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

Stephen Rice  
*Florida Institute of Technology, srice@fit.edu*

Rian Mehta  
*Florida Institute of Technology, rmehta2009@my.fit.edu*

Lisa A. Steelman  
*lsteelma@fit.edu*

Scott R. Winter  
*Florida Institute of Technology, scott.winter@mac.com*

Follow this and additional works at: [http://commons.erau.edu/ijaaa](http://commons.erau.edu/ijaaa)
For the past few decades, humans have, for the most part, accepted air travel as the dominant, practical and most viable mode of long distant transportation. Many research studies have attempted to gauge the consumers’ or passengers’ willingness to fly on board different kinds of aircrafts, with different hypothetical scenarios (e.g. Rice, et al., 2014). The perception of the passenger is critical in the aviation industry, as they are the end consumer, and must be willing to use the airlines’ product. The industry’s primary concern is the passengers’ willingness to fly on board the aircraft, and therefore a better understanding of the consumers’ mental model of willingness to fly is key in many respects. The passengers’ willingness to fly on board an aircraft is most certainly going to be affected by their perception of the pilot, and their overall trust in the pilot to operate the flight safely.

This study seeks to create a scale of trustworthiness in a commercial pilot that could be applied in the field of aviation within India. This particular study will only deal with developing a scale for Indian consumers, in large part due to the heavy surge in commercial aviation in India. Recent research showed that approximately 50 million passengers fly in and out of India on an annual basis (Carrerio, n.d.). The scale is being constructed specifically for one country and not over the industry as a whole in order to minimize generalizations, and to be able to account for cultural differences that may play a factor in passenger’s trust that could possibly alter the scale. The current study was conducted in stages as described in the methodology. Initially, a master list of related trust and trustworthiness terms were amassed using consumers as participants, and as the study progressed, this list was refined to identify key characteristics and traits that exemplify the consumers’ model of trustworthiness in a pilot. The latter portion of the study deals with testing the newly developed scale in order to establish validity, reliability and discriminability.

**Trust and Trustworthiness**

Trust is best described as a psychological construct, and can be scientifically defined in several different ways. In the context of this research, the most apt definition in social psychology states that trust is the predictability of another person (Deutsch, 1958; Eckel & Wilson, 2004; Ergeneli, Saglam, & Metin, 2007). It is important to differentiate that trust is a construct of the truster, while trustworthiness is a construct of the trustee; however, for the purposes of this literature review, we use the terms interchangeably given that they are so highly correlated (Rice, Richardson & Kraemer, 2014).

People have countless interactions with each other on a daily basis, and trustworthiness can have a significant role in how those interactions play out. From the previous definition, we can deduce that trust allows someone to predict whether another individual will do what is expected of him or her. If a person has a high level of trustworthiness, it will usually result in a positive interaction, because one individual believes that the other will do what is expected (Lee & See, 2004). Rotter (1967) defined trust in a slightly different yet relatable manner. He stated that the expectation of someone’s word or an agreement with an individual could be relied upon was the basis of the attitude of trust. In summation, several works (Barber, 1983; Rampel et al., 1985; Rotter, 1967) summarize trustworthiness as a perception of expectancy dependent on the performance of the other individual, the chances of certain events occurring, and lastly, responsibility, monetary or otherwise.
Mechanic (1996) said that trust functions in numerous ways due to its cultural and sociological characteristics. Trustworthiness, being subjective to a person’s feelings, can be a volatile function. Research has shown that though trust is an extremely powerful psychological occurrence, it can be easily affected. Once trust in someone or something is influenced or lost, it can be virtually impossible to completely overcome. Even over extend time frames, trust is sometimes impossible to be completely rebuilt to its previous level. Slovic (1993) went further to state that in certain situations, once trust is lost, it might never be regained. An important aspect of trustworthiness that should be well noted is that it is relative, and varies significantly from person to person. One’s trustworthiness is based on a variety of different factors such as individual personality traits, cultural characteristics, and most of all past personal experiences (Hassan & Semerciöz, 2010).

Different people from different walks of life are going to have varying levels and values associated with trust. A person, who has had negative interactions in a specific scenario, is less likely to be trusted in a similar situation in the future. Conversely, a person that has been trustworthy in the past is more likely to be considered trustworthy in the future. In the same manner, people may have varying views on what traits they believe are required to be considered trustworthy. Additionally, cultural upbringing and societal characteristics may be influencing factors in the development of an individual’s model of trust. The predominant differences are seen when comparing citizens of collectivistic societies, like those of India, to individualistic societies as seen primarily in the United States. People that are raised in collectivistic cultures are brought up from their infancy to be more interdependent on one another; are taught to always keep the community’s best interest above all else; discouraged from questioning authority; and are taught to totally trust without question (Wu & Jang, 2008). Collectivistic cultures therefore also teach one to be trustworthy, so that the fellow citizens of the community will be able to lay their trust in the individual. On the other hand, a person that is being enculturated in an individualistic community is primed to be more self-focused and trusting of one’s self over others. An emphasis is placed on being wary of being too trusting, and of trusting other individuals. 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 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 of trustworthiness.

As mentioned earlier, trustworthiness is a psychological construct that varies from individual to individual. A study conducted in the Cuddalore district of Tamil Nadu, India, concluded that perceptions of trustworthiness are based not just on the facts and analyses that are said to constitute information, but more on the context within which information is accessed (Srinivasan, 2007). A study of consumer perceptions of trustworthiness from India showed integrity/honesty, communication/similarity, shared values, expertise, and ability/consistency as significant predictors of overall trustworthiness (Roy, Eshghi, & Shekhar, 2011).

The aviation industry is a very consumer based market, and the traits of trustworthiness are relatively important in any consumer-oriented field. Sethi and Allen (1984) stated that the traits of ability, interpersonal warmth, trustworthiness, interpersonal strength, motivation, and
family orientation are highly desirable in Indian society. In turn, these are traits highly desirable in the aviation marketplace, especially when considering the traits desirable in a pilot. In many aviation related research settings, it may be desirable to gather consumer perceptions on areas related to trust in pilots. Consumer’s views toward the pilot of their commercial aircraft may significantly influence the trust that they place in their pilot. Trust has been defined earlier, but Meyer et al. (1995) stated that trust could be termed as a willingness to be vulnerable to another. This is relevant in the aviation context of a passenger relying on a pilot and being vulnerable to their actions during control of the flight.

Previous Trust Scales

Multiple scales have examined trust related to people. In 1964, the University of Michigan developed a three-item trust scale in conjunction with the 1964 election. The instrument was designed to measure post-election levels of trust and has become a commonly used instrument in national surveys since its development. In 1986, Yamagishi developed a five-item questionnaire for examining the level of trust towards another person. In 1994, Yamagishi and Yamagishi built a six-item questionnaire on general trust to measure participant beliefs regarding the honesty and trustworthiness of others. Items examined included how trusting of others, how honest, kind, and trustworthy persons were as measured on a Likert scale. Rempel, Holmes, and Zanna (1985) created a scale on trust in close relationships. This 17-item measure had three sub-sections on predictability, dependability, and faith. The purpose of this scale was to determine the level of trust one relationship partner has in the other.

Current Study

Lacking within the literature is an empirically developed scale that may be used to measure pilot trustworthiness, a topic that has received much attention in recent years. The recent disappearance of a Malaysian Airlines Boeing 777 demonstrated scrutiny of the flight crew as explanations for the disappearance were sought. Additionally, events such as the alleged mental breakdown of a JetBlue airline pilot on a flight from New York to Las Vegas in March 2012 (Hunter & Patterson, 2012) or numerous examples of commercial airline pilots being removed from the cockpit before flight due to alcohol related issues are all issues that could influence consumer perceptions of pilot trustworthiness. While these events are clearly a minority of cases, it does highlight the need to have a valid measure that could be available to the research community to help provide accurate ratings of consumer’s trust in their pilot.

In the following sections, we outline how we developed this trustworthiness scale. The initial part of the study involved developing the terms and items by which trust in airline pilots could be measured. This is completed by means of the first three studies in the line of research. The latter half of the project involved two additional studies to test and validate the developed measure of trust. The methodology section shows a step-by-step process by which we arrived at a concise, valid and reliable scale of pilot trustworthiness that can be used with Indian consumers.
Methodology

Stage 1: Word Generation

The purpose of Study 1 was to begin the word generation phase of the scale. While solely soliciting items from experts in the field has been a method used to develop some scales, our goal was to also solicit items from actual consumers, given that it will be consumers themselves who will respond to the scale items upon completion. We believe that this helps to increase construct validity in the process. Thus, in this first stage, we solicited items from potential consumers, experts, and other related scales of trust and trustworthiness.

Participants

Seventy-two (25 females) participants from India participated in the first part of the study. The mean age was 30.81 (SD = 8.79). Participants were recruited via a convenience sample using Amazon’s Mechanical Turk (MTurk). MTurk provides an online source of participants that are willing to complete human intelligence tasks in exchange for a small amount of compensation. Previous research has shown that data from MTurk is as reliable as normal laboratory data (Buhrmester, Kwang, & Gosling, 2011; Germine, et al., 2012). All online participants identified as being airline consumers who spoke English as a native language. This was also the case for each following study.

An additional 10 participants with expertise in aviation were recruited from the [blinded for review] community. Lastly, the literature was reviewed from related scales of trust or trustworthiness (e.g. Jian, Bisantz & Drury, 2000), and words were added accordingly.

Materials and Stimuli

Participants first gave electronic consent via FluidSurveys and were then 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 phrase.” Once participants provided the list of 5 words or phrases, they were debriefed and dismissed.

There were a total of 172 unique words or phrases generated from this exercise (e.g. efficient, experienced, hard-working, etc.). These words were then reviewed for correct spelling and all words were de-capitalized so they would have equal saliency in the following steps.

Stage 2: Nominal Paring

The purpose of study 2 was to narrow down the initial list of items by eliminating words or phrases that were not perceived by participants as being related to the construct of trustworthiness.
Participants

Sixty-two (29 females) participants from India participated in the study. The mean age was 32.10 (SD = 9.69). Participants were recruited via a convenience sample using Amazon’s® Mechanical Turk® (MTurk).

Materials and Stimuli

In this stage, the 172 words generated in the first stage were presented to participants one at a time 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.” Words that were found to be related to trustworthiness by at least 85% of participants were included in the next stage. This resulted in 30 words being included in the next stage.

Stage 3: Likert-scale Paring

The purpose of study 3 was to continue narrowing down the list of items. However, given that all the items had already been determined to be related to trustworthiness by the vast majority of participants, in this stage, we were seeking a more sensitive measure of this relationship. Thus, we used a Likert-type scale to give us data on how related to trustworthiness each item was.

Participants

Forty-six (21 females) participants from India participated in the study. The mean age was 30.57 (SD = 8.84). Participants were recruited via a convenience sample using Amazon’s® Mechanical Turk® (MTurk).

Materials and Stimuli

In this stage, the 30 words left over 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 were asked to give a choice on a Likert-type scale from “Not at all similar to Trustworthiness” (0) to “Extremely Similar to Trustworthiness” (+3). An average score was determined for each item across participants and words that scored an average of 2.0 or higher were kept for the next stage. An average score of 2.0 was equivalent to the average participant saying that this item was at least “quite similar to trustworthiness”. This resulted in 7 words being carried over to Stage 4.

Stage 4: Scenario-based Testing

The first 3 stages of this project were designed to help us generate items that related to trustworthiness and then to narrow those items down to a concise list of the most relevant words or phrases. The purpose of stage 4 was to begin to collect validity and reliability evidence for the
newly created measure. The 7 words were then crafted into statements that could be rated on a Likert-type agreement scale.

Participants

Three hundred and five (109 females) participants from India participated in the study. The mean age was 30.90 (SD = 8.32). Participants were recruited via a convenience sample using Amazon’s ® Mechanical Turk ® (MTurk).

Materials and Stimuli

In this stage, participants were given the following scenario: “Please try to remember the last commercial airplane flight that you flew on. Think about the pilot of that aircraft. You may not have met him or her personally, but you know how the flight went for you. Please respond to the following statements below regarding that pilot.” Participants were then given the questionnaire (see Appendix A; note that past tense was used for this scenario) 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 one factor for the condition of trustworthiness and all items strongly loaded on this factor. A Cronbach’s Alpha test was conducted to measure internal consistency within the scale, resulting in a value of 0.88, indicating high internal consistency. Guttman split-half coefficient was 0.87.

Stage 5: Scenario-based Experiment

In the previous stage, evidence for the validity and reliability of the newly created scale was presented. The following experiment was conducted for three reasons. First, we wanted to replicate the findings of the factor analysis using a more specific scenario indicative of the kind that researchers might use in aviation consumer perception research. Second, we wanted to test the ability of the scale to discriminate between pilots who might be described as trustworthy or untrustworthy. Third, we wanted to test whether the scale correlates well with a common outcome variable that is used in aviation consumer perception research; that is, ‘willingness to fly’ in certain scenarios.

Participants

Two hundred and six (80 females) participants from India participated in the study. The mean age was 31.33 (SD = 9.78). Participants were recruited via a convenience sample using Amazon’s ® Mechanical Turk ® (MTurk).
Materials and Stimuli

Participants were randomly assigned to one of two scenarios. In the first scenario, participants were told: “Imagine a situation where you are on a commercial airline flight from one major city to another. The pilot of the airplane in known by his friends, family and colleagues to often be dishonest about his personal affairs, and sometimes cuts corners in his work performance.” In the second scenario, participants were told: “Imagine a situation where you are on a commercial airline flight from one major city to another. The pilot of the airplane in known by his friends, family and colleagues to always be honest about his personal affairs, and never cuts corners in his work performance.” Participants were then given the new seven-item trustworthiness measure (see Appendix A) and asked to provide statements of agreement or disagreement on a 5-point Likert-type scale ranging from -2 (strongly disagree) to +2 (strongly agree) with a neutral zero option. Lastly, participants were asked to respond to statements of their willingness to fly (see Appendix B), which are the type of questions used in prior studies with ‘willingness to fly’ outcomes (e.g. Rice, et al., 2014).

Scale Development

For the trustworthy condition, a factor analysis using the principle components and varimax rotation resulted in one factor for the condition of trustworthiness and all items strongly loaded on this factor. A Cronbach’s Alpha test was conducted to measure internal consistency within the scale, resulting in a value of 0.85, indicating high internal consistency. Guttman split-half coefficient was 0.92. The correlation between the trustworthiness scale and willingness scale was \( r(100) = .72, p < .001 \), indicating that the two scales strongly correlated with each other.

For the untrustworthy condition, a factor analysis using the principle components and varimax rotation resulted in one factor for the condition of trustworthiness and all items strongly loaded on this factor. A Cronbach’s Alpha test was conducted to measure internal consistency within the scale, resulting in a value of 0.89, indicating high internal consistency. Guttman split-half coefficient was 0.87. The correlation between the trustworthiness scale and willingness scale was \( r(102) = .61, p < .001 \), indicating that the two scales strongly correlated with each other.

A comparison of the two groups revealed a significant difference in scores on the trustworthiness scale, \( t(204) = 10.68, p < .001, d = 1.49 \), revealing that the scale was able to discriminate effectively between the trustworthy (\( M = 0.97, SD = 0.68 \)) and untrustworthy (\( M = -0.16, SD = .83 \)) conditions. There was also a significant difference in the ‘willingness to fly’ scale scores, \( t(204) = 10.52, p < .001, d = 1.47 \), between the trustworthy (\( M = 0.97, SD = 0.90 \)) and untrustworthy (\( M = -0.48, SD = 1.07 \)) conditions.

Discussion

As stated earlier, trust is a critical factor in the rationale behind a passenger’s choice to use air travel. Intuitively, if the passenger feels that the pilot in command is trustworthy, the passenger is more likely to be willing to fly aboard a commercial airliner. The aviation industry is interested in understanding what affects a passenger’s willingness to fly, and therefore learning their consumers’ feelings of pilot trustworthiness will be beneficial. Herein arises the need for
this newly created scale. In the past, ratings of pilot trust have been haphazardly chosen, with little evidence that questions have been tested, validated, and standardized for use by the industry. For this reason, the current research has been conducted in order to fill this gap and create a valid and reliable scale that could serve the purpose of testing a Trustworthiness of Commercial Airline Pilots (T-CAP) Scale for Indian consumers.

It is important to note that this study has chosen only positive words to be used in the scale, as it avoids cognitive confusion of switching between positive and negative choices. Previous research has shown that this practice is superior for scales intended for real world use. Harrison and McLaughlin (1991) stated that reverse-scored items “…can have detrimental effect on psychometric properties of a measure”. The researchers believe this will increase real-world effectiveness of the scale.

Validity

For a scale to be useful to the scientific and aviation communities, evidence for its validity must be presented. The current study conforms to best practices in scale development (Hinkin, 1995). The methodology section delineates the step-by-step process of conducting this research, and it begins with the generation of words relating to trustworthiness from actual consumers as well as from experts and previous scales of trust and trustworthiness. From there, new sets of potential consumers were asked to help narrow the list down over a two-stage process. We believe that the final seven items are representative of the consumers’ idea of trustworthiness in a pilot given that the majority of input and information come from the consumers themselves. Secondly, a factor analysis using the principle components and varimax rotation produced a single factor for the condition of trustworthiness, and all the items strongly loaded on this one factor.

Reliability

In addition to validity of the scale, its reliability is equally important. We tested reliability using the Cronbach’s Alpha test, which reported extremely high internal consistency, and a Guttman split-half coefficient calculation, which was equally high.

Discriminability

The last area of focus was the ability of the scale to discriminate and differentiate between a trustworthy and an untrustworthy pilot, in order to prove useful in the aviation industry. Since the scale was able to discriminate well between trustworthy and untrustworthy pilots, it showed its versatility, and its effectiveness to be used accurately by the industry, once again emphasizing its usefulness in assessing trustworthiness of a commercial pilot.

Practical Applications

The last study in this line of research was also important in that it tested the practical application of this scale in terms of the willingness of a consumer to get on board an aircraft. The scale correlated strongly with the willingness outcome, and showed the relationship between the
ratings of trustworthiness in a commercial pilot, and consumers’ willingness to fly on the aircraft. While correlations do not prove causal relationships, we believe that this evidence still adds to the usefulness of the scale in the aviation industry. Further research should, of course, focus on examining a possible causal relationship between the two constructs; however, that is beyond the purview of this study.

Once this scale is deployed, it will aid several aspects of the aviation industry. Using this scale, airlines will be able to measure consumers’ perceived trust in pilots based on a variety of characteristics, traits, hypothetical or real-world situations (e.g. assessing consumer reactions to incidents or accidents such as the Malaysian Airlines Flight MH370). It will also allow the airlines to make comparison between pilots, training programs, etc., for future decision making about hiring, firing and/or promoting pilots.

Different scales, using modifications based off of this established scale could be created for use in different fields. The universality of this scale is limited to the words relating to trustworthiness in a commercial pilot by Indian consumers, but could be amended, with research, to be used in other consumer-oriented fields regarding trustworthiness. As mentioned earlier, there are varying perceptions and values that affect trustworthiness based on culture, and therefore this scale serves the purpose of differentiating on the basis of nationality and culture. This scale can therefore be amended and utilized in several Indian industries that require the passengers to place trust within an operator. For instance, this trustworthiness scale could be modified to aid in the research of consumer perceptions of trustworthiness in a ship captain, train operator, bus driver, etc. With a large amount of responsibility, in terms of passenger safety, placed on operators of public transportation, ratings of trustworthiness are very useful information to have.

Limitations

Despite the findings associated with this study, the research design does have certain associated limitations that must be addressed. Firstly, actual in-field success will need to be conducted over time to replicate the current findings of validity and reliability. As was mentioned earlier, this scale was developed for use by Indian consumers only, and therefore has its geographic limitations. Future scales can be developed to emulate this one, and be tailored around different nationalities or cultures.

A reasonable argument can be made that use of online surveying tools such as Amazon’s ® Mechanical Turk ® (MTurk) is not ideal for sampling participants. The tool enables convenient sampling, but does not necessarily provide data from all aviation consumers in India. Lastly, the scale is a consumer-based scale, and while that provides some advantages, it may not be as useful when experts are asked to rate pilot trustworthiness. For example, if a student pilot is asked to rate the trustworthiness of her or his certified flight instructor (CFI), then the scale might need to be modified to include items more relevant to that scenario.
Conclusions

A Trustworthiness of Commercial Airline Pilots (T-CAP) Scale was created for the purpose of deepening the understanding of the relationship between passenger and pilot. A passenger’s perception of pilot trustworthiness is critical in a field that involves one individual to place so much control in another. The scale was created with an eye towards being able to extract the true perceptions of trustworthiness of the passengers from an Indian culture. The scale was developed to not only be valid and reliable, but also to be of practical usefulness to the aviation industry in India. It has the potential to allow airlines to understand their passengers better and therefore improve the overall experience of the customer. The end result of using this scale should be to increase the level of trust of the passengers in the pilots, and therefore increase their comfort and willingness to fly on the aircraft, and in turn on that particular airline. This research opens the door for several studies to be conducted in this sphere, and improve the understanding of trust in pilots.
References


Appendix A

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

1. The pilot is qualified.  
   | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
2. The pilot is talented.  
   | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
3. The pilot is reliable.  
   | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
4. The pilot is efficient.  
   | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
5. The pilot is experienced.  
   | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
6. The pilot is active.  
   | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
7. The pilot is trustworthy.  
   | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
### Appendix B

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