data, in the analysis and discussion/revision of existing theories. Extended case methodology combines grounded theory and traditional research approaches.

- The grounded theory approach sets the framework for the project’s primary sources, while the traditional research methodology provides the project’s secondary sources. Grounded theory is a qualitative methodological approach that is an inductive hypothesis generating, as opposed to a deductive quantitative hypothesis-testing approach. Hypotheses are generated from observing patterns that emerge out of ethnographic data, rather than extracting hypotheses from existing theories and then collecting data to test these hypotheses. Using grounded theory, the project aimed to answer research questions by using questionnaires, surveys, interviews with participants, field notes, and observations. These qualitative methods provide the ethnographic approach dimension to the study.

- In contrast to grounded theory, traditional research tests theories in a deductive manner. In this project, the research team immersed itself in the current literature on human automation trust development, with an emphasis on automation capability, cultural, and organizational factors. This included doing extensive literature review on the cultures and organizations involved with the development and testing of Auto-GCAS, and using the results of this review to generate hypotheses to be tested after the synthesis of the literature review was completed.

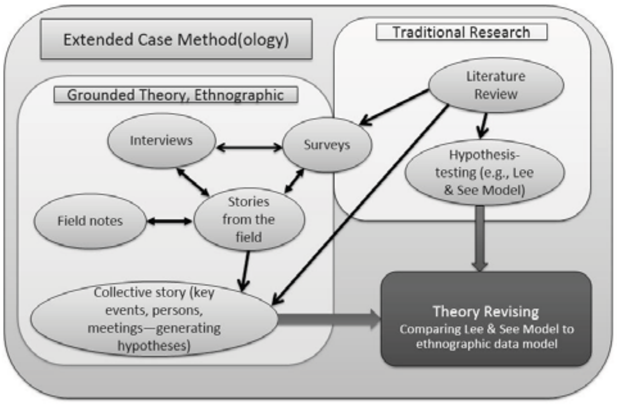


**Figure 1**: Research Design Process and Strategy

Mirroring this dichotomous methodological approach, we divided the research team into two groups. One group did an in-depth literature review on human-automation trust, Auto-GCAS and its predecessors, and the cultures and organizations involved with its development. Another group worked on open-ended interviewing and survey questions. NVivo, an ethnographic research software package, was used to code the data collected from interviews, questionnaires, field notes, observations, and the literature review. The team then used this to help identify trends and to establish a collective story out of the key events, people, and meetings that contribute to the development of trust in Auto-GCAS, and generate theories and hypotheses. Using both primary and secondary data, comments on existing theories about trust in automation can then be posed. Once theories were generated from both the traditional literature review, and from the grounded theory methods, these theories and hypothesis were compared to see if they converged or diverged from each other.

### Unique Challenges and Solutions

To capture the influence of cultural and organizational factors, the research design process had to address a number of unique challenges. The first was the research team's lack of familiarity with ETP culture. ETPs reside within a small, tight-knit community in which cultural studies are not often conducted. Since the literature on ETP is limited, to effectively study how ETP culture influences trust development in automation, we had to first study and define ETP culture. However, this then led to another challenge, which was gaining access to selective personnel and recruiting ETPs to participate in the research, in coordination with the government contacts at the Flight Test Center and NASA. Since this study was done within the context of the ETPs Air Force duties, traditional incentives (e.g., monetary) for experimental participants could not be offered. In addition, by the nature of their profession, ETPs are small in number and have incredibly busy work schedules. The environment in which Auto-GCAS has been developed and tested also demanded respect from the research team in terms of understanding the history of the development of Auto-GCAS, and the political and sensitive nature of the respective military and government organizations involved. The research team also needed to quickly gain an adequate understanding of a highly technical topic (Auto-GCAS), along with a complex organizational structure (with respect to US Air Force and NASA technology development process) in order to effectively process literature and findings and to converse with personnel and participants. Geography also was a challenge because the research team was based in Los Angeles, California, while ETP locations ranged from Edwards, California to Arlington, Virginia. Part of the methods used in the study, as explained in the next section, involved interviews, which were ideally conducted face-to-face and required traveling to several air force bases. As stated by Lee & See [6], there is a lack of extensive literature regarding cultural and organizational factors in human-automation trust development. This provided less of an immediate literature foundation for the research team. In addition to these issues, the project also attempted to capture a large quantity of information within a short time frame of 18 months.

To address these challenges, three main solutions were developed and implemented: 1) an agile implementation of methods, which allowed the research team to adapt to the opportunities emerged during our investigation and to the uniqueness of the population groups; 2) the utilization of key people who could assist in gaining entrée to participants and research sites, as well as providing information and insights; and 3) the research team developed cultural and technical competency in order to effectively interact with the ETP population and work within the constraints of the complex organizational structures.

## RESULTS

### Methodological Lessons Learned

The extended case study research design and implementation described were effective for obtaining the necessary data and for analyzing the data to reveal foundational lessons about ETP trust development in Auto-GCAS. These were based upon the three main solutions implemented in this study.

- Specifically, the agile implementation of the methods proved to be highly useful and beneficial. Initially, we planned for a 2-hour interview with the participants; however, after discovering the time constraints of participants, the method was restructured to include an online questionnaire that could be completed at each participants' convenience with a follow-up 1-hour interview. Interviews were restructured to include open-ended questions and follow-up questions based upon how the respondent completed the questionnaire. Field observations which were initially intended to just include observing mission testing were extended as permitted by the circumstances and access gained by the research team. As a result, the research team was able to immerse itself deeper into ETP culture by conducting numerous observations not only at the work place (i.e., during mission testing and at AFBs), but by formal congregations (e.g., ETP conferences), and informal gatherings (e.g., meals and post-work bonding social events).

- The project also owes a large part of its success to the assistance from key personnel with entrée into the ETP community and various organizations. These key personnel were able to provide valuable insights and feedback during the research process. They facilitated in obtaining access to several AFBs including Edwards, Nellis, Shaw, and Hill, as well as organizing visitation rights to observe mission tests (from pre-mission briefing to post-mission de-briefing). They also assisted in recruiting 14 male and 1 female Caucasian, US born ETP participants, with ages ranging from 28 to 65 years old. All had flown Auto-GCAS and had a median of 2.5 years of experience working with Auto-GCAS. The association with these participants also helped the research team establish credibility and trustworthiness in the communities of the participants.
- Building and having a deep understanding and cultural competency with respect to the ETPs and their community also proved to be highly critical for, and beneficial to, the conduct of the study. The substantial efforts made by the research team to learn ETP culture not only helped the team understand the values and norms of the culture, but also helped them converse and interact with ETPs using their technical language and vernaculars. This, in turn, helped the team to gain respect, build rapport, and gain the trust of the ETPs. By strengthening that connection between the research team and the ETPs, the research team was able to gather necessary data.

**Influence of Automation Capability, Cultural, and Organizational Factors on Trust Development**

A number of key lessons and best practices regarding how automation capability, as well as cultural and organizational factors, influence ETPs trust in Auto-GCAS were identified and are summarized in this section. These findings are pulled from the analysis of interviews, field notes, participant observations, surveys and literature review.

- For safety systems such as Auto-GCAS, determining the right threshold to take control away from the operator is critical for developing a system that does not impede or interfere with the operator. Auto-GCAS was designed to automatically execute a roll to wings-level, 5-G recovery maneuver when the time available to avoid ground impact (TAAGI) is 1.5 seconds. This 1.5-second budget, established through flight tests in which ETPs flew towards terrain and rated the TAAGI in relation to their anxiety, was a design requirement that helps eliminate nuisance activations. Nuisance activations that take away the control of the aircraft is aversive to both ETPs and operational F-16 fighter pilots, and the elimination of these false alarms is highly valued. Having a nuisance free system builds appropriate trust on the system. The 1.5 second budget is specifically designed to address the needs of the operator: F-16 pilots. Pilots stated in interviews that Auto-GCAS does not activate sooner than they would execute an evasive maneuver which increases their trust that Auto-GCAS will not interfere with their flight mission. This identified lesson is extremely valuable for developers of future safety systems with similar characteristics to Auto-GCAS.
- It was found that national differences may have an influence on Auto-GCAS reliance/compliance. Ideally, participants were to include international pilots with Auto-GCAS experience, but those pilots were not able to be recruited for this study. Based on the responses drawn from the US born pilots, there was evidence indicate that national differences between air force pilots can influence reliance in automated technologies. A few pilots with exposure to international cultures cited that pilots from non-Western or Soviet nations who fly according to commands from centralized ground control stations may be more likely to comply with Auto-GCAS due to cultures that prefer conformity. It was also stated that Eastern Asian cultures with heavy conformity and adherence to hierarchy may also heavily rely on such autonomous technologies. This phenomenon is interesting and warrants further exploration.
- In a survey rating their opinions of the Auto-GCAS engineers and managers/leaders, the majority of the ETPs rated "strongly agreed" that they trusted the Auto-GCAS engineering and management teams and that the engineers and managers were competent and they could have honest conversations with them. Pilots stated they had an increased trust in Auto-GCAS before flying just by the positive opinions of the competency of the teams that worked on developing it. This opinion is consistent with the mutual respect that the research team observed among the ETPs, engineers, and managers in their interviews and personal relationships. Thus, the nature of relationships among members of different occupations/organizations can have positive or negative influence on trust development. In the Auto-GCAS case, having respect and a positive opinion of other people's work, profession, and organization, facilitated a positive trust in the automation they produce.

- Generational differences affect the acceptance/use of automation. Participants often stated in their interviews that there were differences in terms of acceptance of automation between the older "fly-by-the-seat-of-your-pants" generation and the young "Nintendo" generation. The younger generation of pilots grew up with more exposure and familiarity with automation which makes them more likely to accept new automated technologies. It was also observed during observations at SETP symposia that older pilots emphasized a need for relying on the pilot and that automation cannot replace the pilot. Thus, younger pilots then are more likely to accept and trust Auto-GCAS. Given that the current composition of F16 pilots is made of mostly younger generation, it would be interesting to gain deeper understanding of the generational issues by observing their behavior and actions in using Auto-GCAS in actual operation.
- The approval of Auto-GCAS's funding was influenced by how its business case is developed and marketed to users and decision makers [15]. Auto-GCAS was often thought of as a safety system. In the interviews, it was revealed that combat systems often have a higher funding priority than safety systems. We learned of the idea that "CFIT weeds out the weak." There is still an ideology with some decision makers that good pilots don't crash, and therefore would not need a technology like Auto-GCAS. To address this, Auto-GCAS was marketed as a combat and safety improvement to both decision makers and end users. The background and credentials of the individual(s) who make and present this business case are also important [15]. For example, ETPs stated in the interviews that they are more likely to trust a new technology when it is presented by a peer pilot (instead of a manager or an engineer). Thus, it can be seen that inherent organizational ideological differences may create a tier system of importance for different types of automated technologies (e.g., preference for combat capability improvements over safety capability improvements). This tier system can foster distrust in systems lower on the totem pole.
- ETPs reported that in the survey and interviews, healthy skepticism (also known as automation suspicion) is an integral part of their training, and is an essential part for developing and calibrating their trust as they gain more exposure and knowledge that the system works properly. The trust continuously calibrates to an appropriate level over repeated positive/negative exposures and experiences with the system. This notion of trust calibration is consistent with and substantiated by data collected on the ETPs' trust evolution. Specifically, we asked ETP participants to mark their initial and final trust of Auto-GCAS as a function of its perceived capability on a graph where a 45-degree line represents ideal trust calibration, i.e. a 1:1 relationship between trust and capability. As such, this line separates the over-trust region (above the line) and the under-trust region (below the line). The results are shown in Figure 2 for 11 ETPs who completed the graph. The trend shows that pilots tend to begin at under-trust points but as they gain experience and exposure to Auto-GCAS, they began to calibrate their trust close to the ideal trust line. However, their final trust is still slightly under the ideal line, indicating that ETPs maintain a dose of healthy skepticism in their trust calibration. This tendency of skepticism can be attributed to a number factors, such as avoiding strong negative consequences of over-trusting, not having enough experience with the system, or the training philosophy inculcated in them during test pilot school. It would be important in future field studies to determine which factors are the most dominant and at what point they are most relevant during the trust calibration process.
- While Auto-GCAS was designed to prevent up to 98% of historical incidents and to work only for situations when the pilot is saturated with tasks or disoriented or unconscious (due to high acceleration), pilots can misuse the system if they over-trust it and believe that "Auto-GCAS will always save them." This misbelief can cause the pilot to fly more aggressively or brazenly, and misuse Auto-GCAS as a combat tool. Of the 15 participants, three commented that having Auto-GCAS would change the way they fly. Two participants also had a long history of exposure to Auto-GCAS and had opportunities to observe changes in flying of other pilots. One pilot stated that he "actually few lower more comfortably" as a result of flying with Auto-GCAS. A plausible explanation for this is that the pilot occupational culture pushes pilots to strive to become better fighters and to use every tool that they have to their advantage. Thus, Auto-GCAS then can be misused by future operational pilots. It would be beneficial to follow Auto-GCAS during actual operation to learn whether misuse would occur and the circumstances and reasons when this happens.

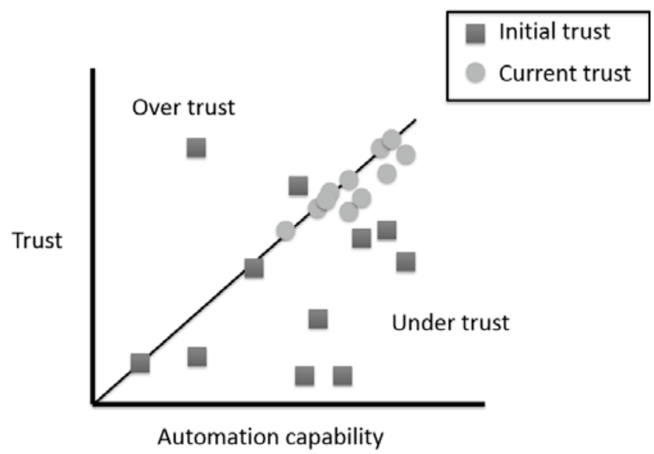


**Figure 2**: Self report of trust evolution of ETPs with Auto-GCAS

- Pilots are trained to be independent, self-reliant, and make life-or-death decisions using real-time data, thus it would be reasonable to hypothesize that they would dislike letting automation take away control of the aircraft. To gain a better understanding into this attitude, and its implications for automatic systems such as Auto-GCAS, participants were asked to rate the highest level of Auto-GCAS automation that they would be comfortable with. The Sheridan and Verplank [13,14] taxonomy on 10 automation levels were used, where level 1 is lowest level (Human does the whole job up to the point of turning it over to the computer to implement) and level 10 is the highest level (Computer does the whole job if it decides it should be done, and if so, tells human, if it decides that the human should be told). The majority of pilots indicated that they would be comfortable with level 7 (Computer does whole job and necessarily tells human what it did). This result at first seems surprising, but from a perspective of trust calibration and evolution, it reinforces the notion that having an extensive and positive experience with a technology facilitates acceptance of that technology's high automation levels. It gives evidence that pilots can be accepting of a highly autonomous system that takes control away from them. Again, this result needs further validation with data from actual operation.

- Automation transparency was an important topic that emerged in this study as it affected how pilots develop and calibrate their trust. Auto-GCAS has several different display features that contribute to its transparency to a pilot, one of which is the use of chevrons on the heads up display (HUD) as shown in Figure 3 (the chevrons are two arrow heads pointing at each other). The chevrons will appear from the left and right sides of the HUD when the Auto GCAS systems detects a potential collision. These chevrons move toward the middle of the HUD and intersect at a TAAGI of 1.5 seconds, when a pull up maneuver occurs. The interviews revealed a debate over whether the inclusion of chevrons is a positive or a negative feature of Auto-GCAS. Pilots provided two opposing arguments on whether chevrons should be an included feature on the fielded version of Auto-GCAS and if the chevrons should default to "on." The chevrons provide feedback and help increase pilot awareness of an impending pull-up maneuver, however, it is a concern that the presence of the chevrons can alter the way in which pilots fly. For example, it can encourage them to become more brazen and aggressive. Pilots may also attempt to fly closer to terrain and use the chevron symbology as an indicator of how close they are to collision. However, some pilots argue that the chevrons are beneficial, as they give feedback to the pilot and allow the pilot to execute his/her own maneuver before the automation. The decision to have the chevrons be on or off by default during actual operation is a hotly debated topic. Typically, Developmental Test (DT) pilots believe that chevrons should default off while Operational Test (OT) pilots believe they should default on. These debates suggest that the design for automation transparency needs to take into account and balance the benefits of having information and the misuse of the information. It also seems reasonable to hypothesize that if ETPs have already calibrated trust during the developmental stages with transparency-driven features such as the chevrons, then operational pilots may not need to have the chevrons. This also suggests that differing groups of pilots may need differing types of transparency to facilitate appropriate trust. These are questions that deserve further examination.

- By the high standards of their profession, only the best of the best pilots are selected to fly fighter aircraft. As such, these pilots possess a mindset of supreme confidence and low vulnerability, which they need in order to be effective in high-risk combat environments. This was supported by participant observations and responses in the surveys and interviews. By the extension of this mindset, a number of pilots who commented in their interviews that Auto-GCAS will be beneficial for their colleagues but they will not need it themselves (because they are competent/good). Pilots had a tendency to emphasis the need for Auto-GCAS as a benefit to the fighter pilot community, but not as a need for themselves. This attitude can lead to potential disuse of the system, and needs further investigation


**Figure 3**: HUD image with Chevrons

- Misinformation can have detrimental effects on the opinion and trust of a system because it affects the user's initial perception of the system [5]. During the course of the project, a number of instances of misinformation with respect to Auto-GCAS appeared during interviews and participant observations. This included gossip that the system doesn't work (when in fact Auto-GCAS was not installed on a crashed aircraft), or pilots believing that they have flown Auto-GCAS (when in fact they have flown with other systems). The misinformation can be attributed to the confusion from the heavy use of acronyms within government groups, the fielding of different systems with the same or similar names, the long developmental history of Auto-GCAS, and potential rumors spread by Auto-GCAS opponents. Consequently, distrust can be formed due to misinformation, creating a hurdle in fostering positive trust development.

## DISCUSSION & FUTURE RESEARCH

### Discussion

The research design of the extended case study and the implementation strategies were effective for conducting the study in a military environment that possessed numerous challenges. The methodological approach developed can be applied to study the cultural/organizational dimensions of human-automation trust in other technological systems.

The lessons learned, best practices, and research issues emerged from this study add to the body knowledge in trust research in a number of ways. First, the present results speak to the importance of transparency [7] as an influence on trust of automated systems. Specifically, the current results show that avoiding false alarms by the system was beneficial to the trust development of the ETPs. The system was designed and tested to approximate the appropriate threshold of constraints (i.e., time available to avoid ground impact) before initiating an action. This relates to the social intent facet of transparency described by Lyons [7], and demonstrates the importance of system design having a focus on the needs/wants of the user. Not only does this suggest that systems should evidence "benevolent intentions" but it also shows the costs of false alarms within automated systems, particularly for safety systems that assume control of one's aircraft. Second, the study shows that trust of automated systems can be calibrated through exposure and experience in real-world scenarios. In this instance, after pilots had experienced the successful initiation of the Auto-GCAS system, they were more likely to have positive trust perceptions associated with it. This consistent with prior research by Hancock and colleagues [4] who found that system performance was the strongest predictor of human trust of robotic systems. Finally, the present study confirms the potential for anomalous unintended consequences associated with automated systems [11]. Safety systems such as Auto-GCAS are designed to promote safe flight operations, and it is very effective at doing just that. Yet, the current study suggests that there may be times where a pilot engages in more risky flight patterns due to the understanding that the Auto-GCAS will protect them from a crash. This finding is an exemplar for Parasuraman and Riley's concept of "misuse" [11], caused by an over trust of a system. The use of a transparent HUD display (i.e., the Chevrons) was implicated as a potential root cause driving the future possible overreliance (misuse) on Auto-GCAS in operational F-16 pilots. This raises important issues for the research community to investigate what the appropriate mechanisms are to foster appropriate transparency for safety systems that are used in high-risk operations, such as combat. Warfighters naturally seek to attain the highest level of performance possible, and as such, real-time transparency inputs may be used to calibrate one's sense of boundary conditions afforded by one's aircraft. There are a number of ways to instantiate transparency into human-machine contexts [7], and future research is certainly warranted to help engineers anticipate how humans will respond to novel designs. This study's findings can be used as the basis to explore how transparency mediates trust development. This can be accomplished with empirical studies in the context of the chevrons in Auto-GCAS and its sister Air Force Auto-Air Collision Avoidance (Auto-ACAS) system, which is currently under development and testing. These empirical results could then be used to spearhead applied research in creating transparency-oriented design principles and for infusing them into various phases of system development cycle.

### Limitations

While the current study provided rich data from which to evaluate the utility and constraints associated with a complex form of automation, there were several limitations resulting from the methods employed. First, the target sample for the initial work involved test pilots and engineers associated with the Auto GCAS development and testing. The involvement of these individuals, while necessary, may have introduced a positivity bias into the results. Further, pilots without the extensive background/history with Auto GCAS may react differently to the technology than the ETPs who were very experienced in the technology. Future research is needed to examine reactions to Auto GCAS from pilots who were not involved in the development and testing of the system. A second limitation involves the lack of experimental methods used in the study. Given that the data were qualitative, one cannot derive causal inference from the data. Experimental studies are needed in order to identify causality within this domain, and one such experiment is briefly described below.

### Conclusions & Future Research Avenues

The findings in this study suggest that Auto-GCAS design was in alignment with pilot culture and organizational mission, indicating a strong potential for appropriate trust development in operational pilots. However, these findings are based on pre-deployment data. Because trust calibration is a dynamic process, spanning from the time when the system is first conceived until it goes into operation and retirement, it would be beneficial to conduct a field study of the deployment of Auto-GCAS with operational pilots. This is in order to collect data to validate the hypotheses and to examine the various research issues raised (e.g., potential Auto-GCAS misuse/disuse due to pilot occupational culture and/or operational circumstances, trust evolution from beginning of deployment to stages when opinions are stabilized). Such a field study would generate data and results that could influence and improve the design of the class of systems that take away control from the operator while eliminating nuisance activations and preventing interference with the mission.

Future research may also focus on better understanding trust in automation disparities originating from cross-cultural differences, generational differences, and individual differences. Cultures align to various forms of social norms, expectations, and value sets. Prior research suggests that there are cultural differences in the domain of interpersonal trust [15], therefore differences in trust of automated systems in different cultural groups warrants future research consideration. Similar to individuals of a different cultural background, individuals of different age groups may also have differing expectations and comfort levels associated with novel technology. Research has shown that younger and older samples report different trust levels for novel automated tools involving anthropomorphic features, with younger individuals reporting higher trust relative to older individuals [10]. Thus, the speculation of younger pilots being more "comfortable" or "accepting" of the Auto-GCAS should be taken seriously and these effects should be given more systematic analysis. Lastly, future research should examine how individual differences influence reliance on automated tools. One promising domain that links to the current study is the construct of suspicion [1]. ETPs reported an inherent suspicion of new tools, which may be indicative to their career field or training. Given recent advances in the domain of state suspicion in IT contexts [1] further research should examine what features of the technology, training, culture, or organizational doctrine are related to suspicion.

#### Potential Experimental Design

The qualitative data suggested that the use of Chevrons as a form of transparency be examined further to determine if this manipulation has a positive or negative impact on pilots use of the system. This experiment could explore how the presence of the Chevrons changes pilot behavior. For instance, using a repeated measures experiment researchers could provide some pilots with the Chevrons and others not during a simulation of the Auto GCAS system and explore whether one group evidences significant changes in altitude relative to a baseline condition. Other key metrics that could be included in this study are: trust in automation, efficacy in piloting at low altitudes, perceived risk, and attitudes toward safety.

## ACKNOWLEDGMENTS

This project was funded via NASA Dryden Research Flight Center by the Air Force Office of Scientific Research, Trust and Influence Portfolio. We'd like to thank Mark Wilkins of the DOD, members of the 416th division at Edwards AFB, and Tim McDonald and the SETP for their support; and students who have been involved in this project.

## REFERENCES

1. Bobko, P., Barelka, A., & Hirshfield, L. The construct of state-level suspicion: A model and research agenda for automated and information technology contexts. *Human Factors*. (in press).
2. Burawoy, M. *Ethnography Unbound: Power and Resistance in the Modern Metropolis.* University of California Press, Berkeley CA, 1991.
3. Colquitt, J.A., Scott, B.A., LePine, J.A. Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology, 92*, (2007). 909-927.
4. Hancock, P.A., Billings, D.R., Schaefer, K.E., Chen, J.Y.C., Ewart, J., Parasuraman, R. A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*(5), (2011). 517-527.
5. Jones, G. R., & George, J. M. The experience and evolution of trust: Implications for cooperation and teamwork. *Academy of Management Review, 23*, (1998). 531–546.
6. Lee, J. D., & See, K. A. Trust in technology: Designing for appropriate reliance. *Human Factors, 46,* 2 (2004), 50-80.
7. Lyons, J.B. Being transparent about transparency: A model for human-robot interaction. *Proceedings of the AAAI Spring Symposium on Trust in Autonomous Systems* (Palo Alto, CA, March 2013), pp. 48-53.
8. Mapes, P. B. Fighter/Attack Automation Collision Avoidance Systems Business Case. Air Force Research Laboratory, Wright Patterson Air Force Base. (2006).
9. Mayer, R.C., Davis, J.H., & Schoorman, F.D. An integration model of organizational trust. Academy of Management Review, 20, (1995). 709-734.
10. Pak, R., Fink, N., Price, M., Bass, B., & Sturre, L. Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics, 1-14*. (2012).
11. Parasuraman, R. & Riley, V. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors, 39*, 2 (1997), 230-253.
12. Ring, P. S., & Vandeven, A. H. Structuring cooperative relationships between organizations. *Strategic Management Journal, 13*, (1992). 483–498.
13. Sheridan, T. B. Considerations in modeling the human supervisory controller. *In Proceedings of the IFAC 6th World Congress.* (1975). Laxenburg, Austria: International Federation of Automatic Control.
14. Sheridan, T. B., & Verplank, W. Human and Computer Control of Undersea Teleoperators. Cambridge, MA: Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT (1978).
15. Yuki, M., Maddux, W.W., Brewer, M.B., & Takemura, K. Cross-cultural differences in relationship- and group-based trust. Personality and Social Psychology Bulletin, 31(1), (2005). 48-62.