

## **VALIDATING THE COMMERCIAL AVIATION SAFETY SURVEY IN CHINESE CULTURAL CONTEXT: A CASE STUDY**

Yongjuan Li<sup>1</sup>, Terry L. von Thaden<sup>2</sup>, Feng Li<sup>1</sup>, Li Jiang<sup>1</sup>, Lei Dong<sup>1</sup>

<sup>1</sup>Institute of Psychology, Chinese Academy of Sciences, Beijing

<sup>2</sup>Human Factors Division, University of Illinois at Urbana-Champaign

The aim of the present study was to validate the Commercial Aviation Safety Survey (CASS), developed at the University of Illinois, in the context of Chinese culture. Specifically two goals of this study included translating/developing a Chinese version of CASS (CASSCN) based on the English version, which would prove comparable with CASS, and exploring the specific cultural and safety contexts of CASSCN. Initial results for the four dimensions measured demonstrate adequate reliability, and the elementary confirmative factor analysis demonstrates that the survey can be dependably applied in the Chinese context. The general results demonstrated relatively low perceptions of organizational safety culture, with the exception of the instructor/trainer subscale. A particular focus of the study examined confidential reporting systems (CRS) comparing the airline's current reporting systems (RS1), the pilots' attitude to a possible CRS which would be controlled by the airlines they serve (RS2), or by a third party (RS3). Results illustrate that pilots' attitude to RS1 were significantly lower than the other two reporting system options. With RS2, Captains showed more reservation using the system than Co-pilots; with RS3, Captains and Co-pilots alike were in agreement that they would report information regarding their own performance and that of other pilots. This data represents a preliminary study concluding the CASS can be successfully applied in the Chinese context, but attention should be paid to measurement equivalence before the Chinese version is widely used.

### Introduction

Research interest in safety culture is generally believed to have begun with the Chernobyl nuclear disaster in 1986, when both the International Atomic Energy Agency and the Organization for Economic Cooperation and Development cited "poor safety culture" as a contributing factor in the accident enquiry (Mearns & Flin, 1999; Cox & Flin, 1998; Pidgeon, 1998). The commercial aviation industry has displayed a strong interest in organizational safety culture, in part precipitated by the role attributed to a lack of corporate safety culture in the crash of Continental Express Flight 2574 in 1991 (NTSB, 1992). The resulting interest in organizational safety culture has given rise to widespread definitions and measures of the construct. Wiegmann, Zhang, von Thaden, Sharma, & Mitchell (2002) reviewed numerous definitions of safety culture to identify the commonalities among them. They concluded that safety culture can be defined as "the enduring value and priority placed on worker and public safety by everyone in every group at every level of an organization" (p. 9). They noted that (1) safety culture was typically defined as a group-level construct, based on shared values; (2) safety culture was especially related to management and supervisory systems, but (3) everyone in the organization was a part of the culture; (4) safety culture was expected to influence the behavior of members of the organization; (5) perceived

contingencies or rewards for safe or unsafe behavior were critical components of safety culture; (6) safety culture and learning culture were closely connected; and (7) safety culture endures and persists over time. Investigating the construct of safety culture is a precondition to study the effect of safety culture on safety performance. Indeed, numerous organizational indicators that reflect an organizational safety culture have been proposed, with some estimates ranging from as few as two to as many as 19 (Flin et al., 2000). Wiegmann et al. (2002; 2004) reviewed the organizational safety culture literature across a number of industries and originally identified five core indicators of an organization's safety culture, including: organizational commitment; managerial involvement, employee empowerment, accountability, and reporting system. The initial validation of the five indicators found that there were high correlations between two of the five dimensions indicating the subscales might measure the same construct (Wiegmann et al., 2003). After further testing, Gibbons, von Thaden and Wiegmann (2004) validated the correlation structure of the safety survey resulting in an improved version of Commercial Aviation Safety Survey (CASS) (Gibbons, et al., 2005). While the reliability and validity of CASS have already been tested at airlines in US, it remained to be seen whether the CASS could be applied in a completely different culture, such as China.

### Method

Safety Across High-Consequence Industries Conference, St. Louis, MO.

March 13-15, 2007

## Translation work

The first aim of our study was to develop the Chinese version of the CASS based on the English version developed at the University of Illinois. To make it sure Chinese version expressed the precise meaning of the English version, we took two steps to finalize the Chinese version:

### *Step one: Translation/Back-translation:*

1) One bilingual researcher, native Chinese speaker, familiar with safety culture, translated the English version into Chinese. Two additional Chinese researchers reviewed and refined it.

2) Another bilingual native Chinese speaker, who had not seen the original English version of the survey, translated the Chinese version back into English. After this translation, another researcher reviewed and refined the English version.

3) *Two members of the program, one a native English speaker, the other a native Chinese speaker, compared the original English version and the back-translated English version (from the Chinese), identifying problem or idiomatic phraseology.*

4) The first two reviewers refined the Chinese version remedying the problems identified in step 3).

### *Step two: Aviation terminology transformation*

Since some of the key aviation terms in the original study were not the same expressions used by the local Chinese aviation practitioners, we transformed the terms in the translated survey to reflect the Chinese aviation vernacular. The process included:

1) *10 native Chinese pilots filled out the survey critically pointing out terms they didn't understand very well or recognize as aviation terms.*

2) The researcher in charge of the English translation modified the terms identified as ambiguous or non-typical Chinese aviation expressions.

3) Back-translation of the refined survey revealed the problematic translation issues had been solved subsequently finalizing the Chinese version of CASS (CASSCN) and completing the conversion.

## Research Process

The research process included three parts. To begin, the CASSCN survey included the four major factors identified by Gibbons, von Thaden and Wiegmann (2005) of Organizational Commitment (OC), Operations Personnel (OP), Formal Safety System (FS), and Informal Safety System (IS) and the

subfactors associated with each. Next, given that Chinese commercial aviation does not have a formal confidential reporting system (CRS), we added two subfactors which were from the Reporting system subfactor of original CASS to the survey to understand the pilots' attitude to a possible CRS (the first item was deleted since it was not relevant to the subjects at hand). One subfactor was to ask pilots' attitude toward a possible CRS developed and maintained by the airlines they serve, and the other was to ask their attitude toward a possible CRS developed and maintained by a third party (e.g. Institute of Psychology, Chinese Academy of Science). The items retained from the English CASS pertaining to reporting system were used to evaluate the current reporting system, which is not anonymous and confidential. Finally, demographic items were established according to the Chinese aviation system.

*Procedure.* Participation was voluntary. The surveys were distributed randomly to airline employees who attend to the routine administrative meeting (once a week). An envelope that could be sealed was distributed with the survey. One blank box was put in the lobby of the airline's building, in which those who finished the survey could deposit the sealed, anonymous survey. After a period of four weeks, the surveys were collected.

*Sample.* A total of 430 surveys were distributed at Airline A, with 190 surveys returned, for response rate of 44%. Some of the surveys returned were unusable (based on the criterion of the amount of unanswered questions), resulting in a valid response rate of 43%.

The distribution according to position of respondents was as follows: Co-pilot (38.4%), Captain (32.4%), Instructor (9.7%), Trainee (8.6%), Inspector (1.1%), Manager-not-flying (1.1%), and missing information (8.6%). Age distribution of the respondents was as follows: 20-30yrs (33.5%), 31-40yrs (44.3%), 41-50yrs (8.6%), 51-60yrs (2.7%), 60+yrs (0.5%), with 10.3% missing this data. Distribution of the number of years respondents spent working for their current airline were as follows: <1yr (7%), 1-5yrs (15.7%), 6-10yrs (38.9%), 11-15yrs (18.4%), 16-20yrs (2.7%), 20yrs+ (4.9%), and missing this information, 12.4%.

## Results

### Factor Reliability

The CASSCN displayed suitable reliability results for the four dimensions measured (Table 1). However, we still needed to validate the model of safety culture

in the Chinese context. Since the model has already been validated using data from US airlines, we performed a confirmative factor analysis only, which was carried out using Amos 4.0 (Arbuckle & Wothke, 1999). Evaluation of the models was based on the following statistics commonly used in prior research: The Normed Fit Index (NFI), Comparative Fit Index (CFI) and the Tucker-Lewis index (TLI), with values greater than .95 indicating good fit and values less than .95 but greater than .90 indicating adequate fit; the Root Mean Square Error of Approximation (RMSEA) and its 90% Confidence Interval (CI), with values of RMSEA smaller than .05 indicating good fit and values greater than .05 but less than .08 indicating adequate fit. For each model, the chi-square statistic ( $\chi^2$ ) and its degrees of freedom (*df*) were conventionally reported but were not considered for evaluation of individual models, due to its sensitivity to sample size.

**Table 1.** Reliability of the four-dimension CASSCN

Subscale	# of Items	Alpha
Formal Safety Program	10	.84
Informal Aspects of Safety	14	.81
Operational Personnel	8	.84
Organizational Commitment	14	.80

According to the criteria above, we tested the subscales with Amos4.0 first. The corresponding results of the subscale evaluation appear in Table 2. Due to a printing error, two subscales (response and feedback, and chief pilots) were accidentally missed, thus the CFA contains results for only 10 subscales. The results show that only one subscale, instructors/trainers, didn't appear to fit the model very well. The item "Instructors/trainers teach shortcuts and ways to get around safety requirements" had low or negative correlation with the other three questions in the subscale. The test result with the problem item deleted is shown in Table 2.

The four factor model fit the data from Airline A to some extent:  $\chi^2 (29, N = 185) = 50.23, p=.01, NFI = .99, TLI = .99, CFI = .99, RMSEA = .06$ . The four factor model without the item in the instructors/trainers subscale also fit the data:  $\chi^2 (29, N = 185) = 62.55, p=.00, NFI=.99, TLI=.99, CFI = .99, RMSEA = .08$ . The three factor model without the operation personnel dimension also fit the data:  $\chi^2 (17, N = 185) = 19.75, p=.29, NFI=.99, TLI=.99, CFI = .99, RMSEA = .03$ . According to the CFA results, the model fit the data adequately, so we can indeed

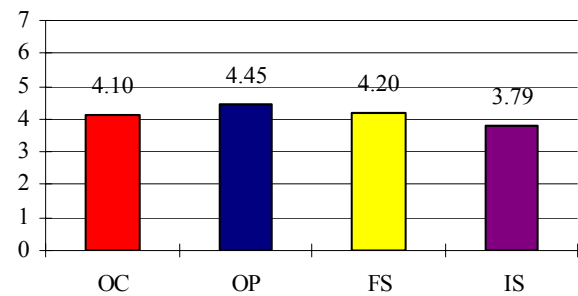
indicate that the CASSCN structure may be applied in the Chinese context.

**Table 2.** Results of CFA for CASSCN.

Subscale	$\chi^2$	P	NFI	TLI	CFI	RMSEA
Reporting System	39.03	.00	.98	.95	.98	.19
Safety Personnel	24.4	.00	.99	.97	.99	.15
Accountability	8	.00	.99	.99	.99	.10
Authority	5.30	.07	.99	.99	.99	.07
Professionalism	12.5	.03	.99	.98	.99	.12
Dispatchers	19.1	.00	.99	.98	.99	.12
Instructors	25.4	.00	.98	.92	.99	.25
Safety Values	0		1.0		1.0	1.32
Safety Fundamental	22.3	.00	.99	.97	.99	.14
Beyond Compliance	6	.00	.99	.98	.99	.14
	22.0	.00	.99	.98	.99	.14
	1	.01	.99	.98	.99	.13
	8.30	.02	.99	.98	.99	.13

Descriptive results

The following data analysis was performed at Airline A. The scale midpoint is 4. Scores are considered positive when above this value and negative when below. The mean score on the four dimensions measured using CASSCN revealed a rather mediocre to poor safety culture at Airline A (Figure 1), with

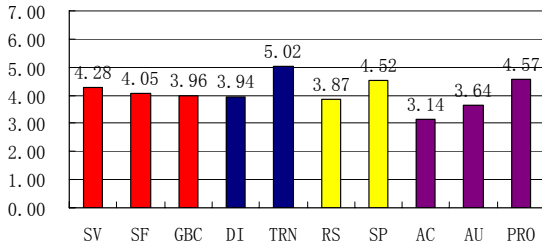


Informal Aspects of Safety rating lowest at 3.79.

**Figure 1.** Mean scores for the major factors of CASSCN for Airline A.

Further analysis of the airline scores at the subfactor level revealed specific areas that scored low, which included Accountability (AC, 3.16), Pilot's Authority (AU, 3.64), Reporting System (RS, 3.87); some areas scored mediocre, which included Dispatch (DI, 3.94), Safety Fundamentals (SF, 4.05) and Going Beyond

Compliance (GBC, 3.96); Some areas scored relatively high, which included Safety Values (SV, 4.28), Instructors/ Trainers (TRN, 5.02) and Safety Personnel (SP, 4.52).



**Figure 2.** Mean scores for Airline A on the subfactors measured with CASSCN.

*Comparative analysis in reporting system.* As different from U.S civil aviation, China only has non-confidential reporting system, though currently the Civil Aviation Safety Research Institute and the safety office of the regulatory commission (CAAC) are attempting to establish a local CRS like that of the Aviation Safety Reporting System in the US. Given this, our interest was to determine if it is possible and acceptable to the potential users of the system to develop a CRS in China, and which party could possibly run a successful CRS, we developed items in the CASSCN to identify this. In the following results, RS1 refers to the airline’s current reporting systems, RS2 refers to a possible CRS which would be controlled by the airline the pilots’ serve, and RS3 refers to a CRS controlled by a third party. Results illustrate that pilots’ attitude to RS1 were significantly lower than the other two reporting system options (Table 3).

**Table 3.** Comparison of reporting systems (N=171).

Reporting system	M.	S.D.
RS1	3.97	1.10
RS2	4.38	1.05
RS3	4.51	1.02

The repeated one-factorial F test showed that  $F(2,342)=23.48, p<.001$ , revealing significant difference among pilots’ perception of the different reporting systems. Post-Hoc tests demonstrated that the pilots’ perception of the CRS was directly associated with whomever was in charge of the system. (Table 4).

**Table 4.** Post - Hoc comparison among the different reporting systems suggested (N=171).

Reporting System	M.	S.E.
RS1 RS2	-0.40*	0.07
RS1 RS3	-0.51*	0.09
RS 2 RS3	-0.11	0.08

Note: \*The mean difference was significant at the .05 level.

Comparison between self -reported information and reporting others’ information. *Two items in the survey expressed the extent to which pilots would most likely report their own information (mistakes) and others’ information. The results of these items in the three reporting systems are shown in Table 5.*

**Table 5.** Comparison between self report and reporting others’ information (N=168).

Reporting System	Item	M.	S.D
RS1	Reporting others	3.74	1.61
	Reporting Self	4.47	1.74
RS2	Reporting others	4.18	1.59
	Reporting Self	4.73	1.61
RS3	Reporting others	4.60	1.49
	Reporting Self	4.68	1.58

Two-factors in the repeated F test showed significance. Factor 1 related to the different reporting system: RS1, RS2, and RS3; and Factor 2 related to the perception of reporting information on self vs. others. The main effect of Factor1 revealed significance,  $F(2,342)=12.90, P<.001$ ; the main effect of Factor2 also revealed significance,  $F(1,171)=25.06, p<.001$ ; with the interaction effect revealing significance,  $F(2,342)=12.74, P<.001$ . The simple effect showed that with RS1, the pilots’ interest in reporting others is significantly lower than self-reporting,  $F(1,171)=31.80, P<.001$ . With RS2, the pilots’ indicate they would report others is significantly lower than self-reporting,  $F(1,171)=18.81, P<.001$ . There was no significant difference in the pilots’ will to report others and self with RS3,  $F(1,171)=.46, P=.501$ .

*Comparison among different positions.* We compared the results to identify any significant difference between Co-pilots (n=71) and Captains (n=60). The

results are shown in Table 6. Two-factor mixed F test showed that, the main effect of job title was significant,  $F(1,122)=6.73$ ,  $p=.01$ . The interactive effect between the two factors was marginally significant,  $F(2, 244)=2.78$ ,  $p=0.06$ . Simple effect analysis demonstrated that with RS1, the Co-pilots' evaluation was significantly more positive than the Captains',  $F(1,122)=6.38$ ,  $p=.01$ . With RS2, Captains showed more reservation using the system than Co-pilots;  $F(1,122)=7.99$ ,  $p=0.05$ . With RS3, Captains and Co-pilots alike were in agreement that they would report information regarding their own performance and that of other pilots,  $F(1,122)=0.50$ ,  $p=0.483$ . There were no significant differences in regard to the reporting systems items revealed among the pilots according to age.

**Table 6.** Difference between Co-pilot and Captain.

Reporting System	Title	M.	S.D.
RS1	Co-pilot	4.16	1.02
	Captain	3.61	1.06
RS2	Co-pilot	4.54	0.93
	Captain	4.01	1.11
RS3	Co-pilot	4.58	0.84
	Captain	4.39	1.04

#### Discussion

##### Validation of the CASS in Chinese context

The CFA demonstrated that most of the subfactor indicators fit the data, revealing that the CASS model could be applied in Chinese contexts, while there are still some aspects in need of refinement. One problematic subfactor relating to instructors/ trainers didn't fit the data well. Another aspect in need of attention is the measurement equivalence. The CASS model was developed and validated according to US cultural norms, so care should be exercised in regard to its application to another culture before the measurement equivalence can be tested.

##### Basic results of Chinese airline

According to the results shown in Figures 1 and 2, generally, the scores of Chinese airline A were low. The pilots evaluated their professionalism and their instructor/trainers highly. Scores relating to safety fundamentals and safety personnel were also above neutral. Pilots evaluated management (accountability, safety value, and going beyond violence) and supporting staff (dispatchers), as well as the participation of pilots to safety issues (reporting system and authority) relatively low.

Chinese civil aviation has rapidly developed in recent years, and has experienced typical representations of Chinese reformation, changing from a central economy to a market orientation in the past 20 years<sup>70</sup>. The State-owned companies began reformation at the beginning of the 21st century. During this time, three large airlines merged with some of the local airlines. Privately owned airlines have been permitted to coordinate operations in the past few years, and foreign-investment airlines will enter China in the near future. The macro- and micro-environment has been turbulent, and the public media frequently features news stories about the conflict between pilots and management, between passengers and airlines, and the competition among older airlines and newer airlines for pricing and recruits, et al<sup>71</sup>. These major change upheavals increase the potential risks to flight safety. On the one hand, management at most Chinese airlines understands that training programs offered in China are not as high-quality as programs offered elsewhere, so Chinese pilots are sent outside the country to receive their advanced technical training. While management takes positive steps to correct one weakness, on the other hand, they still must contend with the relatively slow changes in organizational morals and the supporting systems in the "real" world Chinese aviation system. The traditional Confucian values that emphasize the class system, obedience, the doctrine of mean, and *renqing* (kindness/reciprocity) still influence contemporary Chinese behaviors (Tsui, et al., 2004). The requirements of flight safety: cooperation, coordination, communication, cross-supervision, et al., challenge the traditional cultural values of the pilots to some extent. These factors must be considered in regard to their effect on safety culture.

##### Reporting system

Due to the difference between US airlines and Chinese airlines, we analyzed Reporting System (RS) separately. Two items in RS were below neutral. One item, "Pilots can report safety discrepancies without fear of negative repercussions," may make sense if we consider the score in relation to the penalty system in Chinese aviation. Comments revealed that the current penalty system was the reason pilots hesitate to report events. The other low scoring item, "Pilots are willing to report information regarding marginal performance or unsafe actions of other

70

<http://finance.sina.com.cn/g/20021012/0844265300.html>

71 <http://news.sina.com.cn/c/2006-10-25/011311322674.shtml>

pilots,” also makes sense in the context that Chinese civil aviation does not have a systematic confidential reporting system. If a pilot reports another, the participant and other peers could easily know this information, which burdens the reporter. On the other hand, the Confucian value of *renqing* and keeping the in-group harmony is also an important value in China (Oyserman, Coon, & Kemmelmeier, 2002). Nisbett et al. (2001) has similarly described Eastern ontologies and epistemologies in terms of holism, in which greater attention is paid to the perceptual field, the situational context, and relationships among objects and events in the environment. Conversely, Western epistemologies, rooted in Aristotelian formal logic (Peng & Nisbett, 1999), emphasize order and constancy in the world, immutable laws and truths, and decontextualized facts and ideas. Given this, it may explain why Chinese pilots may be reticent to report others' information relating to safety.

One interesting finding arises from the significant difference between self-reporting and reporting others in RS2. The results reflect pilots are apprehensive giving this information to an airline CRS system. The difference ceases when the CRS is potentially charged by a third party. It appears that third-party confidence could motivate pilots to offer more information, especially in the absence of post-report pressure, without the cultural concern that the report may break airline group harmony. Comparing the different positions reveals that the younger generation of pilots prefers, and is open to, the reporting system. This also reflects the transition and reformation of Chinese culture.

### Conclusion

Results of this study indicate the CASS can be applied in the Chinese context. Results reveal that pilots at the test airline have a negative perception of management, supporting departments and staff, while perceptions of their own professionalism and that of their instructors/trainers is positive. The typical professional culture of pilots is pride in their job and themselves (Helmreich, 1997). A confidential reporting system controlled by a third party could greatly aid pilots to report safety issues without fear of negative repercussions and moral trouble. Our study to explore the structure of Chinese aviation's safety culture and the measurement equivalent requires more research. We have taken steps to understand the pilots' perception of safety culture, which is only part of our overall research goal. Further investigation is needed regarding the perception of those in other positions to arrive at a more general conclusion.

### Acknowledgement

This work was supported in part by a grant from the Federal Aviation Administration under Award No. DTFA 01-G-015, Jennelle Derrickson, technical monitor, and in part by a grant from National Natural Science Foundation of China (NSFC, 70401018). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the FAA or the NSFC.

### References

- Arbuckle, J. L., & Wothke, W. (1999). *Amos 4.0 User's Guide*. Chicago, IL: Smallwaters Corporation.
- Cox, S., & Flin, R. (1998). Safety culture: Philosopher's stone or man of straw? *Work & Stress*, 12(3), 189-201.
- Flin, R., Mearns, K., O'Connor, P., & Bryden, R. (2000). Measuring safety climate: Identifying the common features. *Safety Science*, 34, 177-192.
- Gibbons, A. M., von Thaden, T. L., & Wiegmann, D. A. (2005). Development and validation of a survey to assess safety culture in airline maintenance operations. *Proceedings of the 13th International Symposium on Aviation Psychology*.
- Gibbons, A. M., von Thaden, T. L., & Wiegmann, D. A. (2004). Exploration of the Correlation Structure of a Survey for Evaluating Airline Safety Culture. *University of Illinois Human Factors Division Technical Report AHFD-04-06/FAA-04-3*.
- Helmreich, R. (1997). Managing human error in aviation. *Scientific American*, 62-67.
- Mearns, K. J., & Flin, R. (1999). Assessing the state of organizational safety--Culture or climate? *Current Psychology: Developmental, Learning, Personality, Social*, 18(1), 5-17.
- Nisbett, R., Peng, K., Choi, I., & Norenzanan, A. (2001). Culture and system of thoughts: Holistic versus analytic cognition. *Psychological Review*, 108, 291-310.
- Oyserman, D., Coon, H., Kemmelmeier, M. (2002). Rethinking Individualism and Collectivism: *Evaluation of Theoretical Assumptions and Meta-Analyses*. *Psychological bulletin*, 128(1), 3-72.

Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, 54, 741-754.

Pidgeon, N. (1998). Safety culture: Key theoretical issues. *Work & Stress*, 12(3), 202-216.

Tsui, A.S., Wang, H., Xin, K., Zhang L., Fu, P.P. (2004). Let a thousand flowers bloom – Variation of leadership styles among Chinese CEOs. *Organizational Dynamics*, 33 (1), 5–20.

Wiegmann, D. A., Zhang, H., von Thaden, T. L., Sharma, G., & Mitchell, A. A. (2002). A Synthesis of Safety Culture and Safety Climate Research. *University of Illinois Aviation Research Lab Technical Report ARL-02-03/FAA-02-2*.

Wiegmann, D. A., von Thaden, T. L., Mitchell, A. A., Sharma, G., & Zhang, H. (2003) Development and Initial Validation of a Safety Culture Survey for Commercial Aviation. *University of Illinois Human Factors Division Technical Report AHFD-03-03/FAA-03-1*.