Training student air traffic controllers to trust automation

Adriana Miramontes*, Andriana Tesoro, Yuri Trujillo, Edward Barraza, Jillian Keeler, Alexander Boudreau, Thomas Z. Strybel, Kim-Phuong L. Vu

Center for Human Factors in Advanced Aeronautics Technologies (CHAAT), Department of Psychology, California State University, Long Beach

Abstract

The Next Generation Air Transportation System (NextGen) will implement new automation tools to allow controllers to effectively manage the projected increase in air travel over the next decade. In order for the implementation of NextGen tools to be successful, it is important that air traffic controllers (ATCos) develop appropriate levels of trust in these automated tools. The present study investigated whether students could be trained to trust automation in a NextGen environment, and whether their trust in automation would affect their air traffic management performance. We also examined if personality traits influence the student’s likelihood to trust automation. We found an early benefit in terms of efficiency in air traffic management for students who were trained to trust automation. We also found that people with high emotional stability reported higher levels of trust in automation.

Keywords: Trust in automation; NextGen; Training; Air traffic controller performance; Personality

1. Introduction

Air travel is expected to increase over the next decade, which could result in increased demands on air traffic controllers (ATCOs) managing aircraft in the National Airspace System (NAS). To assist ATCos with managing the increased traffic levels, the Next Generation Air Transportation System (NextGen) is developing automation tools for operators [1]. One benefit of implementing automated tools in NextGen is that these tools should produce reductions in ATCo’s workload, meaning that ATCos can maintain or increase traffic safety despite the increased...
traffic levels. For example, Vu et al. (2012) found that when ATCos were given automation tools to use, they were able to manage higher traffic levels while reporting lower workload levels and making fewer safety violations (i.e., fewer losses of separation, LOS) [2]. However, the benefit of automation cannot be obtained if the ATCos do not trust and use the automation appropriately.

Parasuraman and Riley (1997) identified one type of misuse of automation, called disuse. Disuse is the operator’s unwillingness to use automated tools because he/she does not trust the automated system. An operator is usually unwilling to use automation when they feel the system is unreliable or when there are high false alarm rates. For systems that have high costs for a miss, a user is less likely to rely on the automated system. A possible solution to increase use of automation is to give the operator more exposure to the automated system.[3]. Therefore, one method for influencing operator trust in automation is through training. Higham et al. (2013) found that student ATCos could be trained to trust automation and that student ATCos who were trained to trust the automation reported lower workload levels under some conditions compared to students who were not trained to trust the automation [4].

Another factor that may affect an individual’s likelihood to trust automation is personality type. Although, we are not aware of any research directly relating ATCo personality traits to his/her tendency to trust automation, Hoff and Bashir (2014) found that people with high emotional stability and low neuroticism are more likely to rely on automated systems in general [5]. This relationship between neuroticism and trust in automation has been found in different domains. For example, Szalma and Taylor (2011) found that individuals with high levels of neuroticism were less likely to agree with automation while completing a threat detection task[6]. A possible explanation for why individuals with high neuroticism lack trust in automation is that individuals with high neuroticism are more likely to be cautious in decision making[7]. With regard to air traffic control, personality traits have shown to predict ATCo performance. For example, Durso et al. (2006) found conscientiousness predicted ATM performance, and Schroeder and Dollar (1997) found that emotional stability was an important determinant of the success of new ATCo applicants [8,9].

In the present study, we investigated whether trust in automation could be trained in a NextGen environment, similar to Higham et al. (2013). However, we also looked at other factors related to trust. Specifically, we investigated whether trust in automation changed over time as a function of training and whether the students’ level of trust in automation affected their air-traffic-management performance in terms of safety and efficiency. The relationship between trust and personality was also examined to determine if conscientiousness and emotional stability could predict trust in automation since they were shown previously to be predictive of ATCo performance. Our research questions were:“Does trust in automation change over time as a function of training?”,”Does trust training affect ATCo performance?”, and “Are conscientiousness and emotional stability predictive of trust in automation?”.

2. Methods

2.1. Participants

A secondary analysis was performed using data from 48 student air traffic controllers from the Aviation Sciences Program at Mount San Antonio College, an FAA CTI institution. These students participated in a 16-week radar simulation internship at the Center for Human Factors in Advanced Aeronautics Technologies (CHAAT) in one of four semesters (two calendar years). Demographics were available for 44 participants (37 males and 7 females; Age $M = 24.75$ years, $SD= 2.74$ years). Since data were collected over 4 semesters, there were some differences in the internship content that limited the available data for some tests. Data from all participants were analyzed to determine the relationship between personality traits and subjective measures of trust in automation. However, only 19 participants’ data were available to determine the effects of training trust in automation.
2.2. Materials and apparatus

Participants were trained and tested using the Multi Aircraft Control System (MACS) software [10]. MACS is a medium fidelity simulator and it simulated a transitional sector, ZID-91, which included overflights, departures from and arrivals to Louisville International (SDF) Airport. During training sessions, students in the internship took turns as ATCos or pseudopilots. Pseudopilots responded to and executed the ATCos clearances. For the experimental sessions that occurred during the midterm and final weeks, confederate researchers served as the pseudopilots. Participants were trained to use the following NextGen tools: conflict probe, Data Comm, and trial planner. The conflict probe uses algorithms to detect conflicts (pairs of aircraft that would lose separation unless one or both changed its flight plan.) Data Comm allows ATCos to send digital messages for changing frequencies, and uplinking clearances. The trial planner allows ATCos to adjust an aircraft’s route by “clicking and dragging” and uploading it to the aircraft through Data Comm. Use of these NextGen tools is expected to increase ATCo efficiency by helping operators managing traffic more effectively. The tools can also improve ATCo safety by providing them with a visual alert to detect and move conflict aircraft more effectively.

Participants were given the Potential for Complacency Questionnaire, a measure of trust in automation that is based on a 5-point scale with 1 indicating low trust in automation and 5 indicating high trust in automation [11]. They also were given the Revised NEO Personality Inventory, in which participants rated their conscientiousness and emotional stability [12]. The personality questionnaire was also a 5-point scale, with 1 being low on these personality dimensions and 5 being high. The personality questionnaire was administered at the beginning of the course and the trust in automation questionnaire was administered at the end of the course.

2.3. Procedure

The radar simulation course consisted of 3.25 hours of radar lab training and 1.5 hours of classroom lecture each week. A retired, radar-certified ATCo taught all sessions of the internship and supervised the training. Students were enrolled in one of two lab sessions, which differed in the type of trust training the participants received. For the first 8 weeks of the internship, the trust-training group was trained on scenarios having 75% NextGen equipped aircraft and the control group was trained on scenarios having 25% equipped aircraft. Therefore, the trust training group had more practice with the NextGen tools and the control group had more practice with the traditional manual/voice-based tools. After the 8th week, both groups trained on scenarios having 50% NextGen equipped aircraft. However, the trust training group received feedback from the instructor when they moved one of a pair of near miss aircraft, while the control group continued to receive normal training. A near miss occurred when a pair of equipped aircraft came close to the separation minima but never lost separation; therefore, they would not be alerted by the automated tools. Moving a near miss aircraft was used as a behavioral indication of mistrusting the automation because it would lead to less efficient air traffic management and possibly reduce safety when unnecessary changes to aircraft flight plans were made.

Participants were tested at midterm during the 8th week and final during the 16th week of the internship. For the purposes of this study, only a subset of the test scenarios were analyzed to answer our research questions. In particular, we examined the 50% equipped scenarios, which contained one near miss. Each participant had one 50% equipped scenario and therefore there was only one near miss per participant. To measure the effectiveness of trust training we recorded the number of near misses moved. To examine performance measures of safety we recorded the number of LOS and for performance measures of efficiency we computed the average time through sector. The personality and trust questionnaires were used to assess the relationship between personality traits and trust in automation.

3. Results

We set our alpha level at .10 for a more liberal criterion to increase power because of the small sample size. Three mixed factorial ANOVAs were run with trust training as a between subjects factor and exam period as a within subjects factor.
3.1. Trust

To assess whether trust in automation changed over time as a function of training, a 2 (trust training: trust/control) x 2 (exam period: midterm/final) mixed factorial ANOVA was performed with the average number of near miss aircraft moved as the dependent measure. The main effects of trust training, $F(1,17)=6.79$, $p=.018$, and exam period, $F(1,17)=4.08$, $p=.06$, were significant, but the interaction of these two factors, as shown in Figure 1, was not significant, $F(1,17)=0.56$, $p=.46$. The trust training group ($M=0.50$) moved fewer near miss aircraft than the control group ($M=0.94$) over both exam periods. All students moved more near miss aircraft at the midterm ($M=0.82$) than at the final ($M=0.62$).

3.2. ATCosafety performance

To assess whether trust training had an effect on ATCo performance measures of safety a 2 (trust training: trust/control) x 2 (exam period: midterm/final) mixed factorial ANOVA was performed with the number of LOS as the dependent variable. No significant effects were obtained, $F$s $< 1.40$, $p$s $>.25$.

3.3. ATCoefficiency performance

To assess whether trust training had an effect on ATCo performance measures of efficiency a 2 (trust training: trust/control) x 2 (exam period: midterm/final) mixed factorial ANOVA was performed with average time through sector as the dependent measure. The main effects of trust training, $F(1,17)=0.60$, $p=.45$, and exam period, $F(1,17)=0.31$, $p=.58$, were not significant. However, a significant interaction between trust training and exam period, $F(1,17)=2.87$, $p=.10$, was obtained. Simple effects analysis of trust training on exam period indicated that trust training affected average time through sector for midterm, $F(1,17)=3.67$, $p=.07$, but not for final, $F<1.0$ (See Figure 2). Specifically, at midterm the trust group ($M=620.93$ s) was more efficient than the control group ($M=668.21$ s). At the final exam, the differences between training groups in time through sector were not significant.
3.4. Personality

To assess whether emotional stability and conscientiousness could predict scores on the Potential for Complacency trust in automation scale, a simultaneous multiple regression was run. It was found that emotional stability and conscientiousness accounted for 12.9% of the variability in trust in automation ($R^2=.129$, adjusted $R^2=.090$) and both were significant predictors of trust, $F(2,45)=3.32$, $p=.045$. Although, conscientiousness was not found to be a significant unique predictor ($\beta=-.044$, $p=.819$) of trust in automation, emotional stability was a significant factor ($\beta=.387$, $p=.049$).

4. Discussion

The main goal of our study was to determine whether trust in automation could be trained in a NextGen environment. Our results indicate that trust in automation can be trained in a NextGen environment because the trust training group moved fewer near miss aircraft (a behavioral indication of trust in automation) than the control group. Furthermore, student ATCos moved more near-miss aircraft at the midterm than at the final. This indicates that with more experience with the automation over the course of the semester, both groups developed better trust in automation.

We also wanted to determine if trust in automation could affect an ATCo’s performance. We examined performance measures of safety (LOS) and efficiency (average time through sector). We found that trust training did not have an effect on ATCo safety. However, this may be due to a floor effect since there were very few LOS ($M=0.61$). In terms of efficiency, we found an early benefit for students who were trained to trust the automation. The ATCo students in the trust training group were more efficient in air traffic management at the midterm because they relied on the automation tools for air traffic management. This is likely due to the high early exposure to NextGen tools during the internship.

Finally, we were interested in the personality factors that could affect an individual’s likelihood to trust automation. We did not find that conscientiousness was a unique predictor of trust in automation. Instead, we found that ATCo students with high emotional stability reported higher levels of trust in automation in a NextGen environment. This finding is consistent with those obtained by Hoff and Bashir (2014), who found that individuals with high emotional stability were more likely to rely on automated systems in general.

In summary, we found that student air traffic controllers could be trained to trust automation in a NextGen environment. Furthermore, we found that trust training did not affect performance measures of safety but there was
an early benefit for trust training, because ATCo students were more efficient in managing air traffic. Finally, we found that high emotional stability is related to having higher levels of trust in automation.

Acknowledgements

The project was supported by NASA cooperative agreement NNX09AU66A, Group 5 University Research Center: Center for Human Factors in Advanced Aeronautics Technologies (Brenda Collins, Technical Monitor).

References

[1] Joint Planning and Development Office (JPDO)(2010). Concept of operations for the next generation air transportation system version 3.2. Accessed from http://jpe.jpdo.gov/ee/dpcs/conops/Nextgen_ConOps_v3_2.pdf.
[2] Vu, K. P. L., Silva, H., Ziccardi, J., Morgan, C. A., Morales, G., Grigoleit, T., Lee, S., Kiken, A., Strybel, T. Z., & Battiste, V. (2012). How does reliance on automated tools during leaning influence students’ air traffic management skills when the tools fail?. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting 56(1), 16-20.
[3] Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors: The Journal of the Human Factors and Ergonomics Society, 39(2), 230-253.
[4] Higham, T. M., Vu, K. P. L., Miles, J., Strybel, T. Z., & Battiste, V. (2013). Training air traffic controller trust in automation within a NextGen environment. Human Interface and the Management of Information. Information and Interaction for Health, Safety, Mobility, and Complex Environments, 76-84.
[5] Hoff, K. A., & Bashir, M. (2014). Trust in automation integrating empirical evidence on factors that influence trust. Human Factors: The Journal of the Human Factors and Ergonomic Society, 0018720814547570.
[6] Szalma, J. L., & Taylor, G. S. (2011). Individual differences in response to automation: the five factor model of personality. Journal of Experimental Psychology: Applied, 17(2), 71.
[7] Matthews, G. (2008). Personality and information processing: A cognitive-adaptive theory. The SAGE Handbook of Personality Theory and Assessment, 1, 56-79.
[8] Durso, F., Bleckley, M., & Dattel, A. (2006). Does situation awareness add to the validity of cognitive tests?. Human Factors: The Journal of the Human Factors and Ergonomics Society, 48(4), 721-733.
[9] Schroeder, D. J., & Dollar, C. S. (1997). Personality characteristics of pre/post-strike air traffic control applicants. Federal Aviation Administration Washington D.C. Office of Aviation Medicine.
[10] Prevot, T. Exploring the many perspectives of distributed air traffic management: The Multi Aircraft Control System: MACS. International Conference on Human-Computer Interaction in Aeronautics, HCI-Aero 2002, 23-25 October, MIT, Cambridge, MA (2002).
[11] Verma, S., Kozon, T., Ballinger, D., Lozito, S., & Subramanain, S., (2011). Role of the controller in an integrated pilot-controller study for parallel approaches. Paper presented at the 30th Digital Avionics System Conference, Seattle, WA.
[12] Costa, P. & McCrae, R. (1992). The NEO Personality Inventory (revised) manual. Odessa, FL: Psychological Assessment Resources.