A Trust in Air Traffic Controllers (T-ATC) Scale

Bradley S. Baugh  
*Emory-Riddle Aeronautical University, Bradley.s.baugh@gmail.com*  
Scott R. Winter  
*Emory-Riddle Aeronautical University, scott.winter@mac.com*

Follow this and additional works at: https://commons.erau.edu/ijaaa

Part of the Aviation Safety and Security Commons, and the Human Factors Psychology Commons

**Scholarly Commons Citation**

This Article is brought to you for free and open access by the Journals at Scholarly Commons. It has been accepted for inclusion in International Journal of Aviation, Aeronautics, and Aerospace by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu, wolfe309@erau.edu.
Introduction

Safety within the National Airspace System requires a high degree of cooperation between system actors facilitated through trust in system components. The component most vulnerable to errors is the human components which is why it is vital to understand areas where human errors may occur. The interaction between pilots and air traffic controllers can be affected by a bevy of potential human factors issues. One area of interest is the trust relationship between pilots and controllers. Uniquely, this relationship is facilitated by communicating through radiotelephony without the benefit of prior personal relationships or non-verbal cues assessed visually. The relationship between pilots and controllers requires trust; lack of trust may lead to errors. However, trust between pilots and controllers is not well researched in the literature. Further, there appears to be no specific instrument in the literature to measure this facet of trust in aviation. Therefore, the purpose of this research is to develop a valid and reliable instrument to measure a pilot’s trust in air traffic controllers.

Literature Review

Pilots bear the responsibility to safely operate their aircraft (General Operating and Flight Rules, 2019). However, operating conditions and airspace complexity sometimes require pilots to rely on assistance from air traffic controllers. Pilots must have a level of trust in air traffic controllers; otherwise, safety and efficiency of the system are degraded. The risk is that misplaced trust can spell disaster. Consider, for example, the US-Bangla Airlines Flight 211 crash in Kathmandu, Nepal (Accident Investigation Commission, 2018). The aircraft crashed, resulting in 51 fatalities. According to the report, the controller failed to communicate clear instructions (including providing missed approach procedures) and failed to monitor the aircraft flight path.

Given the gravity of broken trust between pilots and air traffic controllers, it is striking the lack of literature on pilot-controller trust. Trust can be understood in many ways. The current study adopts the understanding of Mayer, Davis, and Schoorman (1995) which involves vulnerability, risk, and expectations to act in a certain way without the ability to directly control the other party.

Trust and Trustworthiness

Trust is a complex concept with many nuances and variations that can be situationally dependent. Studies beginning in the mid-20th century began experimentally identifying and cataloging trust factors and antecedents of trust. While there are dozens of trust theorists, there are a few that continually figure prominently in the literature including Deutsch (1958), Giffin (1967), Rotter (1971), Luhmann (1979), and Lewis and Weigert (1985).
Trust and suspicion was the focus of one of the earlier studies by Deutsch (1958) using a set of experiments to study trust (where the outcome is expected to positive) and suspicion (where the outcome is expected to be negative). He wrote that trust has elements of expectation, predictability, motivation, intentions, and obligation. Using a two-person non-zero-sum-game where wins and losses relate to choices of both individuals in a variation of the prisoner’s dilemma, Deutsch (1958) found that individuals who had a co-operative orientation produced mutually beneficial outcomes. Conversely, competitive orientation produced a mutual loss. Participants who were individualist oriented were more likely to trust if they could freely communicate with each other.

Giffin (1967) built on Deutsch’s (1958) work focusing on the relationship of communication in interpersonal trust defining trust as “reliance upon the characteristics of an object, or the occurrence of an event, or the behavior of a person in order to achieve a desired but uncertain objective in a risky situation” (p. 105). Trust is based on impressions, the impressions can be measured, and there is a suggestion that trust is situational and operates on a continuum. Further, trust judgments in the communications process are partially based on the credibility of the source. Interpersonal trust contains judgments of expertness, reliability, intentions, dynamism, and personal attraction.

Rotter’s (1971) definition of trust adds the element of expectancy that the verbal or written statements of the other party can be relied on suggesting trust is action-oriented. Underlying the expectancy is a judgment of the believability of the other party. Luhmann’s (1979) contribution relative to the current study is the purpose of trust. Humans must trust in a broad sense as a mechanism to cope with the complexities of life. People often have the freedom to choose who and what to trust, and without trust, humans would get overwhelmed. According to Luhmann (1979), “a complete absence of trust would prevent [a person] from even getting up in the morning” (p. 5).

Lewis and Weigert (1985) fused the works of Luhmann (1979) and others to explain trust from a sociological standpoint as comprised of two dimensions; cognitive and emotional. Building on notions that trust is foundational to a functioning society, the authors note the ever-present component of risk. Further, they acknowledge the role of trust as a mechanism for reducing complexity (see also Luhmann, 1979). The cognitive aspect of trust relates to categorization and judgments of who to trust and when. The emotional aspect of trust explains investments in the relationship.

In the psychology and sociology literature, trust differs from trustworthiness. Trust implies the willingness to take a risk in the other party, whereas trustworthiness refers to the actions, behaviors, and manifestations of the other party who has an awareness of the trust placed in them (Deutsch, 1958). In common usage, trust and trustworthiness are used interchangeably (Rice &
Kramer, 2014; Rice, Mehta, Winter, & Oyman, 2015; Rice, Winter, & Tokarski, 2015), and are positively associated with each other (Rotter, 1971). Levine et al. (2018) summarized the relationship between trust and trustworthiness explaining the benefits of trust are only realized “when trust is well placed, in targets who are trustworthy” (p. 468).

Trust in Air Traffic Control

No specific studies between pilots and air traffic controllers were located. However, there have been various trust studies in the air traffic control and aviation domain. Mehta, Rice, and Rao (2016) researched trust in air traffic controllers from a traveler’s perspective. A recent study by Svennsson, Ohlander, and Lundberg (2019) researched internal trust dynamics in air traffic control teams. Other recent studies in aviation looked at the relationship of communication and trust in aviation maintenance (Chatzi, Martin, Bastes, & Murray, 2019) and passenger perceptions of trust, risk, and flying in adverse weather (Chang, 2017). Predominant in the literature are studies related to trust in automation in aviation (Gontar, Homans, Rostalski, Behrend, Dehais, & Bengler, 2018; Lim, Gardi, Sabatini, Ramasamy, Kisten, Ezer,…& Bolia, 2018; Lyons et al., 2017; Rice, Winter, Deaton, & Cremer, 2016).

Delphi Group Consensus

The Delphi method is a technique to elicit group consensus from experts in a manner, reducing the likelihood of bias effects between team members. It was developed in the 1950s and 1960s through work with the U.S. Air Force and the RAND Corporation (Dalkey, 1969; Dalkey & Helmer, 1963; Linstone & Turoff, 1975). The Delphi method is characterized by several rounds of anonymous questioning (through questionnaires and interviews), controlled feedback, and statistical group response resulting in a general consensus on a topic (Dalkey, 1969). The Delphi methodology seeks to avoid direct confrontation and reduces a tendency towards the negative aspects of groupthink (Janis, 1982). It promotes greater participation, creativity, and open-mindedness to novel ideas. Misconceptions can be explored and corrected without fear of embarrassment.

Early Delphi studies were conducted in person or using mail sent through the postal system. Over the past half-century, Delphi techniques have evolved, and not all Delphi methods strive for consensus but more of a preferred direction (Hasson & Keeney, 2011). More recently, Delphi studies have employed technology email and other online technology to facilitate research (Hasson & Keeney, 2011; Weise, Fisher, & Trollor, 2016)

Methodological rigor is an important part of any technique. One consideration in the literature in the size requirements of the expert panel. References to panel size range from four (Skulmoski, Hartman, & Krahn, 2007)
to well over 1,000 (Powell, 2003). Dalkey and Helmer’s (1963) study employed a panel of 7 experts. The key point is to have a diversity of viewpoints from varied backgrounds (Powell, 2003), should be knowledgeable in their sphere, and impartial to any findings (Jairath & Weinstein, 1994).

**Methodology**

The current study models the methodology used to design the Trustworthiness of Commercial Airline Pilots (T-CAP) scale (Rice et al., 2015). Because constructs such as trust are not directly measurable, summative scales are developed using words and phrases that when taken as a whole represent the latent construct. The methodology involves word and phrase generation followed by statistical testing to assess the validity and reliability of the constructs. The goal at the outset was to create a six to eight question scale for use in assessing a pilot’s trust in air traffic controllers. A team of 11 aviation experts and practitioners was assembled and words generated using a virtual Delphi method. The Delphi method is used to obtain consensus through anonymity (Weise et al., 2016). None of the experts knew the others, only that there were 10 other members of the team.

**Stage 1: Word Generation**

During the first stage, words and phrases representing trust and trustworthiness in air traffic controllers were generated using an anonymous online questionnaire.

**Participants.** A team of 11 aviators (2 female, 9 male) was used for word generation. Ages ranged from 18-50 ($M = 33.4$, $SD = 10.87$). Aeronautical certifications held comprised Private Pilot (3), Commercial Pilot (4), Airline Transport Pilot (4), and military certification (4). (Note: Certifications were greater than 11 because four pilots held both civilian and military certifications). Two pilots identified as certified flight instructors and four as certified flight instructor-instrument. Flight training sources varied from 14 CFR Part 61 operations (2), 14 CFR Part 141 schools (5), and military (4). Current affiliation spanned a flight school (1), university/academia (4), airline (2), and military (4). Finally, 10 pilots had fixed-wing aircraft experience, and one had rotor-wing aircraft experience.

**Materials and stimuli.** The participants were asked to consider the following scenario: “Imagine an air traffic controller who is trustworthy. In the context of the air traffic controller mentioned previously, please enter five (5) characteristics of a trustworthy air traffic controller in the spaces proved below. Each answer should include only one word or a short phrase.” After reading the scenario, the participants were allowed to record their responses in a free-text manner.
Stage 2: Nominal Paring
The second stage was used to pare the list of 33 distinct words developed in Stage 1.

Participants. The same team of 11 aviators used for word generation participated in the nominal paring exercise.

Materials and stimuli. The participants were asked to consider the following scenario: "In the context of an air traffic controller, please rate whether each word below is related to (similar to) air traffic controller trustworthiness, not related to (not similar to) air traffic controller trustworthiness, or you do not know." The participants were then presented with the 33 words.

Stage 3: Scenario-Based Testing
Stage 2 produced a list of six terms with 100% agreement among the Delphi team. Because this met the study goal of creating a list of six to eight terms, further paring was not required. The six words were then set together into a proposed scale. Scenario-based testing was used to assess the underlying construct of the proposed scale.

Participants. Over 200 pilots took part in Stage 4 using a convenience sample from online and university pilots with U.S. (FAA) flight certifications. Upon initial review, 201 responses were deemed usable (22 female, 174 male, 5 prefer not to answer) with two removed for missing data or study ineligibility. Online pilots volunteered to participate through a solicitation in the Curt Lewis Newsletter, a daily news service with worldwide readership of approximately 100,000 subscribers. University pilots from a large southeastern U.S. school were solicited in person. The pilots ranged in age from 18-79 (M = 42.9, SD = 19.69). Aeronautical certifications held comprised Private Pilot (47), Commercial Pilot (64), Airline Transport Pilot (77), military only (1), and both military and civilian certifications (12). Initial flight training sources varied from 14 CFR Part 61 operations (93), 14 CFR Part 141 schools (80), military (25), and non-US flight schools (3). Current affiliation spanned flight schools (27), university/academia (59), airline (47), corporate (23), general aviation (41), and military (4). Delphi team members were not used in Stage 3.

Materials and stimuli. Participants were asked to consider the following scenario: “Please try to remember the last time you flew an airplane. Think about the air traffic controller who worked you in the pattern and cleared you to land. You may not have met the air traffic controller personally, but you know how the flight went for you. Please respond to the following statements regarding the air traffic controller.” After reading the scenario, the participants were given the
questionnaire containing the six statements, as shown in Appendix A, using a 5-point Likert-type measurement ranging from strongly disagree to strongly agree.

**Scale development.** The data were subjected to Principal Components Analysis (PCA) with a varimax rotation. The correlation matrix showed all variables had a correlation coefficient > 0.3, indicating a linear relationship between variables as required for reliable PCA results (Laerd Statistics, 2015). Kaiser-Meyer-Olkin (KMO) was 0.902 overall, indicating a “meritorious” classification (Kaiser, 1974), suggesting that the data were suited for factor analysis. Individual KMO measures were all >.88, indicating that each individual variable was suited for factor analysis. Bartlett’s Test of Sphericity was statistically significant ($p < .001$), indicating the data were factorizable. The PCA indicated a single component with an eigenvalue > 1.00 explaining 69.06% of the total variance. Scree plot analysis of the inflection point (see Figure 1) confirmed the appropriateness of the one component solution (Cattell, 1966).

![Scree Plot](image)

**Figure 1.** Scree plot used to visually suggest the number of factors to retain.

Two additional tests were conducted to assess validity and reliability. A Cronbach’s alpha of 0.90 showed the scale to have high internal consistency, and a Guttman split-half test result of 0.91 indicated the scale had high reliability (Field, 2018).
Discussion

The purpose of this study was to develop a valid instrument to measure a pilot’s trust in air traffic controllers. While trust has been studied in many domains, there is scant research regarding pilots and air traffic controllers. Further, current instruments to measure trust are domain specific and do not appear to be valid for measuring trust between pilots and controllers. Summative scales have the ability to measure latent concepts that cannot be measured directly. The benefit of an instrument developed by the population of interest (pilots) is it improves the probability the measurements will provide useful information for researchers. A concise scale was developed using pilots as subject matter experts and study participants.

The words used in the scale represent domain specific antecedents of trust, and using pilots to produce the words representing trust was a vital component of this project. The Delphi experts were not given any previous literature to base their contributions ensuring ownership of the concepts they believed represented trust in controllers. That the words used in the T-ATC scale loaded strongly on one factor is likely tied directly to the fact that pilots developed the scale and pilots participated in validating the scale.

The current study differs slightly from previous studies in that the antecedents of trust are stated as adjectives rather than nouns. The T-ATC scale used the terms attentive, confident, consistent, knowledgeable, professional, and direct. Strikingly, in their noun forms these terms are similar to those found in previous studies, yet they differ in the combination or in the variations of concepts supporting the argument that the terms need to be tailored to the population. For example, Butler (1991) used consistency, but also used integrity, competence, and fairness among others. Cook and Wall (1980) used ability and trustworthy intentions. Good (2000) also used ability and intention. Johnson-George and Swap (1982) used reliability. Finally, Mishra (1996) used competence, openness, caring, and reliability. Despite the differences, the commonality of the six T-ATC terms with other terms from other studies provides support to the conclusion that the one factor-six statement construct in the T-ATC scale do represent a pilot’s trust in air traffic controllers.

Limitations and Recommendations

There are a few limitations inherent in this study. First, the scale is U.S.-based. That is, the words were generated by U.S. pilots and the data gathered through a U.S sample. It is possible different results would be obtained from international pilots (for word generation) and an international sample (for scale validation). Second, the project did not consider flight hours, which has been shown in various studies to correlate with different outcomes (Ison, 2014). However, this is somewhat mitigated by analyzing pilot certification level, which
requires specific training in addition to a certain amount of experience. Thirdly, the project did not consider how recent the pilot had flown, possibly complicating the results due to memory issues. While the results indicate the scale is valid, the value of the scale will be improved through replication. Specifically, a discrimination analysis through an experiment can provide an indication of the strength of the scale.

**Conclusions**

Safety within the National Airspace System requires a high degree of cooperation between system actors facilitated through trust in system components. Advanced automation within the system assists, however, it is with the human element where most errors occur. The components of interest in the current study are the pilots and the air traffic controller who must work in concert to achieve high levels of safety and efficiency. The purpose of this study was to create a deliberate scale to assess a pilot’s trust in air traffic controllers where one did not exist previously. In a system reliant upon trust, there is little knowledge of trust between pilots and controllers and therefore, no foundation for understanding and addressing trust dynamics. The benefit of this project is in its contribution to researchers who rely on validated instruments with which to conduct scholarly research. The validated T-ATC scale can now be used by aviation researchers to close the gap in knowledge regarding a pilots trust in controllers.
References


Appendix A

Trust in Air Traffic Controllers (T-ATC) Scale
Please respond to the following statements according to how strongly you disagree or agree with the following statements.

1. The controller was attentive.
   Strongly Disagree  Disagree  Neutral  Agree  Strongly Agree

2. The controller was confident.
   Strongly Disagree  Disagree  Neutral  Agree  Strongly Agree

3. The controller was consistent.
   Strongly Disagree  Disagree  Neutral  Agree  Strongly Agree

4. The controller was knowledgeable.
   Strongly Disagree  Disagree  Neutral  Agree  Strongly Agree

5. The controller was professional.
   Strongly Disagree  Disagree  Neutral  Agree  Strongly Agree

6. The controller was direct.
   Strongly Disagree  Disagree  Neutral  Agree  Strongly Agree