

2015

A Valid and Reliable Safety Scale for Passenger's Perceptions of Airport Safety

Stephen Rice

Rian Mehta

Scott Winter

Korhan Oyman

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2015



Part of the [Other Psychiatry and Psychology Commons](#)

Repository Citation

Rice, S., Mehta, R., Winter, S., & Oyman, K. (2015). A Valid and Reliable Safety Scale for Passenger's Perceptions of Airport Safety. *18th International Symposium on Aviation Psychology*, 242-247.
https://corescholar.libraries.wright.edu/isap_2015/66

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2015 by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu, library-corescholar@wright.edu.

A VALID AND RELIABLE SAFETY SCALE FOR PASSENGER'S PERCEPTIONS OF AIRPORT SAFETY

Stephen Rice, Florida Institute of Technology
Rian Mehta, Florida Institute of Technology
Scott Winter, Florida Institute of Technology
Korhan Oyman, Florida Institute of Technology

Previous research has developed various customer satisfaction scales in many applied areas; however, to our knowledge, there is not a validated scale for measuring commercial airline passengers' ratings of personal safety based on airport security. The current study seeks to address this missing gap by developing a valid and reliable safety scale for commercial airline passengers (SS-CAP). We first solicited words and phrases that are related to a passenger's feeling of safety from potential consumers and experts in the field. We then narrowed down the list to 7 remaining items. Lastly, we tested the scale using participants from Amazon's Mechanical Turk, which is a reliable source for participants in online surveys. A principal components factor analysis with varimax rotation revealed that all items loaded strongly on one factor, accounting for 78% of the variance in the model. A Cronbach's alpha test revealed high internal consistency, $r = 0.95$. A Guttman split half test showed high reliability, $r = 0.95$. These results provide strong evidence for a valid and reliable scale of passenger ratings of personal safety. The scale statements include: I feel safe, I feel secure, I feel protected, I feel guarded from danger, I feel shielded from harm, I feel at ease, I feel sheltered from threats. Participants should respond on a 5-point Likert type scale scored from strongly disagree (-2) to strongly agree (+2).

Introduction

Safety is of the utmost importance, especially in high consequence industries such as aviation (Janic, 2000; Maurino, 2000; Sarter & Alexander, 2000). As aviation continues to grow, countries are investing in developing, expanding, and building new airports. The Atlanta Hartsfield-Jackson airport recently opened a fifth runway and Abu Dhabi and Istanbul are investing in the development of new airports. As this growth continues so must the safety record within this industry. Numerous authors (Patankar & Sabin, 2010; Sarter & Alexander, 2000) express concern that this growth could result in an increase in accidents.

Safety is defined as being “free of harm” and “the state of being safe” (Merriam-Webster, n.d.). Safety is also framed within the social construct (Maurino, 2000) and is related to risk. What is deemed safe or risky in one society may be outside the tolerances of another. Therefore, safety is somewhat of a continuum, but there is no question that in high consequence industries, such as aviation, the margin of error is extremely low.

Airports play a key role in the aviation system. The majority of all flights originate and arrive at airports. These facilities serve as the backbone of a large network of flights where passengers and cargo are loaded and aircraft are maintained. The complex maze of taxiways and runways must provide for safe movement of aircraft, and if a mistake is made, the results could be catastrophic. Of interest in the current study was to develop a valid and reliable scale that could be used to measure consumer’s perceptions on airport safety.

Human error remains the leading cause of safety infractions in aviation. Experts estimate that as many as 70% to 80% of accidents are attributed to human error (Sarter & Alexander, 2000). However, it is possible that accidents are the results of compounding issues. Poor airport design or layout could compile the human related factors. Many airports are now being designed to reduce the number of possible locations where a runway incursion (possible collision or loss of separation) could occur. Recent studies have focused on the measurement of safety management systems and safety culture (Gill & Shergill, 2004; von Thaden & Gibbons, 2008), locus of control (Hunter, 2002), and commercial aviation safety culture (von Thaden, Wiegmann, Mitchell, Sharma, & Zhang, 2003).

Studies by Gill and Shergill (2004), von Thaden and Gibbons (2008), and von Thaden, Wiegmann, Mitchell, Sharma, and Zhang (2003) reviewed safety management and views toward safety culture. However, these studies were focused on the perceptions of operators within the system and not the consumers of these processes. Operators have more extensive levels of training within the system and therefore different perspectives. Hunter (2002) studies the measure of locus of control. Locus of control is defined as the level a person believes that the outcomes of certain situations are under their personal control. Hunter’s measure was pre-existing and adjusted to be a valid and reliable instrument that could be used in the aviation field.

While these studies all provide for accurate measures of safety by operators, there appears to be a gap in the literature related to 1) consumer perceptions of safety and 2) their views toward the safety of airports. Consumer perceptions can have powerful influence within aviation. Many consumers did not like flying on turboprop aircraft, and now few of them are used for commercial transportation. In fact, some consumer ticketing websites have checkboxes to remove those options that include turboprops as part of the flights (e.g. Kayak). Therefore, having a valid and reliable instrument to measure consumer perceptions towards airports may provide valuable insights into those items that are most important to consumer’s views toward airport safety.

Methodology

Stage 1: Word Generation

In Stage 1, we began generating items for the scale. We first solicited words and phrases from experts in the field and similar scales in the literature. We then selected participants for an online survey in order to generate items from people who might actually use the scale in the future. Since consumers are the ones who will be responding to the scale, we felt that this increased the validity of the process.

Participants. Seventy-two (28 females) 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 32.98 ($SD = 9.35$). Three additional participants with expertise in safety and/or security provided further items for review. Lastly, the trust literature was reviewed, and items were added accordingly.

Materials and Stimuli. Participants were presented with the following scenario: "In the context of a commercial airport, please enter 5 characteristics of safety 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 232 unique words or phrases. These items were then reviewed for correct spelling and de-capitalized when necessary to ensure uniformity.

Stage 2: Nominal Paring

In Stage 2, we began eliminating words or phrases that were not perceived by participants as being related to the construct of safety as it relates to a commercial airport.

Participants. Forty-nine (23 females) participants from the United States were recruited via a convenience sample using Amazon's Mechanical Turk (MTurk). The mean age was 32.65 ($SD = 11.26$).

Materials and Stimuli. Each of the 232 items generated in the first stage were presented to participants, along with the following statement, "In the context of a commercial airport, please rate whether each word below is related to (similar to) safety, not related to (not similar to) safety, or you don't know." Forty-two items were chosen to be related to trustworthiness by at least 85% of participants.

Stage 3: Likert-scale Paring

In Stage 3, we continued narrowing down the list of items that would be retained for the final scale. Looking for a more sensitive measure of the relationship between the items and safety, we used a Likert-type scale instead of a nominal scale.

Participants. Forty-eight (22 females) participants from the United States were recruited via a convenience sample using Amazon's Mechanical Turk (MTurk). The mean age was 32.23 ($SD = 10.59$).

Materials and Stimuli. The 42 items retained from Stage 2 were presented to participants with the following statement, "In the context of a commercial airport, please rate how strongly each word below is related to safety." Participants responded based on a Likert-type scale from "Not at all related to safety" (0) to "Extremely related to safety" (+3). Seven items averaged 2.5 or higher and were retained for the final scale.

Stage 4: Scenario-based Testing

In Stage 4, we began collecting evidence of validity and reliability for the new scale. The seven items were converted into statements that could be rated on a 5-point Likert-type scale from strongly disagree (-2) to strongly agree (+2).

Participants. Two hundred and twenty-two (83 females) participants from the United States were recruited via a convenience sample using Amazon's Mechanical Turk (MTurk). The mean age was 31.09 ($SD = 10.55$).

Materials and Stimuli. In this stage, participants were presented with the following scenario: "Please try to remember the last commercial airplane flight that you flew on. Think about the airport security that you interacted with. In the context of the airport security described above, please respond to the following statements to the best of your ability." 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 0.95 indicated high internal consistency. A Guttman split-half test was conducted as well. The resulting coefficient of 0.95 indicated high reliability.

Discussion

The purpose of this study was to create a valid and reliable instrument for obtaining airline consumer ratings of personal safety based on airport security. Human beings value their personal safety and security, and is oftentimes of the highest priority when making decisions. This remains true when passengers consider which airports they use for their travels. If a passenger believes an airport has a lower level of security as compared to a neighboring airport, it can be assumed that they would choose the safer of the two options. The consumers' perception of the safety factor of an airport can have a significant influence on airport traffic. An additional purpose of this research was to fill a void in the aviation related literature regarding a valid and reliable measure that can be used to capture consumer perceptions on airport safety.

Only positively scored items related to aircraft safety were enlisted to develop the scale. In doing so, it prevents the need for the participant to cognitively switch between negative and

positive words. This in turn eliminates the possibility of cognitive confusion. Research has indicated that a negative psychometric effect could be observed with the use of reverse scored items (Harrison & McLaughlin, 1991).

The creation of a valid and reliable scale is of practical value to the airport operations field as well as the research community. The study fills a gap by developing a metric that measures a vital airport consideration. The additional benefit lies in the fact that the developed scale is has been statistically proven for validity, reliability, and discriminability. While knowledge of the inner working of the industry are important, it is relevant to note that this scale was developed using words generated by consumers and not industry experts. A consumer perception scale specifically for airport security could be of valuable assistance to most airport management staffs across the country. Additionally, the creation of such a scale lays the foundation for future scales to be developed, within the realm of security, as well as for other airport facets.

While it is beneficial to create a consumer scale using actual consumers, the means of data collection have certain limitations. Each stage of the generation and statistical analysis phase use responses generated by participants for compensation from Amazon's ® Mechanical Turk ® (MTurk). Research by Buhrmester, Kwang, and Gosling (2011) state that this form of data is a reliable as laboratory data. Since the recruited participants were only recruited from the United States, the generalizability of the findings is limited to U.S. airports and consumers. Future steps along this line of research could seek to field test the instrument in person at airports around the United States. Lastly, further studies could seek to better understand the relationship, if any, between the frequency of a consumer's travels and their perception of airport security.

Conclusion

The aim of this research endeavor was to fill a gap in scientific community by creating a valid and reliable scale to measure consumer perceptions of airport security. In doing so, it allows airport operators to utilize such an instrument in order to better understand, and perhaps address, certain concerns or misconceptions of the passengers. Consumer responses through a multi-stage process using word generation, word paring, and scenario-based examples were used to create the instrument. It is the hope that this scale could be used as a tool for data collection on perceptions of airport security, and enable improvements in that sector.

References

- 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.
- Gill, G. K., & Shergill, G. S. (2004). Perceptions of safety management and safety culture in the aviation industry in New Zealand. *Journal of Air Transport Management*, 10, 233-239.
- 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. *Proceedings of the 1991 Academy of Management annual meetings*, 310-314.
- Hunter, D. R. (2002). Development of an aviation safety locus of control scale. *Aviation, Space, and Environmental Medicine*, 00(0), 000-00.
- Janic, M. (2000). An assessment of risk and safety in civil aviation. *Journal of Air Transport Management*, 6, 43-50.
- Maurino, D. E. (2000). Human factors and aviation safety: What the industry has, what the industry needs. *Ergonomics*, 43(7), 952-959.
- Merriam-Webster (n.d.) In Merriam-Webster online dictionary. Retrieved from: <http://www.merriam-webster.com/dictionary/safety>
- Patankar, M., S., & Sabin, E. J. (2010). The safety culture perspective. In Salas, E. & Maurino, D. (Eds.), *Human Factors in Aviation*, (95-122). Burlington, Elsevier.
- Sarter, N. B., & Alexander, H. M. (2000). Error types and related error detection mechanisms in the aviation domain: An analysis of aviation safety reporting system incident reports. *The International Journal of Aviation Psychology*, 10(2), 189-206.
- von Thaden, T. L., & Gibbons, A. M. (2008). The safety culture indicator scale measurement system (SCISMS). Human Factors Division Institute of Aviation. Savoy, Illinois.
- von Thaden, T. L., Wiegmann, D. A., Mitchell, A. A., Sharma, G., & Zhang, H. (2003). Safety culture in a regional airline: Results from a commercial aviation safety survey. Proceedings of the *International Symposium on Aviation Psychology*, Dayton, Ohio.

Appendix A

Please respond to each of the following statements:

I feel safe.

Strongly Disagree DisagreeNeutral Agree Strongly Agree

I feel secure.

Strongly Disagree DisagreeNeutral Agree Strongly Agree

I feel protected.

Strongly Disagree DisagreeNeutral Agree Strongly Agree

I feel guarded from danger.

Strongly Disagree DisagreeNeutral Agree Strongly Agree

I feel shielded from harm.

Strongly Disagree DisagreeNeutral Agree Strongly Agree

I feel at ease.

Strongly Disagree DisagreeNeutral Agree Strongly Agree

I feel sheltered from threats.

Strongly Disagree DisagreeNeutral Agree Strongly Agree