

The Flight Risk Perception Scale (FRPS): A Modified Risk Perception Scale for Measuring Risk of Pilots in Aviation

Scott R. Winter
Embry-Riddle Aeronautical University

Dothang Truong
Embry-Riddle Aeronautical University

Joseph R. Keebler
Embry-Riddle Aeronautical University

Risk and risk perception remain focal areas of research within the aviation domain. The purpose of the current study was to assess an existing measure of a 26-item self-risk perception scale for pilots. A sample of 490 participants was used in the present study, and a confirmatory factor analysis was conducted on the original 26-item instrument. The findings indicated that there was a poor model fit of the original instrument. Through the use of modification indices, a new 13-item scale was produced, which resulted in a second-order CFA model. Flight risk was shown to be the second-order construct with general flight risk, high risk, and altitude risk as the first-order constructs. The new model reported good psychometric values of GFI of 0.933, AGFI of 0.893, CFI of 0.947, NFI of 0.923, normed chi-squared of 3, and RMSEA of 0.071. The findings produce a new 13-item scale that can be used by aviation researchers who wish to conduct studies related to the pilot's self-assessment of risk perception.

Recommended Citation:

Winter, S. R. Truong, D., & Keebler, J. R. (2019). The flight risk perception scale (FRPS): A modified risk perceptions scale for measuring risk of pilots in aviation. *International Journal of Aviation Research*, 11(01), 57-72.

Risk is an ever-present factor in everyday life. When it comes to flying, a pilot's perception of risk may influence their decision to operate in certain conditions (Molesworth & Chang, 2009; Molesworth, Wiggins, & O'Hare, 2006). As a result, being able to measure a pilot's self-risk assessment is a valuable metric in aviation research. Hunter (2006) published a 26-item self-risk assessment scale for use with pilots. This instrument has been used in several prior studies (Drinkwater & Molesworth, 2010; Hunter, Martinussen, Wiggins & O'Hare, 2011; You, Ji, & Han, 2013). The purpose of this paper was to reassess this original instrument, which demonstrated mathematical and conceptual issues, to determine if the factor structure and items hold or if new data suggest a modified risk perception scale should be created to effectively measure self-risk assessments in pilots.

Risk perception presents many opportunities for future research within the aviation industry. Since pilots take risks every flight, it is an area where researchers may seek to gain a better understanding. For instance, why did the pilot continue flying into deteriorating weather conditions, try to land on a runway which was too short, or run an aircraft out of fuel? Risk perception plays an influential role in the decision-making and judgments of pilots, and thus, it is frequently seen as a desired variable to be measured during research studies. Risk may also be related to hazardous attitudes in pilots, which could be precursors to the types of pilots who may be willing to take more risks during flight. These concepts justify the need for a valid instrument to measure risk perception in pilots.

Literature Review

Risk Perception of Pilots in the Aviation Industry

Risk is something humans are exposed to daily (Hansson, 2005), with humans frequently taking on high-risk activities. Risk is commonly defined as the "possibility of a loss" (Merriam-Webster, 2019, n. p.), and it is described by the Federal Aviation Administration (FAA) as the balance or matrix between the likelihood of a particular outcome by the severity of that outcome (2009). For example, a plane crash is somewhat unlikely to occur, but the severity level due to total damage costs, injuries or fatalities would be quite high, thus increasing the associated risk level. Similarly, there are certain phases of flight, specifically takeoff and landing, which are considered the two critical phases of flight due to their increased level of risk since the aircraft are relatively slower and closer to the ground than during other parts of the flight (FAA, 2009).

Related to risk is a pilot's risk perception, which remains a focal point of aviation research. Risk perception is defined as the cognitive ability of a pilot to both recognize and accurately assess the level of risk compared with their personal skills to handle the situation (Hunter, 2002). A disconnect between these two assessments can result in a potentially dangerous situation. The National Transportation Safety Board (NTSB) noted that within between 1989 and 2009, around 85% of aviation accidents were attributed to pilot error (FAA, 2009), and updated metrics from Oster, Strong, and Zom (2013) found 83% of aviation accidents from 1990-2013 were a result from which human error was a least a contributing factor, while the most recent Nall Report cites pilot-related issues in 74% of accidents (AOPA, 2018). Part of these errors is likely related to poor risk perception and the resulting decisions. A prior study by Orasanu, Fischer, and Davison (2002) found that pilots typically tend to have optimistic attitudes

toward risk where they overestimate their abilities and underestimate the risk of loss, sometimes referred to as overconfidence bias. This finding may be useful in explaining pilot related errors to events such as plan continuation errors (PCE), where, for example, pilots continue flight into deteriorating weather conditions.

In some more recent studies on pilots and risk, Joseph and Reddy (2013) reviewed risk-taking perceptions of helicopter pilots in the Indian Army. Their findings suggest participants with low self-confidence and safety orientation scores were more likely to take on higher levels of risk. Knecht and Frazier (2015) examined the risk of pilots using graphical weather tools in their flight planning. They found that higher motivation to complete the flights increased the risk-taking of pilots, and pilot generally had higher risk tolerance levels than desired. Wiggins, Hunter, O'Hare, and Martinussen (2012) used a sub-set of the existing scale (only 10 of the 26 items) to assess a pilot's decision to continue flight into instrument meteorological conditions. Lastly, using the existing Risk Perception Scale, Ji, Yang, Li, Xu, and He (2018) found that risk perception scores mediated the relationship between trait mindfulness and likelihood to be involved in an incident. These studies also demonstrate the need for additional research to be conducted related to risk perception in aviation.

Assessment of Risk Using a Revalidated Scale

Risk remains a construct of interest in aviation research. It is linked closely with decision-making and studies have focused on investigating the risk, decision, and judgment of pilots for many years (Jensen & Benel, 1977; Ji, You, Lan, & Yang, 2011; Hunter, 2002; Molesworth & Chang, 2009; O'Hare, 1990; You, Ji, & Han, 2013). Since risk is a latent construct, researchers need to rely on valid scales to provide measurement. Latent variables are those variables which are not directly observable (Byrne, 2010); flight risk is an example of one such variable. These latent variables are represented by manifest variables, which are directly observable variables. In another example, IQ is the latent construct, and the IQ test provides the manifest variables. Scale validation helps ensure the scale being used by the researcher is measuring the construct under investigation. An initial assessment using the current dataset found areas of concern with the original risk perception instrument, which resulted in questionable validity. Specifically, there were individual items which loaded onto multiple latent factors. This threatens the validity of the overall scale (Blunch, 2013) and fails to produce unidimensional measures (Hair et al., 2010). Therefore, this warranted a further and thorough investigation to revalidate the original risk perception scale to verify its validity for use to measure risk levels in aviation.

When conducting studies, there is immense value in using a validated scale over an invalid scale (Wilson & Joye, 2017). A construct-valid measure has been statistically shown to measure the construct under investigation, whereas, a newly created scale may not, in fact, be measuring the construct intended by the researchers. For aviation researchers wishing to investigate risk, using an instrument which has not demonstrated the psychometric properties of validity could cause multiple threats to the internal validity of the findings. However, there can also be a contradiction between the length of the scale and its usefulness. If time were not an issue, many researchers would prefer larger, multi-item instruments to ensure good validity and

reliability of the constructs under investigation (Gosling, Rentfrow, & Swann Jr., 2003); however, scarce resources and limited time usually are real challenges incurred by researchers.

Achieving a balance between scale validity and its length leads to the notion that in scale development, researchers should strive for parsimony. The goal of parsimony is for researchers to develop scales with the minimum number of items needed to represent the desired construct (Hair et al., 2010). As a result, there is value in developing an instrument that is both valid and efficient to administer. With the original instrument having individual items load on multiple factors, this threatens the parsimony of the scale, along with the validity. Additionally, from a procedural standpoint, a shorter instrument will also help reduce participant fatigue, especially in studies where multiple measures are proposed. It will also help if intended to be administered longitudinally or in scenarios such as a pre- and post-test (Robins, Hendin, & Trzesniewski, 2001). Prior studies have shown that short scales can have just as much validity as longer scales (Burisch, 1984, 1997). Based on the threats to the existing scale, these concepts are relevant to the current study as Hunter's original 26-item scale is reviewed for validity and to determine if any further item reduction is possible. This original instrument was determined to consist of five factors, namely, general flight risk, high risk, altitude risk, driving risk, and everyday risk. Ten items each loaded onto both the general flight risk and high risk factors. Altitude risk consisted of 7 factors while driving risk and everyday risk consisted of 3 and 4 items, respectively. This immediately presents a concern as the total of all factor items equals 34, which is more than the 26 items on the scale, and our initial analysis of the 26-item scale showed single items loading on multiple factors, which is problematic in regards to construct validity (Blunch, 2013). Based on these issues, there was grounding to revalidate the original scale with the objective of producing a construct valid measure of Flight Risk.

Current Study

The purpose of the present study was to examine the factor validity of the Risk Perception – Self scale initially developed by Hunter (2006). This 26-item assessment was previously shown to measure risk perception using a five-factor structure. The Confirmatory Factor Analysis model in this study used Hunter's five factors and these 26 question items. The five factors identified in the original model were: general flight risk, high risk, altitude risk, driving risk, and everyday risk. Using a new sample of participants, the current study conducted a confirmatory factor analysis to see if 1) the factor structure held as proposed in the original scale, and 2) if any further items could be reduced while maintaining high levels of validity. Both construct validity and construct reliability will be evaluated. Construct validity consists of convergent and divergent validity. Convergent validity demonstrates that the items on a construct are highly correlated with one another, while divergent validity demonstrates that each factor in the measurement model differs from the others. Lastly, construct reliability is evaluated to ensure adequate levels of reliability exist for the scale.

Methods

Participants

Four hundred and ninety (20 females) members of the Aircraft Owner's and Pilot's Association's (AOPA) Air Safety Institute (ASI) participated in the study. An email requesting participation in the study was sent to approximately 9,800 members, and the response window remained open for approximately three weeks, indicating a response rate of around 5%. The average age of participants was 60.46 years ($SD = 13.58$), and they reported an average of 3,357.07 ($SD = 5353.95$) total flight hours ($MDN = 1,246.50$). Figure 1 depicts the certificates and ratings held by participants in the study. The majority of pilots were private pilots (52.6%), and a majority of participants also held instrument ratings (56.2%). Most pilots indicated that they flew Part 91 recreationally (78.7%), with a few flying Part 121 (6.7%), Part 91 business/corporate (8.5%), Part 135 (2.0%), military (0.8%), Part 91K (0.2%), or other (3.1%).

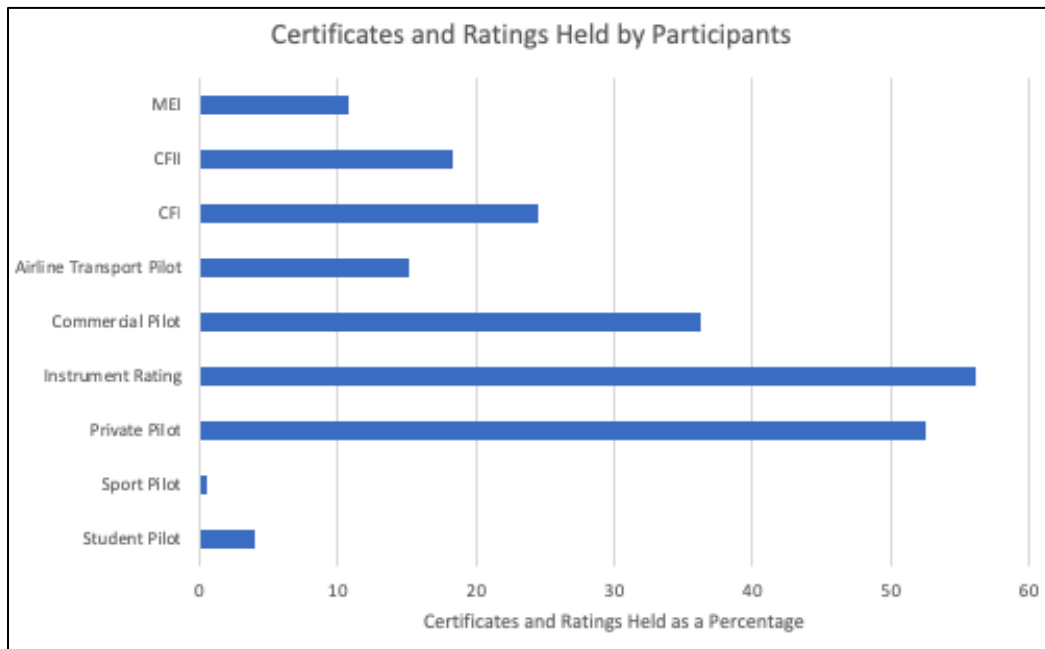


Figure 1. Demographic data of participants showing the percentage of certificates and ratings held. Participants were able to select multiple certificate and ratings.

After an initial screening of the data, 370 cases were deemed valid for use in the data analysis. The main reason for a case being removed was due to incomplete responses in the questionnaire. All responses with incomplete data were removed prior to data analysis. The minimum suggested sample size to conduct the data analysis was 229 usable cases based on the assumption of a small to medium effect size (0.25), statistical power of 0.80, 5 latent variables, 26 observed variables, and an alpha level of 0.05 (Soper, 2019).

Procedures and Materials

Participants were solicited through an email notification. The data collection window remained open for approximately three weeks, and ASI sent a reminder email at roughly the halfway point. Within the email was a link to the questionnaire, which was hosted using Google Forms®. After clicking on the link, participants were first presented with an electronic consent form, which they had to accept to complete the study. Following this, they were presented with instructions, and they were then presented a series of instruments, specifically, the 44-item Big-

Five Inventory (BFI, Benet-Martinez & John, 1998; John & Srivastava, 1999); the 27-item Aviation Safety Attitude Scale (Hunter, 2005); five questions related to their participation in AOPA ASI safety materials such as videos, online training courses, and seminars; the 26-item Risk Perception – Self scale (Hunter, 2006), and lastly demographic information. The study took approximately 15 minutes for participants to complete. The complete instrument was administered to produce a full structural model of pilot attitudes toward taking risk. However, the first step in assessing a structural model is to verify adequate factor structure. Due to the validity problems associated with the Risk Perception – Self scale, the researchers had to revalidate the Risk Perception – Self scale due to a series of issues, specifically, poor model fit and individual items loading onto more than one factor, which heightens validity concerns.

Design

The study used a quantitative method and a non-experimental correlational design. Confirmatory Factor Analysis (CFA) was the statistical procedure used to assess the validity and reliability of the instrument to assess a pilot's self perception of risk. CFA is conducted when the researcher's specify the relationships between latent and manifest variables a priori, in this case, based on the original scale. Through the process of CFA, a measurement model can be determined, and it allows the researcher to make a decision as to the accuracy of the pre-determined model (Hair et al., 2010). Establishing a valid measurement model is frequently the first step in developing a full structural model. In the case of this current study, since the original risk perception scale failed this measurement model stage, a full structural model (the original goal of the study), was unable to be conducted and required a revalidation of the risk scale be completed. The results are presented in the following section.

Results

Confirmatory Factor Analysis (CFA) was utilized for this analysis through IBM SPSS AMOS 24. In order to evaluate the model fit indices, this research uses a Goodness-of-Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Normed Fit Index (NFI), Comparative Fit Index (CFI), Normed Chi-Squares, and Root Mean Square Error of Approximation (RMSEA). GFI, AGFI, and NFI are recommended to be at least 0.9 to ensure the model fit. Additionally, it is recommended that CFI should be greater than 0.93, and RMSEA should be less than 0.06 (Byrne, 2010).

Should the model fit be unsatisfactory, the model respecification would be conducted to determine the best fit factor structure. In this step, Modification Indices (MIs) are examined to determine necessary changes that should be made to the measurement model to improve the model fit. This is an exploratory process in which multiple iterations were conducted by making one change at a time.

Then, reliability of the constructs was assessed using Cronbach's Alpha and Construct Reliability (CR). It is recommended that alpha and CR should be greater than 0.7 to achieve good construct reliability (Nunnally, 1978; Hair et al., 2010). Finally, construct validity, including convergent and discriminant validity, was assessed using Average Variance Extract (AVE) and Maximum Shared Variance (MSV) methods following guidelines provided by

Fornell and Larcker (1981) and Hair et al. (2010). To achieve good construct validity, AVEs should be greater than 0.5 and higher than the corresponding MSVs for all constructs.

The analysis results show that the initial CFA model did not have a good model fit, so MIs were examined to determine necessary changes to the factor structure. Table 1 presents the model fit indices for the initial CFA model and the second CFA model, which reflects the respecification as described next.

Table 1

Model fit indices for the initial CFA model and the final CFA model

Fit indices	Initial CFA Model	CFA Model 2
CMIN/DF	2.963	3.000
GFI	0.856	0.909
AGFI	0.816	0.865
CFI	0.885	0.935
NFI	0.837	0.907
RMSEA	0.073	0.074

An examination of the results indicated that items for Everyday Risk construct have low factor loadings (less than 0.5) and high variance, resulting in low construct reliability (Cronbach's Alpha is 0.429 and CR is 0.152) and low convergent validity (AVE is 0.195). Additionally, the high correlation between Driving Risk and Everyday Risk indicated poor discriminant validity of this construct. These issues brought up the concerns with the relevance and validity of the question items for this construct. Accordingly, Everyday Risk was removed from the measurement model (Blunch, 2013; Hair et al., 2010).

In addition to Everyday Risk, the CFA results also revealed concerns with question items for the High Risk construct. Specifically, the items labeled RPS9 and RPS18 have low factor loadings and high variance, resulting in low reliability and validity of this construct. Further examining these question items, it was noted that these questions seemed ambiguous and overlapped with other questions, which lead to the low factor loadings and inter-item correlations. Hence, these items were removed. Following the same process, further concerned items were identified and removed in an iterative process, including the items RPS14, RPS18, RPS21, and RPS26. It is important to note that the question content and literature were examined to justify removing these items (Blunch, 2013; Hair et al. 2010). Table 2 shows the CFA model with the items for each construct. This CFA model has a satisfactory model fit with GFI of 0.909, AGFI of 0.865, CFI of 0.935, NFI of 0.907, normed chi-squared of 3, and RMSEA of 0.074. AGFI and RMSEA are outside the boundaries of recommended levels but considered acceptable (Hair et al., 2010). Table 2 also shows the construct reliability and validity results. Flight Risk, Altitude Risk, and Driving Risk have satisfactory construct reliability with Cronbach's Alphas and CRs higher than 0.7. These constructs also have good convergent validity with AVEs greater than 0.5 and good discriminant validity with AVEs higher than corresponding MSVs. High Risk is the construct that has mediocre construct reliability and validity results. However, given the importance of this construct and the satisfactory factor loadings, it was decided to keep this construct in the measurement model.

Table 2

Confirmatory Factor Analysis Results, Construct Validity, and Construct Reliability for Model 2.

Construct	Items	Factor loading	Variance	Average Variance Extract (AVE)	Construct Reliability (CR)	Cronbach's Alpha	Maximum Shared Variance (MSV)
Flight Risk	RPS1	0.794	0.669	0.544	0.702	0.841	0.174
	RPS3	0.714	1.146				
	RPS5	0.796	1.225				
	RPS13	0.631	1.145				
	RPS23	0.741	1.540				
High Risk	RPS7	0.686	2.649	0.449	0.329	0.603	0.504
	RPS10	0.710	2.536				
	RPS19	0.609	3.008				
Altitude Risk	RPS4	0.843	1.269	0.588	0.616	0.861	0.504
	RPS8	0.761	1.458				
	RPS15	0.826	1.548				
	RPS22	0.722	2.470				
Driving Risk	RPS24	0.669	2.338	0.689	0.632	0.861	0.192
	RPS11	0.895	0.714				
	RPS17	0.701	2.056				
	RPS20	0.880	0.793				

The results of this CFA model shows that Flight Risk, High Risk, and Altitude Risk are all flight-related while Driving Risk is not. To determine if a more parsimonious scale could measure flight risk perception, an alternative measurement model was created as a second-order CFA model, shown in Figure 2, without the factor of Driving Risk. In this CFA model, Flight Risk is a second-order construct attributed by three first-order constructs, General Flight Risk, High Risk, and Altitude Risk. The same process was followed to determine the model with a good fit. Figure 2 shows the final second-order CFA model. This final model has a better model fit than the previous CFA models with GFI of 0.933, AGFI of 0.893, CFI of 0.947, NFI of 0.923,

normed chi-squared of 3, and RMSEA of 0.071 (see Table 3). Additionally, Table 4 presents the factor loadings, AVE, Cronbach's alpha, CR, and MSV for this final CFA model. The results are very similar to the previous CFA model 2 as presented in Table 2, confirming that this model has good construct reliability and validity. This data suggests that the final CFA model is a better and more valid assessment of risk perception than the previously existing model as depicted in Tables 3, 4, and Figure 2. A version of the new scale can be found in Appendix A - Flight Risk Perception Scale (FRPS).

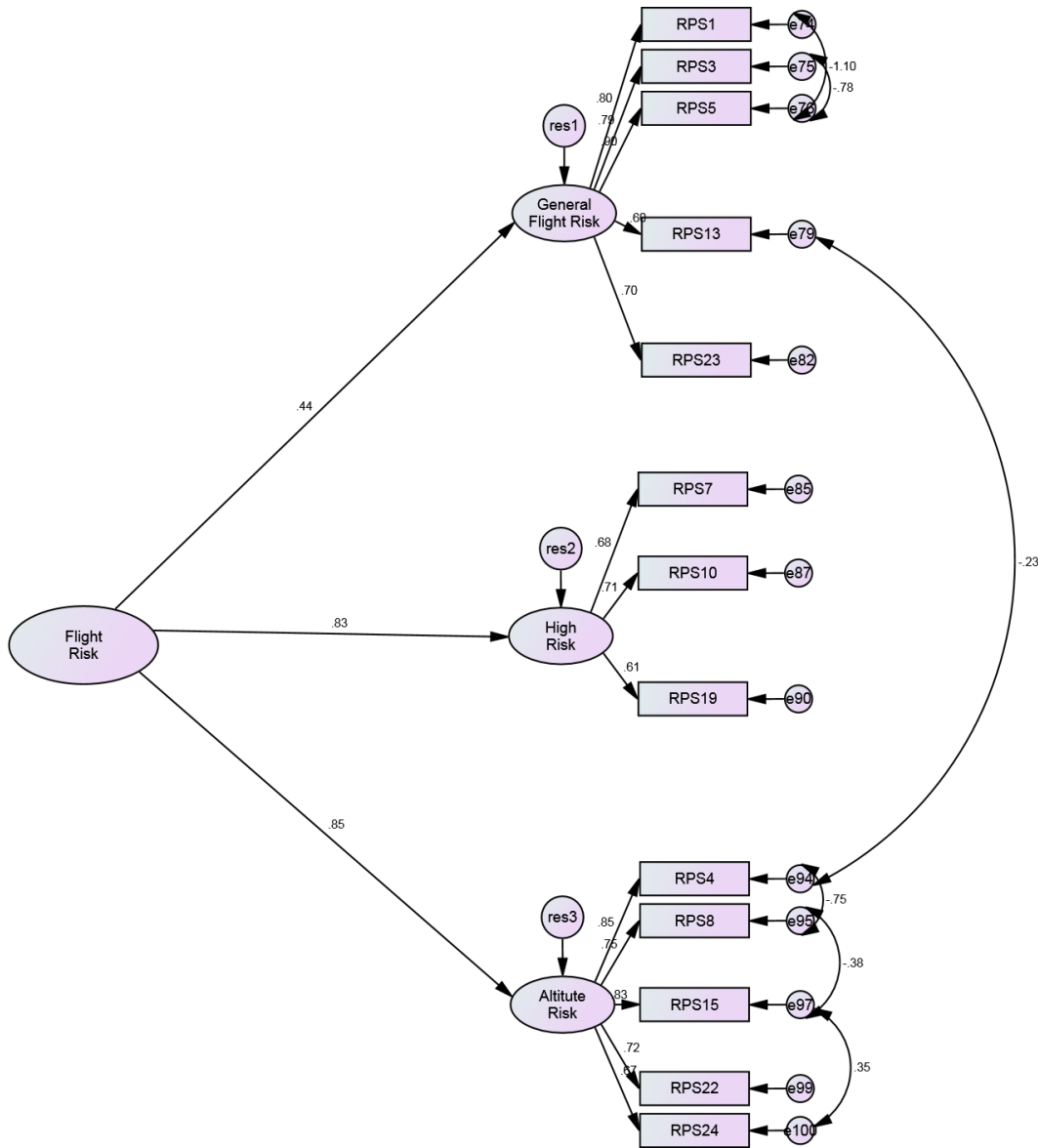


Figure 2. Final CFA Model depicting a second-order structure.

Table 3*Model fit indices for the second CFA model and the final CFA model*

Fit indices	CFA Model 2	Final CFA Model
CMIN/DF	3.000	3.000
GFI	0.909	0.933
AGFI	0.865	0.893
CFI	0.935	0.947
NFI	0.907	0.923
RMSEA	0.074	0.071

Table 4*Confirmatory Factor Analysis Results, Construct Validity, and Construct Reliability of the Final Model*

Construct	Items	Factor loading	Variance	Average Variance Extract (AVE)	Construct Reliability (CR)	Cronbach's Alpha	Maximum Shared Variance (MSV)
Flight Risk	RPS1	0.790	0.682	0.544	0.703	0.841	0.172
	RPS3	0.710	1.161				
	RPS5	0.805	1.179				
	RPS13	0.629	1.151				
	RPS23	0.742	1.533				
High Risk	RPS7	0.683	2.67	0.449	0.329	0.603	0.503
	RPS10	0.712	2.523				
	RPS19	0.611	2.997				
Altitude Risk	RPS4	0.846	1.243	0.588	0.616	0.861	0.503
	RPS8	0.752	1.517				
	RPS15	0.828	1.538				
	RPS22	0.724	2.459				
	RPS24	0.672	2.324				

Discussion

The purpose of the current study was to assess the Risk Perception – Self scale published by Hunter (2006). That study produced a 26-item scale which was demonstrated for use in pilot self-assessment of risk. The results show that Hunter’s instrument, as reflected in the initial CFA model, does not have a good model fit and fails the construct reliability and validity tests. The current study used confirmatory factor analysis to identify a second-order CFA model with Flight Risk as the second-order construct and general flight risk, high risk, and altitude risk as the first-order constructs. The resulting new 13-item scale demonstrated good fit, construct validity, and reliability.

Risk remains a constant and ever-present threat, and this is particularly true in the aviation industry where events, such as weather, can change quickly. A tool such as the one revalidated in this study to measure pilot’s self-assessed risk perception can be valuable to the aviation research community.

There is also value added through having a valid instrument which can be used in the aviation field to measure risk perception. The process to develop a new instrument is rigorous and requires a large sample of participants, which may often be difficult for researchers to build before conducting the main study, especially within the aviation domain. Within the aviation field, it can frequently be a challenge to gather participants to complete studies, and having to collect hundreds of participants to validate an instrument before completing the actual proposed study can often be a non-starter. This challenge may also result in researchers using scales which have not been properly validated, which could threaten the overall findings of the study. Therefore, there is clearly value in using a validated scale. This scale can help fill that void as the psychometrics of this study demonstrate a valid scale to measure self-assessed risk perception in pilots.

Lastly, this updated scale balances the length and usefulness of the scale, resulting in a parsimonious tool. While longer scales typically have good validity, their length can sometimes make them rather impractical to use conveniently in research. This could result in a highly valid instrument, but one which is somewhat restrictive for use in actual studies. The current 13-item scale attempts to balance the length and usefulness while also maintaining high levels of validity (Gosling, Rentfrow, & Swann Jr., 2003). This Flight Risk Perception Scale can be completed in a few minutes, which makes it ideal if researchers want to investigate multiple measures along with risk or if they want to use the risk perception measure as part of a pre-test/post-test design (Robins, Hendin, & Trzesniewski, 2001).

Practical Applications

The current study produced a valid 13-item Flight Risk Perception Scale for the use of self-assessment of risk perception for pilots. The primary outcome of this study is a valid instrument that may be used by other researchers in the investigation of their proposed research topics on flight risk perceptions by pilots. It is hoped that other researchers in aviation will use this tool to further our understanding in areas of inquiry such as pilot risk taking, risk perception, decision-making, and judgment which remains a major focal point of aviation research. This

scale is updated from the original 26-item scale developed by Hunter (2002, 2006). This shortened instrument has been updated, and the additional item reduction has resulted in a more parsimonious scale, which still maintains a high level of validity. The advantages of this updated scale are the ability to be used efficiently in research studies where researchers wish to have multiple instruments or use the risk perception scale as part of a pre-test/post-test design.

Limitations

Several limitations constrained the current study. First, the use of a convenience sample collected through AOPA ASI limits the generalizability of the findings to those types of individuals who are members and subscribers to that safety organization. Additionally, the scale is limited to the experience level of participants represented within this current sample. Further research is needed to replicate the results of this research to verify the new instrument remains valid within a broader population of pilots. The study also had a relatively low response rate of around 5%, which could result in non-response bias in the findings. More active forms of data collection could help increase the response rate of the study. The construct of High Risk demonstrated low validity and reliability after the data analysis. Future studies should revisit the wording of these questions as adjustments to the phraseology may help improve the validity of this construct. Finally, the cross-sectional nature of the data collection resulted in the data for the study being collected over a short three-week window. Therefore, the findings provide data from one point in time. Future research could use the scale in a longitudinal type study to monitor rating levels over a period of time.

Conclusions

The purpose of the current study was to assess the Risk Perception – Self scale. This 26-item scale was designed as a five-factor scale to measure the risk perception as self-assessed by pilots. A review of the original instrument suggested threats to its validity, and therefore, the current study used confirmatory factor analysis, which produced a restructured second-order factor structured scale. Flight risk was shown to be the second-order construct and the three first-order constructs were general flight risk, high risk, and altitude risk. The new scale was able to be reduced to 13-items instead of the original 26-item assessment, and the findings indicated good construct validity and construct reliability for this revised model. The shorter scale may be useful for researchers who wish to measure self-assessed risk perception in pilots, and it could be helpful to studies with multiple measures, such as a pre-test/post-test or in longitudinal studies. The new model with good construct validity could be used in future risk perception studies to evaluate the relationship between Flight Risk with other factors.

Acknowledgments

The authors would like to thank the Aircraft Owner's and Pilot's Association's (AOPA) Air Safety Institute (ASI) for their support of this research project and assistance in data collection.

References

- Aircraft Owners and Pilot's Association (AOPA). (2018). 27th Joseph T. Nall Report: General Aviation Accidents in 2015. AOPA Air Safety Institute: Frederick, Maryland.
- Benet-Martínez, V. & John, O. P. (1998). Los Cinco Grandes! Across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, 75, 729–750.
- Blunch, N. J. (2013). *Introduction to structural equation modeling using IBM SPSS Statistics and AMOS 2nd Edition*. Thousand Oaks, CA: Sage.
- Burisch, M. (1984). You don't always get what you pay for: Measuring depression with short and simple versus long and sophisticated scales. *Journal of Research in Personality*, 18, 81–98.
- Burisch, M. (1997). Test length and validity revisited. *European Journal of Personality*, 11, 303–315.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. 2nd Edition. Routledge Taylor & Francis Group.
- Drinkwater, J. L. & Molesworth, B. R. C. (2010). Pilot see, pilot do: Examining the predictors of pilots' risk management behaviour. *Safety Science*, 48, 1445-1451.
- Federal Aviation Administration. (2009). *Risk management handbook (FAA-H-8083-2)*. Washington, D.C.: U.S. Department of Transportation.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18(1), 39-50.
- Gosling, S. D., Rentfrow, P. J., & Swann Jr., W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504-528.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (7th edition)*. Upper Saddle River: Prentice Hall.
- Hansson, S. O. (2005). Seven myths of risk. *Risk Management: An International Journal*, 7(2), 7–17.
- Hunter, D. R. (2002). *Risk perception and risk tolerance in aircraft pilots* (Rep. No. DOT/FAA/AM-02/17). Washington, D.C.: Federal Aviation Administration
- Hunter, D. R. (2005). Measurement of hazardous attitudes among pilots. *The International Journal of Aviation Psychology*, 15(1), 23-43.

- Hunter, D. R. (2006). Risk perception among general aviation pilots. *The International Journal of Aviation Psychology*, 16(2), 135-144.
- Hunter, D. R., Martinussen, M., Wiggins, M., & O'Hare, D. (2011). Situational and personal characteristics associated with adverse weather. *Accident Analysis and Prevention*, 43, 176-186.
- Jensen, R. S., & Benel, R. A. (1977). *Judgement evaluation and instruction in civil pilot training*. (Rep. No. FAA-RD-78-24). Washington, DC: Federal Aviation Administration.
- Ji, M., Yang, C., Li, Y., Xu, Q. & He, R. (2018). The influence of trait mindfulness on incident among Chinese airline pilots: The role of risk perception and flight experience. *Journal of Safety Research*, 66, 161-168.
- Ji, M., You, M., Lan, J., & Yang, S. (2011). The impact of risk tolerance, risk perception and hazardous attitude on safety operation among airline pilots in China. *Safety Science*, 49(10), 1412-1420.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102–138). New York: Guilford Press.
- Joseph, C. & Reddy, S. (2013). Risk perception and safety attitudes in Indian army aviators. *The International Journal of Aviation Psychology*, 23(1), 49-62.
- Knecht, W. R. & Frazier, E. (2015). *Pilot's risk perception and risk tolerance using graphical risk-proxy gradients*. Report No. DOT/FAA/AM-15/9. FAA Office of Aerospace Medicine, Washington, D.C.
- Merriam-Webster (2019). Definition of risk. Retrieved on May 5, 2019 from: <https://www.merriam-webster.com/dictionary/risk>
- Molesworth, B. R. C. & Chang, B. (2009). Predicting pilots' risk-taking behavior through an implicit association test. *Human Factors*, 51(6), 845-857.
- Molesworth, B., Wiggins, M., & O'Hare, D. (2006). Improving pilots' risk assessment skills in low-flying operations: The role of feedback and experience. *Accident Analysis & Prevention*, 38, 954-960.
- Nunnally, J.C. (1978) *Psychometric theory* (2nd edition). New York: McGraw-Hill.
- O'Hare, D. (1990). Pilots' perceptions of risks and hazards in general aviation. *Aviation, Space, and Environmental Medicine*, 61, 599-603.
- Orasanu, J., Fischer, U., & Davison, J. (2002). *Risk perception: A critical element of aviation safety*. Paper presented at the 15th Triennial World Congress, Barcelona, Spain.

- Oster, C. V., Strong, J. S., & Zom, C. K. (2013). Analyzing aviation safety: Problems, challenges, opportunities. *Research in Transportation Economics*, 43(1), 148–164.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem scale. *Personality and Social Psychology Bulletin*, 27, 151–161.
- Soper, D. (2019, July). A-priori sample size calculator for structural equation models. Retrived on July 17, 2019 from: <https://www.danielsoper.com/statcalc/calculator.aspx?id=89>
- Wiggins, M. W., Hunter, D. R., O'Hare, D., & Martinussen, M. (2012). Characteristics of pilots who report deliberate versus inadvertent visual flight into instrument meteorological conditions. *Safety Science*, 50, 472-477.
- Wilson, J. E. & Joye, S. W. (2017). *Research methods and statistics: An integrated approach*. Thousand Oaks: Sage.
- You, X., Ji, M., & Han, H. (2013). The effects of risk perception and flight experience on airline pilots' locus of control with regard to safety operation behaviours. *Accident Analysis and Prevention*, 57, 131-139.

Appendix A – Flight Risk Perception Scale (FRPS) – The Modified Risk Perception Scale

Instructions: Please rate the level of risk present in the situation, if YOU were to experience the situation tomorrow. Responses are provided on a scale from 1 (Low Risk) to 9 (High Risk).

Flight Risk Perception Scale

General Flight Risk

1. During the daytime, fly from your local airport to another airport about 150 miles away, in clear weather, in a well-maintained aircraft.
2. Make a two-hour cross-country flight with friends, after checking your weight and balance.
3. At night, take a cross-country flight in which you land with over an hour of fuel remaining.
4. During the daytime, take a cross-country flight in which you land with over an hour of fuel remaining.
5. At night, fly from your local airport to another airport about 150 miles away, in clear weather, in a well-maintained aircraft.

High Risk

6. Fly in clear air at 6,500 feet between two thunderstorms about 25 miles apart.
7. Make a traffic pattern so that you end up turning for final with about a 45 degree bank.
8. Make a two-hour cross country flight with friends, without checking your weight and balance.

Altitude Risk

9. Fly across a large lake or inlet at 500 feet above ground level.
10. Take a two-hour sightseeing flight over an area of wooded valleys and hills, at 3,000 above ground level.
11. Fly across a large lake or inlet at 1,500 feet above ground level.
12. Take a two-hour sightseeing flight over an area of wooded valleys and hills, at 1,000 above ground level.
13. Fly across a large lake or inlet at 3,500 feet above ground level.