

Publications

2015

A Trustworthiness of Commercial Airline Pilots (T-CAP) Scale for American Consumers

Stephen C. Rice
Florida Institute of Technology

Rian Mehta
Florida Institute of Technology

Scott Winter
Florida Institute of Technology, winte25e@erau.edu

Korhan Oyman
Florida Institute of Technology

Follow this and additional works at: <https://commons.erau.edu/publication>

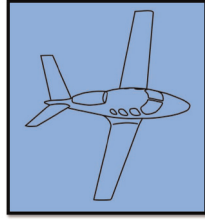


Part of the [Aviation Commons](#), and the [Social and Behavioral Sciences Commons](#)

Scholarly Commons Citation

Rice, S. C., Mehta, R., Winter, S., & Oyman, K. (2015). A Trustworthiness of Commercial Airline Pilots (T-CAP) Scale for American Consumers. *Journal of Aviation Technology and Engineering*, 4(2).
<https://doi.org/10.7771/2159-6670.1108>

This Article is brought to you for free and open access by Scholarly Commons. It has been accepted for inclusion in Publications by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu.



JATE

Journal of Aviation Technology and Engineering 4:2 (2015) 55–63

A Trustworthiness of Commercial Airline Pilots (T-CAP) Scale for American Consumers

Stephen C. Rice, Rian Mehta, Scott Winter, and Korhan Oyman

Florida Institute of Technology

Abstract

The purpose of this study was to create a Trustworthiness of Commercial Airline Pilots (T-CAP) scale that could be used with American participants. Previous research (Rice, Mehta, Steelman, & Winter, 2014) created a similar scale that may be used with Indian participants. However, due to cultural differences, it was necessary to recreate an instrument that could be used with American consumers. In fact, the scale developed by American participants did differ significantly, both in terms of length and items to measure trustworthiness of commercial pilots. Participants were used in the entire process of creating the scale, including item generation and testing. A factor analysis using principle components and a varimax rotation produced a single factor for the condition of trustworthiness. Cronbach's Alpha and Guttman split-half tests verified the reliability of the instrument and an experimental scenario was used to test the discriminability of the scale.

Keywords: pilot, trustworthiness, scale, valid, reliable, American

About the Authors

Stephen C. Rice is an associate professor at Florida Institute of Technology. He received his PhD in Human Factors from the University of Illinois at Urbana-Champaign in 2006. His research interests include aviation psychology, automation, trust, stigmas, and human performance.

Rian Mehta is a graduate student at the Florida Institute of Technology pursuing a Master's degree in Aviation Safety, while continuing on to complete his PhD in Aviation Science. He received his bachelor's degree in Aviation Management with Flight from the Florida Institute of Technology in 2013. His research interests include automation, pilot fatigue, and decision-making in the cockpit. Correspondence concerning this article should be sent to rmehta2009@my.fit.edu.

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

Korhan Oyman is a professor of aviation management and the dean of the College of Aeronautics at the Florida Institute of Technology. He has a master's degree in aviation safety from UCM in Missouri and a PhD in aviation management from Anadolu University in Turkey. Before academia he worked as Managing Director of SG International airport in Istanbul, Turkey. His research interests include airport and airline management, aviation safety, and flight training.

Introduction

With the affordable mainstream introduction of commercial flight, aviation has become the predominant mode of long-distance international travel for the past several decades. The viability of large distances being covered in relatively short time frames is the key behind the success of aviation. The economic benefits of this factor to passengers have become the main consideration when choosing their preferred mode of travel. However, beyond a certain breaking point in economic benefit, humans will make decisions based on their feelings of safety and comfort with a particular situation (Hughes, Rice, Trafimow, & Clayton, 2009). In the commercial aviation setting, a large aspect of the passengers' comfort and feelings of safety are tied into their perception of the airline, the aircraft, and, perhaps most importantly, the persons operating the flight.

Human perceptions on varying hypothetical scenarios have been studied by several research studies with an aim to gauge the passengers' willingness to fly on board different kinds of aircraft (Rice, Mehta, Steelman, & Winter, 2014). Technological advancements have far exceeded the current comfort levels of passengers. Therefore, the limiting factor is the consumers' trust in the flight's operation. The need for the aviation industry to have a valid and reliable scale to gauge consumer perception regarding the trustworthiness of a commercial pilot has become clearly evident. A previous study has endeavored to develop a trustworthiness scale of commercial airline pilots in India (Rice et al., 2014). This study seeks to expand this line of research in order to create a valid and reliable scale of consumers' perception of trustworthiness in commercial airline pilots using American participants. This will aid in creating a useful asset to the American aviation sector.

The purpose of these scales is to create a better understanding of the passengers' mental model of trust and trustworthiness. Their willingness to fly is the key to unlocking the potential of commercial air travel. This scale's applicability will be tailored specifically to American consumers. It is plausible to create an industry-wide scale for all commercial pilots from all countries, but these studies seek to create separate scales for different nations to minimize generalizations. Additionally, this separation will be able to account for factors that could possibly alter the scale, such as cultural differences, which are known to play a factor in a person's trust.

The study was conducted in a stepwise staged manner as described in the methodology, similar to the previous study. To start, consumers were used as participants in a survey to generate a master list of terms that relate to trust and trustworthiness. From that point, the list was narrowed to accurately represent key characteristics and traits that consumers believed to be a model for a trustworthy pilot. Once the list was sufficiently refined, the latter part of the study sought to establish validity, reliability, and discriminability by testing the newly developed scale.

Cultural Considerations

In an industry such as aviation, an important caveat to remember is that even though commercial airline travel has brought people from different nations together, there are distinct cultural differences seen across the wide spectrum of passengers. Cultural backgrounds and upbringings have the ability to influence the mental model of a consumer, and therefore lay varying levels of emphasis on different areas of importance when dealing with different scenarios. This is the influencing factor behind the need for the creation of separate scales for different nations within the aviation industry.

The basis of conducting research with two different nationalities is to further understand the differences in culture in the aviation industry. Culture could be defined in several different broad ways. The expansive statement put forth by Helmreich (2000) defines culture as norms, values, and practices on a national, organizational, and/or professional level that are shared within a society. Since the purpose of these studies is to generate scales for different cultures, this definition, though fairly broad, is the most apt. It can be applied to differentiate the two countries with respect to their norms, values, and practices as related to their opinions toward aviation.

One crucial aspect of analysis between the two cultures is to explore the differences of a collectivistic society versus an individualistic society as they relate to their trust in commercial pilots, and, in turn, their trust in aviation. Markus and Kitayama (1991) state that India is generally considered to be a collectivistic society, wherein the citizens are interdependent on one another. On the other hand, a society which in large part creates an atmosphere where people are more independent of their fellow citizens can be termed as an individualistic society. The United States is predominantly considered to be individualistic in nature since the country scored the highest with 98 out of 100 on Hofstede's (1980) cultural value by nation index.

The link between the psychological aspects of trust in relation to the differences in independencies of the two cultures is what gives weight and importance to the creation of two scales through these studies. Research studies have shown that participants raised in collectivistic cultures tend to be more trusting than their individualistic counterparts (Hofstede, 1980). When a person is raised in a collectivistic society, he or she is taught from an early age that the norm is to totally trust without question. This ties into other research that further explains these cultural tendencies, as studies have found that citizens of collectivistic societies display a higher concern for others' opinions in their own decision-making processes. These allocentric tendencies could be identified as potential influencing factors in the person's mental model of trust, and thereby alter his or her view on what are appropriate characteristics of trustworthy pilots.

Individuals that develop in an atmosphere of collectivism are taught not to question authority, which in turn leads to a heavier respect and trust for people in positions of authority. Collectivistic cultures, therefore, also teach one to be trustworthy, so that the fellow citizens of the community will be able to place their trust in the individual. Several other research studies have noted that there is a marked difference in levels of trust between extroverts and introverts. It has been stated that extroverts are more willing to trust others as compared to introverts (Gaines et al., 1997; Omodei & McLenna, 2000; Shikishima, Hiraishi, & Ando, 2006). These various studies show the effect of internal character traits and external societal influences on an individual's definition, level, and ability to trust.

Hofstede's research from 1980 inferred that citizens of individualistic cultures manifested traits to suggest that they were less likely to trust new people than those of collectivistic backgrounds. Hofstede's cultural value by nation index scores countries on a scale from 0 to 100 in an attempt to quantify these values and characteristics. Robbins and Judge (2009) showed that the United States scored the highest with 98 out of 100, while India scored a 48. This illustrated that, as compared to India, the United States was predominantly individualistic. Lastly, an area of research to be noted is that conducted by Rice and Kraemer (in press), which showed that, in several research cases, the views toward aviation and trust-related aviation scenarios were more extreme in American participants as compared to those of the Indian participants. All of this research points to the possibility for differential items in a trustworthiness scale.

Trust and Trustworthiness

Trust is crucial in any industry, especially between the consumer and the operator. It is even more important in a field such as aviation where consumers' lives are in the hands of the operator. Trust can be defined in numerous ways. It could be phrased differently to suit the context of the setting; however, at its core, trust is a psychological construct. For the purpose of this study, the research will utilize the most appropriate definition in social psychology, which states that trust is the predictability of another person (Deutsch, 1958; Eckel & Wilson, 2004; Ergeneli, Saglam, & Metin, 2007). Several prior studies have made an effort to differentiate the meanings of the two terms trust and trustworthiness, where trust is considered the construct of the truster, while trustworthiness is a construct of the trustee. On the other hand, this study will use the two terms interchangeably, due to the high correlation between the two (Rice & Kraemer, in press).

People's daily decisions and instincts are based on their interactions with other people, or plainly stated, on the amount of trust they have in the person with whom they are interacting. Using the previous definition as a base model for the purpose of this study, trust heavily influences

people's cognitive and emotional decision-making. A person's trust in someone else can aid one in predicting whether or not another individual will do what is expected of him or her. Previous Research aptly addressed the concept of trustworthiness, where they stated that to obtain a positive interaction between two individuals, one individual must believe that the other will do what is expected, thereby creating a high level of trustworthiness (Lee and Moray, 1994; Lee and See, 2004).

Trust as a psychological construct is known to be extremely volatile. This is even more so based largely on the fact that it is derived from people's emotions. The paradox that exists when discussing trust is that it is an extremely powerful construct or emotion; however, it is also extremely easy to break. Trust can be broken, lost, or affected by the slightest shift in circumstances or actions, and is easily affected. To overcome decreased or lost trust in someone or something, it can oftentimes prove to be a large task, and is virtually affected permanently (Lewick & Bunker, 1996). Research has found that even through massive effort, and extended periods of time, trust in a person or situation is nearly impossible to rebuild to its prior form or level. A study conducted by Slovic (1993) went to the furthest extreme to categorically state that once trust was lost or affected, it could never be regained no matter what. It is crucial to keep in mind the variability that exists between different people from different walks of life when dealing with trust and trustworthiness.

It is imperative that it is understood that the qualities of trustworthiness are not concrete or permanent. However, they do have lasting impacts on future decisions similar to the psychological phenomenon of anchoring or the "first impression" phenomenon. If a person proves to be trustworthy to an individual on their first encounter, he/she is more likely to be considered trustworthy in subsequent interactions. On the other hand, once primed by a negative interaction with an individual or particular situation, the individual is less likely to be trusting of a similar person or situation in the future. Therefore, it is important to note that all people have significantly different levels of trust, even in identical situations, and this is in large part due to their own personal experiences. The cultural differences between people of different nationalities were outlined earlier. Now, emphasis will be laid on the wide-ranging differences between the levels of trust and the constructs of trustworthiness among different cultures. This is a facet that was laid out in detail by Mechanic's research (1996) of trust and the differences in trust presented by cultural and sociological differences. Hassan and Semerciöz (2010) showed that an individual's trust was based on a variety of influencing factors such as individual personality traits, cultural characteristics, and past experiences.

As stated earlier, there is not just one definition of trust. It can be defined in several forms that can be applied to this research. The foundation of the emotional feeling of trust is

when someone can rely on someone else's word or rely upon the agreement with another (Rotter, 1967). In conclusion, the basis of trust lies in the mutual understanding between agreeing parties that the other will perform what is expected of them works (Barber, 1983; Rempel et al., 1985; Rotter, 1967).

The need for this research is urgent, as the foundation of the aviation industry is the consumer and thus, in turn, the consumers' perception of trust and trustworthiness in the industry, the airline, the aircraft, and lastly, the pilot. Trustworthiness is important in the consumer-based field of aviation, as the consumer is placing his or her life in the hands of the industry based on his or her level of trust in the industry. It is therefore important to gather the consumer perceptions of pilots, as these views are direct translations of their trust in the pilots, and, in turn, their trust and willingness to fly within the commercial aviation sector. This can help airlines better understand their consumers, and thus provide a service in which consumers have higher levels of trust. A definition of trust that is appropriate in this context is one that states that trust can be explained as an individual's willingness to be vulnerable to another (Meyer et al., 1995). This is a different definition than the one used previously, but is nonetheless relevant based on the context of the passenger placing trust in the pilot to safely conduct the flight, and taking caution for his or her life. The aviation industry can always benefit from understanding its passengers better. This research seeks to fill that gap in the scientific community.

Methodology

Stage 1: Word Generation

Stage 1 was designed to begin the item generation phase of the scale. The goal was to not only solicit items from aviation and trust experts, as well as prior scales of trust and trustworthiness, but also to solicit items from real-world consumers. Given that consumers are the ones who will be responding to the scale, the authors felt that this increased construct validity of the process.

Participants

Seventy (47 male, 23 female) participants from the United States were recruited via a convenience sample using Amazon's Mechanical Turk (MTurk). MTurk provides participants who complete human intelligence tasks in exchange for monetary compensation. Prior research shows that data from MTurk is as reliable as normal laboratory data (Buhrmester, Kwang, & Gosling, 2011; Germine, et al., 2012). The mean age was 30.81 ($SD=8.91$). Six additional participants with expertise in aviation provided further items for review. Lastly, the trust literature was reviewed (Jian, Bisantz & Drury, 2000), and items were added accordingly. The researchers used all necessary precautions

to ensure that participants from earlier stages did not participate in later stages.

Materials and Stimuli

Participants were presented with the following scenario: "Imagine a commercial airline pilot who is trustworthy. In the context of the commercial airline pilot mentioned above, please enter 5 characteristics of a trustworthy pilot in the spaces provided below. Each answer should include only one word or short phrase." After providing the list of 5 words or phrases, participants were debriefed and dismissed. This exercise generated 177 unique words or phrases. These items were then reviewed for correct spelling and decapitalized when necessary to ensure uniform saliency.

Stage 2: Nominal Paring

Stage 2 was designed to narrow down the initial list of items. In this stage, the researchers began eliminating words or phrases that were not perceived by a large majority of participants as being related to the construct of trustworthiness.

Participants

Forty-one (22 male, 19 female) participants from the United States were recruited via a convenience sample using Amazon's Mechanical Turk. The mean age was 33.17 ($SD=13.87$). The researchers used all necessary precautions to ensure that participants from earlier stages did not participate in later stages.

Materials and Stimuli

Each of the 177 items generated in the first stage were presented to participants, along with the following statement, "In the context of a commercial airline pilot, please rate whether each word below is related to (similar to) pilot trustworthiness, not related to (not similar to) pilot trustworthiness, or you don't know." Thirty-five items were chosen to be related to trustworthiness by at least 85% of participants.

Stage 3: Likert-Scale Paring

Stage 3 was designed to continue narrowing down the list of items to items that would be retained for the final scale. Since the researchers were seeking a more sensitive measure of the relationship between the items and trustworthiness, they used a Likert-type scale instead of a nominal scale.

Participants

Forty-four (24 male, 20 female) participants from the United States were recruited via a convenience sample using Amazon's Mechanical Turk. The mean age was 32.55 ($SD=9.87$). The researchers used all necessary precautions to ensure that participants from earlier stages did not participate in later stages.

Materials and Stimuli

The 35 items retained from Stage 2 were presented to participants with the following statement, “*In the context of a commercial airline pilot, please rate how strongly each word below is similar to trustworthiness.*” Participants responded based on a Likert-type scale from “Not at all similar to Trustworthiness” (0) to “Extremely Similar to Trustworthiness” (+3). Five items averaged 2.5 or higher and were retained for the final scale.

Stage 4: Scenario-Based Testing

Stage 4 was designed in order to begin collecting validity and reliability evidence for the newly created scale. The five items were crafted into statements that could be rated on a Likert-type agreement scale.

Participants

Two hundred and thirteen (147 male, 66 female) participants from the United States were recruited via a convenience sample using Amazon’s Mechanical Turk. The mean age was 28.22 ($SD=7.68$). The researchers used all necessary precautions to ensure that participants from earlier stages did not participate in later stages.

Materials and Stimuli

In this stage, participants were presented with the following scenario: “*PanAm Flight 230 took off from NYC at 9:00 AM en route to Paris, France. About 40% of the way into the flight, a bird strike occurred, whereby one of the engines was slightly damaged by the impact of a bird. The pilot contacted the airline’s maintenance department via company radio to report the incident and was told that it was a judgment call on the part of the pilot whether to divert to a nearby airport or continue on to Paris and land there. The pilot decided to continue on. Please respond how strongly you disagree or agree with the following statements about the pilot.*” Participants were presented with the questionnaire (see Appendix A) and asked to provide statements of agreement or disagreement on a 5-point Likert-type scale (coded from -2 to $+2$).

Scale Development

A factor analysis using the principle components and varimax rotation resulted in all items strongly loading on one factor. A Cronbach’s Alpha test was conducted to measure internal consistency within the scale. The resulting coefficient of 0.93 indicated high internal consistency. A Guttman split-half test was conducted as well. The resulting coefficient of 0.90 indicated high reliability.

Stage 5: Scenario-Based Experiment

The following experiment was conducted in order to accomplish three goals. The first goal was to replicate the

factor analysis and reliability analyses using a specific scenario similar to the kind that researchers might actually use in aviation consumer research (Rice et al., 2014). The second goal was to test the ability of the scale to discriminate between pilots who might be perceived as trustworthy or untrustworthy. Our third goal was to test how well the scale correlates with “willingness to fly”, which is a common outcome variable used in aviation consumer research (Rice et al., 2014).

Participants

Two hundred and five (120 male, 85 female) participants from the United States were recruited via a convenience sample using Amazon’s MTurk and randomly assigned to one of the two groups described below. The mean age was 30.36 ($SD=10.50$).

Materials and Stimuli

Participants were presented with the following scenario: “*Imagine a situation where you are on a commercial airline flight from one major city to another. The pilot of the airplane is known by his friends, family and colleagues to often be dishonest about his personal affairs, and sometimes cuts corners in his work performance.*” A second scenario was presented to another group of participants where the pilot was described as being honest and never cutting corners. Participants were then asked to respond to the Trustworthiness Scale (see Appendix A), and to statements of their willingness to fly (see Appendix B) in these particular scenarios.

Scale Development

For the trustworthy condition, a factor analysis using the principle components and varimax rotation resulted in all items loading strongly on one factor. Prior to performing the t-test, measurement equivalence was tested and established. A Cronbach’s Alpha test was conducted to measure internal consistency within the scale. The resulting coefficient of 0.96 indicated high internal consistency. A Guttman split-half test was conducted as well. The resulting coefficient of 0.91 indicated high reliability. The correlation between the trustworthiness scale and willingness scale was $r(100)=0.55$, $p<0.001$, revealing a significant, and fairly strong, positive correlation between the two variables.

For the untrustworthy condition, a factor analysis using the principle components and varimax rotation resulted in all items loading strongly on one factor. A Cronbach’s Alpha test was conducted to measure internal consistency within the scale. The resulting coefficient of 0.90 indicated high internal consistency. A Guttman split-half test was conducted as well. The resulting coefficient of 0.86 indicated high reliability. The correlation between the trustworthiness scale and willingness scale was $r(101)=0.60$, $p<0.001$, revealing a significant, and fairly strong, positive correlation between the two variables.

There was a significant difference in scores on the trustworthiness scale between the two conditions, $t(203)=28.92$, $p<0.001$, $d=4.06$, indicating that the scale was easily able to discriminate effectively between the trustworthy ($M=1.45$, $SD=0.66$) and untrustworthy ($M=-1.11$, $SD=0.61$) conditions, with notably small standard deviations for each group.

Discussion

The purpose of this study was to produce a valid and reliable instrument that would be able to discriminate among varying levels of consumer trust in his or her pilot using a sample of participants from the United States. A gap was present in the literature for an instrument that could be used with American participants. Previous research (Rice et al., 2014) created a valid and reliable instrument that could be used to measure consumer trust of pilots with samples from India; however, due to cultural differences between India and the U.S. (Markus & Kitayama, 1991), it would not be good practice to use that scale with American participants. The two scales are in fact different from one another. Therefore, the authors believe that there is a need within the scientific community for a scale that can be used by researchers wishing to complete trust studies using participants from the United States.

In the scale development, the authors structured the instrument to only use positive words. Prior research completed by Harrison and McLaughlin (1991) stated that reverse-scored items could have detrimental psychometric effects on the measure. Because of this, in real-world situations, it is deemed preferable to have items scaled in one direction (Hinkin, 1998). It is also anticipated that this will increase the effectiveness of the scale by reducing the cognitive load.

Validity

Validity of an instrument refers to the instrument measuring what it is intended to measure (Field, 2009). To ensure validity, a two-step process was utilized. First, American consumers (the population for which the scale was developed) were used to generate words associated with the trustworthiness of pilots. The strength in using the intended audience of the scale was that these individuals are deemed most appropriate to supply words. Experts may not identify the same trustworthiness attributes as actual consumers. By incorporating these participants into the word generation and paring phases, the final items produced should relate more closely to consumers' views on the trustworthiness of pilots. Second, a factor analysis was completed on the data using the principle components and varimax rotation. Results indicate that the final five items represented a single factor (trustworthiness). Based on these two steps, evidence for construct validity has been

established for the scale, and it provides a valid measure of American consumers' trustworthiness ratings of pilots.

Reliability

The reliability of an instrument refers to how consistently it measures what it is intended to measure (Field, 2009). The final five items on the scale were subjected to a Cronbach's Alpha analysis and values ranged from 0.90–0.96. Additionally, a Guttman split-half analysis was completed and values ranged from 0.86–0.91. Therefore, the data suggests that the final scale items have very high internal consistency and reliability.

Discriminability

Aside from being valid and reliable, a characteristic of a good scale is its ability to discriminate. For the current scale to be effective in the real-world, it must be able to discriminate between conditions in which a pilot is being trustworthy and untrustworthy. The final stage of the scale development provided for two scenarios: a trustworthy and untrustworthy condition. The scale successfully demonstrated its versatility in discrimination with notably small standard deviations in each group. Therefore, it should be a useful tool for the research community on measuring the trustworthiness of pilots.

Differences between Scales

The current study provides an example as to why it is important to examine the validity and reliability of scales across different populations. The scale produced by American participants did differ significantly from the scale that was previously created with Indian participants (Rice et al., 2014). While some items were the same (reliability and trustworthiness of the pilot), other items were different. The scale with Indian participants focused partially on items such as qualified, experienced, active, and talented. The scale developed by American participants addressed items such as dependable, safe, and responsible. Since the participants in each study were used to generate the words used in the scale, it seems that Indian participants focused more toward the skill-based attributes of the pilot (experience, qualifications, active, etc.), while American participants focused on the responsibility of the pilot (responsible, dependable, safe, etc.). Additionally, the scale developed with Indian participants consisted of 7 items while the scale developed by Americans contained 5 items.

Clearly there are some differing views toward the trustworthiness of pilots based upon populations. These differences justify the current study, and the creation of a new instrument to be used with American participants. If additional populations were desired to be studied,

it would be best to recreate an instrument to be used with those populations.

Practical Implications and Limitations

This scale provides for some practical implications. The purpose of developing the scale was to provide the research community with a tool that may be used to measure the relationship between the passenger and pilot based on trustworthiness for American participants. There had been a gap identified within the literature as to a valid and reliable instrument to measure the construct of trustworthiness of consumers toward their pilot. This scale can be used by researchers in many different studies that seek ratings of consumer trustworthiness of pilots. These perceptions may provide valuable insights that could be beneficial in business, marketing, or academic disciplines.

All research is bound by certain limitations—the present study is no different. This scale was designed to be used by American participants. Use of the scale outside of this population may yield invalid results. As with prior research on the development of a trustworthiness scale (Rice et al., 2014), different scales were produced for participants from India and the United States. It is possible that this would also occur if the scale were modified for other nationalities. Future research could adapt the scale to be used with additional populations. Additionally, future research should be completed to verify the findings of this study and verification of the scale.

MTurk was used for data collection. While shown to be as effective as laboratory data (Buhrmester et al., 2011; Germine, et al., 2012), MTurk is not without its own limitations. It may be possible that this population does not fully represent the views of American participants. Finally, this scale was designed to be used by consumers. If experts or those with backgrounds in aviation were to use the scale, the findings may be in question, as these persons may have different views or interpretations of the trustworthiness of a pilot.

Conclusions

The purpose of this study was to create a valid and reliable scale that could be used to discriminate among passengers' trustworthiness of a pilot using American participants. The study built on an earlier scale that was created to measure trustworthiness of pilots with Indian participants (Rice et al., 2014). Due to cultural differences, the current study was justified, as it was expected that the two scales would be slightly different; in fact, this result was found to be the case.

American consumers assisted in the entire process of scale development from word generation through discriminability tests. The result of the study was a valid and reliable scale that successfully discriminates between trustworthy and untrustworthy scenarios. The scale should

help researchers gather the true perceptions of consumers toward the trustworthiness of their pilots. This can help in practical situations as researchers work to understand passengers better and improve customer experiences.

References

- Barber, B. (1983). *The logic and limits of trust*. New Brunswick, NJ: Rutgers University Press.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science*, 6(3), 3–5. <http://dx.doi.org/10.1177/1745691610393980>
- Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution*, 2(4), 265–279. <http://dx.doi.org/10.1177/002200275800200401>
- Eckel, C. C., & Wilson, R. K. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization*, 55(4), 447–465.
- Ergeneli, A., Saglam, G., & Metin, S. (2007). Psychological empowerment and its relationship to trust in immediate managers. *Journal of Business Research*, 60(1), 41–49. <http://dx.doi.org/10.1016/j.jbusres.2006.09.012>
- Field, A. (2009). *Discovering statistics using SPSS*. London, UK: Sage.
- Gaines, S. O., Panter, A. T., Lyde, M. D., Steers, W. N., Rusbult, C. E., Cox, C. L., & Wexley, M. O. (1997). Evaluating the complexity of interpersonal traits and the manifestation of interpersonal traits in interpersonal trust. *Journal of Personality & Social Psychology*, 73, 610–623.
- Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the web as good as the lab? Comparable performance from web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin & Review*, 19(5), 847–857. <http://dx.doi.org/10.3758/s13423-012-0296-9>
- Harrison, D. A., & McLaughlin, M. E. (1991). Exploring the cognitive processes underlying responses to self-report instruments: Effects of item content on work attitude measures. In *Proceedings of the 1991 Academy of Management Annual Meetings* (pp. 310–314) unpublished work.
- Hassan, M., & Semerciöz, F. (2010). Trust In Personal And Impersonal Forms Its Antecedents And Consequences: A Conceptual Analysis Within Organizational Context. *International Journal of Management & Information Systems*, 14(2), 67–84.
- Helmreich, R. L. (2000). Culture and error in space: Implications from analog environments. *Aviation, Space, and Environmental Medicine*, 71(9–11), 133–139.
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1(1), 104–121. <http://dx.doi.org/10.1177/109442819800100106>
- Hughes, J., Rice, S., Trafimow, D., & Clayton, K. (2009). The automated cockpit: A comparison of attitudes towards human and automated pilots. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(5), 428–439.
- Hofstede, G. (1980). Motivation, leadership and organization: Do American theories apply abroad? *Organizational Dynamics*, 9(1), 42–63.
- Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4(1), 53–71. http://dx.doi.org/10.1207/S15327566IJCE0401_04
- Lee, J., & Moray, N. (1994). Trust, control strategies and operators' adaptation to automation. *Ergonomics*, 35, 1243–1270.
- Lee, J. D., & See, A. K. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80. http://dx.doi.org/10.1518/hfes.46.1.50_30392
- Lewick, R., & Bunker, B. B. (1996). Developing and maintaining trust in work relationships. *Trust in organizations: Frontiers of theory and reach* (pp. 114–139). New York, NY: Sage.

- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological review*, 98(2), 224–253. <http://dx.doi.org/10.1037/0033-295X.98.2.224>
- Mechanic, D. (1996). The logic and limits of trust. *Contemporary Sociology*, 455, from, http://serials.unibo.it/cgi-ser/start/en/spogli/df-s.tcl?prog_art=1851483&language=ENGLISH&view=articoli
- Meyer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734.
- Omodei, M. M., & McLennan, J. (2000). Conceptualizing and measuring global interpersonal mistrust-trust. *The Journal of Social Psychology*, 140(3), 279–294. <http://dx.doi.org/10.1080/00224540009600471>
- Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, 49(1), 95–112. <http://dx.doi.org/doi/10.1037/0022-3514.49.1.95>
- Rice, S., & Kraemer, K. (in press). Taking the human pilot out of the cockpit: Passengers reject idea. *Proceedings of the 10th International Conference on Technology, Knowledge and Society*, February 23–24, Berkley, CA.
- Rice, S., Mehta, R., Steelman, L., & Winter, S. R. (2014). A trustworthiness of commercial airline pilots (T-CAP) scale for Indian consumers. *International Journal of Aviation, Aeronautics, and Aerospace*, 1(3), 1–15. Retrieved from <http://commons.erau.edu/ijaaa/vol1/iss3/3>
- Robbins, S. P., & Judge, T. A. (2009). *Organizational behavior (13th ed.)*. Upper Saddle River, NJ: Prentice Hall.
- Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, 35(4), 651–665. <http://dx.doi.org/10.1111/j.1467-6494.1967.tb01454.x>
- Shikishima, C., Hiraishi, K., & Ando, J. (2006). Genetic and environmental influences on general trust: A test of a theory of trust with behavioral genetic and evolutionary psychology approaches. *Japanese Journal of Social Psychology*, 22, 48–57.
- Slovic, P. (1993). Perceived Risk, Trust, and Democracy. *Risk Analysis*, 13, 675–682. <http://dx.doi.org/10.1111/j.1539-6924.1993.tb01329.x>

Appendix A**Trustworthiness Scale**

Please respond how strongly you agree or disagree with the following statements.

1. The pilot is dependable.
Strongly Disagree Disagree Neutral Agree Strongly Agree
2. The pilot is reliable.
Strongly Disagree Disagree Neutral Agree Strongly Agree
3. The pilot is responsible.
Strongly Disagree Disagree Neutral Agree Strongly Agree
4. The pilot is safe.
Strongly Disagree Disagree Neutral Agree Strongly Agree
5. The pilot is trustworthy.
Strongly Disagree Disagree Neutral Agree Strongly Agree

Appendix B**Willingness to Fly Scale**

Please respond how strongly you agree or disagree with the following statements.

1. I would be willing to fly in this situation.
Strongly Disagree Disagree Neutral Agree Strongly Agree
2. I would be comfortable flying in this situation.
Strongly Disagree Disagree Neutral Agree Strongly Agree
3. I would have no problem flying in this situation.
Strongly Disagree Disagree Neutral Agree Strongly Agree
4. I would be happy to fly in this situation.
Strongly Disagree Disagree Neutral Agree Strongly Agree