International Journal of Aviation, Aeronautics, and Aerospace

Volume 8 | Issue 4 Article 4

2021

Measuring Safety Culture: Qualitative and Quantitative Means of Measuring Safety Culture for Safety Management System Optimization

David R. Zubowski
Embry-Riddle Aeronautical University, dzubowski@gmail.com

Safety Management Systems (SMS) became a requirement for United States air carriers in March of 2015 when the Federal Aviation Administration (FAA) established the requirements and codified them in the Code of Federal Regulations under Part 5 titled *Safety Management Sytems* (FAR, 2015). An organizational SMS, according to the FAA (2016a), should show a means of decision making, management capability before a system failure, risk controls through safety assurance, knowledge and data sharing with the FAA, and a sound safety culture. In the implementation and evaluation of an SMS, guidance for the understanding of a sound safety culture has shown to be a challenge. While there are many theories and studies on the development and assessment of safety culture, the FAA oversight and guidance on assessing safety culture was determined by the United States Department of Transportation Office of the Inspector General (USDOT OIG) (2020) to be insufficient.

Background

Before SMS was initially introduced by the International Civil Aviation Organization (ICAO) and subsequently introduced to United States carriers by the FAA, aviation safety had gone through several iterations. Initially, the FAA had a dual focus of both safety and the economic protection of the industry. As the mechanical and technical issues of commercial flight were resolved, the industry moved initially to Crew Resource Management (CRM) to improve safety from the human factors perspective. CRM eventually evolved into the Threat and Error Management program used alongside Line Operation Safety Audits. However, those programs still have disconnects and occasionally fail to mitigate risks. This led to the introduction and eventual mandate of SMS (Petitt, 2017).

The ICAO (n.d.) created and released multiple editions of their Safety Management Manual, the most recent at the time of submission is the 4th edition, to provide guidance material to State and Regulatory bodies on safety management principles and concepts. The FAA (2016a) endorsed these principles and concepts by releasing their version of SMS with the goal of integrating modern risk management and safety assurance into the aviation industry.

Three regulatory ideas presented by the FAA (2014) seem to fall short of the intent of SMS implementation as determined by the USDOT OIG (2020) investigation. The FAA (2014) would continue to be responsible for the enforcement of regulation, would continue to be a technical workforce, and SMS implementation would not replace any FAA oversight. However, the USDOT OIG received complaints about the lack of FAA oversight counter to the message presented by the FAA. The findings by USDOT OIG may reinforce findings from Gill and Shergill (2004) where regulatory oversight was secondary in the role in aviation safety, behind luck. Surprisingly, even with the evolution of CRM and the introduction of SMS, "…'pilots' perceive luck and safety to be the most important factor in aviation safety" (Gill & Shergill, 2004, p. 237). While focused on aviation

in New Zealand, one of the major findings from this study is the perception that employers regard an individual's responsibility for safety as more important than SMS or safety culture development. These perceptions are contrary to the stance taken by the Civil Aviation Authority of New Zealand where safety is an imperitave. One area of further study identified is the attitude of aviation safety being dependent on luck or if the attitude exists through lack of confidence in the SMS (Gill & Shergill, 2004).

Statement of the Problem

The USDOT OIG (2020) received a hotline complaint about the oversight of a major U.S. airline SMS by the FAA in early 2018. One of the concerns identified by the USDOT OIG was the lack of appropriate guidance for FAA inspectors to assess safety culture as part of the oversight of SMS implementation. Through the USDOT OIG investigation into the airine and the FAA, the safety culture of said airline was identified as a concern, but the lack of guidance resulted in safety culture not being a factor in the evaluation of the airline's risk identification and mitigation processes. Recent studies from academia showed the viability of quantitative and qualitative means to assess an organizational safety culture, however, it is unclear if the FAA will use the results of those studies to provide safety culture assessment guidance to inspectors at all levels.

Research Question

What are the benefits, limitations, and assumptions of published quantitative and qualitative means of measuring an organizational safety culture and how could they be applied for the effective FAA oversight of SMS implementation by the aviation industry?

Delimitations and Assumptions

A delimitation of this study was the restriction of sources to peer reviewed articles, FAA website pages, and the ICAO website pages inclusive of links provided by the FAA and ICAO to external sites. A second delimitation was the restriction of peer reviewed article searches that included the term "safety culture" in the search. A third delimitation was the restriction of using a case study from USDOT OIG (2020) findings to establish a scenario that focused on safety culture evaluation.

One assumption made in this study was the lack of bias in the USDOT OIG (2020) findings on the lack of FAA oversight for the the airline's SMS implementation. A second assumption was that enough relevant peer reviewed studies had been published that investigated quantitative and qualitative means of measuring safety culture. A third assumption was that academic studies produced recommendations that could be implemented by the FAA and the Part 121 carriers without process or cost barriers that would make such implementation prohibitive.

Methodology

The method used in this study is a case study on the report from the DOT USOIG (2018) and the deficiencies identified by the USDOT OIG on the ineffective oversight of an airline's SMS by the FAA. The research was conducted by searching for peer-reviewed articles relating to safety culture, SMS, and safety culture evaluation. The search terms for the peer-reviewed articles from were "SMS safety culture," "safety culture," "FAA safety culture," "FAA SMS," "measuring safety culture," and "how does the FAA measure safety culture." The sources were subjectively analyzed based on the article abstract and selected for inclusion based on applicability to the USDOT OIG recommendations, quantitative means of measuring organizational safety culture, qualitative means of measuring safety culture, and research into safety culture and climate. A case study was appropriate for this study as a real or quasi-experiment is not necessary given the amount of research conducted in this area.

Literature Review

The Safety Management International Collaboration Group (SM ICG) is a collective of the regulatory bodies from Spain, Brazil, the Netherlands, New Zealand, Singapore, Hong Kong, Australia, France, Italy, The European Aviation Safety Agency, Switzerland, Finland, Ireland, Japan, the FAA, Canada, the United Arab Emirates, and the United Kingdon. SM ICG (2019) published an evaluation tool to assist regulators and organizations to determine the state of their SMS. It is noted in the tool that it should not be used to develop a score for an SMS nor should the tool be used as pass/fail criteria in SMS implementation. This is a subjective tool with criteria that could require guidance to inspectors from each regulatory body in applying the guidelines of each section across different organizations (SM ICG, 2019). Narrowed to culture, the tool first labels culture as "safety culture," then changes to "positive safety/just culture," and eventually completely changes the label to "just culture" (SM ICG, 2019). The changing of the label for culture could impact the subjective evaluation.

Safety Climate versus Safety Culture

Petitta et al. (2016) examined the understudied difference between safety climate and safety culture. Petitta et al. (2016) postulated that there was a lack of a clear theoretical difference between the two ideas in academia and industry. In defining safety climate, four factors stood out in the literature: the organization member's perception of management concern for the safety of the employee, perception of opportunity to discuss safety issues, adequacy of safety training, and the quality of safety management systems, specifically their effectiveness in preventing work incidents. In comparison, safety culture is "a shared set of safety-related attitudes, behaviors, values, and ingrained assumptions that orient organizational action pertaining to safety" (Petitta et al., 2016, p. 79). The difference between climate and culture was assumed to be individual versus group ideology, respectively, on the subject.

Using a six factor confirmatory factor analysis model for the data collected from 32 Italian organizations, Petitta et al. (2016) had results consistent with prior research on safety climate and safety culture. Additionally, the findings indicated safety climate directly affected employee safety compliancewith safety climate noted as a construct that was distinct from safety culture. The limitations noted in the study was the use of convenince samples that could introduce self-selection bias as well as the use of cross-sectional self-reported data that could introduce common method bias with affects on the validity of the findings (Petitta et al., 2016).

Safety Surveys

Gibbons et al. (2005) conducted a pilot study with a five factor safety survey for maintenance that was similar to a flight operations survey to develop a standard survey to assess organizational safety culture. Safety surveys are commonly used to assess safety climate and safety culture in an organization with mixed feelings about the effectivity of such surveys from individuals in the industry. The study was limited to two Part 121 airlines with a response rate of 13% and 26% for a total of 185 returned surveys. For the overall model fit that was studied, the single factor model and the five factor model had poor model fit. However, Gibbons et al. (2005) did find that two of the five factors they investigated were problematic in the survey. The revised complete model still did not show a good fit with the two factors of organizational commitment and employee empowerment. The conclusion was that those two factors may require further consideration due to the complexity uncovered when analyzing the survey data. The concerted effort on the part of Gibbons et al. (2005) to develop a viable survey demonstrated the difficulties in using surveys to analyze safety culture in an organization.

While the FAA supported the study by Gibbons et al. (2005) with a funding award, the FAA (2016b) provides a link to the Government of Canada (2015) safety culture survey that is based on the James Reason's Checklist for Assessing Institutional Resilience. This survey developed a numerical score that enabled a quick assessment from the management level on the safety culture of the organization. However, this safety culture survey is limited by the focus on management commitment, competence and cognizance (Government of Canada, 2015).

Qualitative Analysis

Robertson (2016) conducted a qualitative analysis of perceptions of safety professionals with regard to SMS and safety culture in flight training organizations. Of those organizations, only two of the five organizations involved in the study had mature SMS. It was noted that while three of the organizations utilized surveys to gather safety data, two of the organizations used Line Operations Safety Assessments, audits, observations, and flight data management to evaluate safety culture. Cultural change was noted for organizations that reported commitment from accountable executives as well as stakeholder involvement when benefits

from cultural assessments and independed hazard reporting were realized. The development and building of a strong SMS was noted to come from an already existent strong safety culture (Robertson, 2016).

Robertson (2016) noted a limitation in the study because of the small, purposeful sample size from the University Aviation Association safety professionals. Further, the study was based in phenomenology and was identified as not being generalizable to a greater population. The recommendations from the study was to continue qualitative and quantitative research into the relationship between safety culture and SMS (Robertson, 2016).

Quantitative Analysis

Stolzer et al. (2018) researched a model using data envelopment analysis to test a survey intended to measure the effectiveness of an SMS. The survey instrument developed was determined to have a high level of convergent validity and was determined to be reliable. The discriminate validity was unable to be confirmed because of insufficient evidence. Stolzer et al. (2018) made mention that there was no clear path for evaluating an SMS, but the model developed may be able to be used by the aviation industry to evaluate their own implemented SMS programs.

The model developed in this study was intended for use only by aviation service providers and would not be an appropriate tool for other aviation operations. With the responses to the survey being anonymous by not collecting personally identifiable information, the assumption made of the collected data was that all answers were truthful. The study was funded by the FAA through a research grant and recommended further refinement of the survey questions along with running the models again with more data for the goal of developing an industry tool (Stolzer, 2018).

Wei et al. (2020) countered the position of Stolzer et al. (2018) of survey usage with their study of 82 different enterprises in 11 separate industries and found that surveys have little effect on the safety culture of an organization. Their results mirrored a study conducted by Stewart (2002), also referenced in their study, during his tenure at DuPont that indicated safety culture is representative of the safety climate. Essentially, the higher safety performing organizations had a greater number of individuals who had a greater understanding of safety performance in common while the poor performing organizations had a greater number of employees with a poor understanding of safety performance (Wei et al., 2020).

Of note, the majority of enterprises included in the study were based in coal production and limited the ability of Wei et al. (2020) to make any industry comparisons. However, the conclusions were consistent with the findings of Stewart (2002), but were identified as not being generalizable to the overall safety culture of China due to the small sample of organizations in relation to the full range of Chinese enterprises.

For Profit Evaluators

Commercially available safety culture tools are available from for-profit organizations and were investigated by van Nunen et al. (2018) specific to the Belgian market to examine the extent with which the studied tools were based in scientific evidence. Of the 68 tools found through internet searches, 15 tools were included based on seven inclusion criteria that focused on organizational or behavioral aspects and appropriateness to occupational safety and health legislation. With the 15 tools not being freely available, van Nunen et al. (2018) contacted each company and received agreeable responses from eight companies. Through the evaluation of the selected tools, subjective expert opinion was the focus of each tool with many having a basis in scientific frameworks. However, none of the tools studied were validated through scientific evidence, the relative importance of the tools to accident rates was unknown, and the internal validity of the tools were not known. A recurrent theme of theory not put into practice and practice not being put into theory was reiterated in the conclusion. The recommendation was for the commercial organizations and academia to work together to close the tool validation gap.

Discussion

Measuring organizational safety culture is benefited through the amount of peer reviewed articles published that have used quantitative and qualitative means of measuring and assessing organizational safety culture. While there are differences in the means of measuring, the understanding of safety climate and safety culture, to include the difference in definition, seems to be agreed upon by most researchers. This provides a strong framework for commercial enterprises to develop tools used to evaluate safety culture in organizations subject to State regulation. This knowledge could be used by the FAA to develop and implement guidance for inspectors to address a recommendation from the USDOT OIG (2020).

There are limitations to the published studies. Many of the studies identified concerns with results that may have caused the researchers to recommend further study. Additionally, reseachers often made note of generalizability based on the sample size or population that could have an affect on the results. This limitation could result in tools that are specific to an industry or organization and may not be a one-size tool for the general aviation industry. Limitations from published scientific works should not be considered a negative for the FAA. Instead, the limitations found through research could help define the constraints of a developed tool or constrain the methodology used based on industry or culture.

The most commonly mentioned assumption, based on most of the studies using surveys to collect data, was that the survey participants were telling the truth. While the researchers took steps to help protect participant anonyminity, violation of this assumption by participants could change the results. This assumption, and

the resultant tailoring of surveys to mitigate a risk of participants providing untrue responses, could be used by the FAA to limit the amount of demographic data collected to mitigate the same risk in industry tool usage.

Conclusion

Addressing the USDOT OIG (2020) recommendation on providing guidance on safety culture evaluation could be a difficult task. While commercially available tools exist, they seem to lack internal validity even being based on previous scientific frameworks. While academia has produced validated means to assess the collected data, concerns with some results seems to create more recommendations for study. For the FAA, funding and facilitiation of cooperation between commercial industry and academia could be the path towards addressing a gap in oversight identified by the USDOT OIG.

Recommendations

One recommendation is based on the findings of van Nunen et al. (2018) where funded research into safety culture evaluation methods lead to the publication of tools. Two published studies included in this work were funded by the FAA, with one study showing promise of a scientifically validated tool, but the theory is not put into practice. The FAA (2016b) website links to aforementioned survey posted by the Government of Canada (2015). The recommendation is for the FAA to continue to fund, develop, and release scientifically validated tools for safety culture evaluation.

A second recommendation is for academia to look for opportunities to work with commercial organizations that provide unvalidated safety tools to industry. By looking for and exploiting partnering opportunities, academia and private industry can work towards closing the gap between theory not becoming practice and practice not becoming theory.

References

Federal Aviation Administration. (2014, August 4). *Safety management system misconceptions*. https://www.faa.gov/about/initiatives/sms/explained/misconceptions/

Federal Aviation Administration. (2016a, July 5). *Safety management system explained*. https://www.faa.gov/about/initiatives/sms/explained/

Federal Aviation Administration. (2016b, July 5). *Safety management system training*. https://www.faa.gov/about/initiatives/sms/reference_library/training/

Federal Aviation Regulations, 14 CFR § 5 (2015). https://www.ecfr.gov/current/title-14/chapter-I/subchapter-A/part-5

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. *2005 International Symposium on Aviation Psychology*, 259-264. https://corescholar.libraries.wright.edu/isap_2005/37

Gill, G. K. & Shergill, G. S. (2004). Perceptions of safety management and safety culture in the aviation industry in New Zealand. *Journal of Air Trasport Management, 10*, 233-239. https://www.doi.org/10.1016/j.jairtraman.2004.02.002

Government of Canada. (2015, July 21). *Score your safety culture – TP 13844.* https://tc.canada.ca/en/aviation/publications/score-your-safety-culture-tp-13844

International Civil Aviation Organization. (n.d.). *Guidance material*. https://www.icao.int/safety/SafetyManagement/Pages/GuidanceMaterial.aspx

van Nunen, K., Reniers, G. & Ponnet, K. (2018). Measuring and improving safety culture in organisations: An exploration of tools developed and used in Belgium. *Journal of Risk Research*, *21* (5), 622-644. https://www.doi.org/10.1080/13669877.2016.1235602

Pettit, K. (2017). SMS, safety culture, and the four pillars of safety applied to airline pilot training: Nextgen demands to improve safety. *International Journal of Aviation Systems, Operations and Training, 4* (2), 45-57. https://www.doi.org/10.4018/IJASOT.2017070104

Petitta, L., Probst, T. M., Barbaranelli, C. & Ghezzi, V. (2016, November 22). Disentangling the roles of safety climate and safety culture: Multi-level effects on the relationship between supervisor enforcement and safety compliance. *Accident Analysis and Prevention, 99* (2017), 77-89. http://www.doi.org/10.1016/j.aap.2016.11.012

Robertson, M. F. (2016). Safety professional's perception of the relationship between safety management systems and safety culture. *Journal of*

Aviation Technology and Engineering, 6 (1), 9-15. http://www.doi.org/10.7771/2159-6670.1137

Safety Management International Collaboration Group. (2019, April). *Safety management system (SMS) evaluation tool*. https://skybrary.aero/articles/sm-icg-sms-evaluation-tool

Stewart, J. M. (2002). *Managing for world class safety, first ed*. Wiley.

Stolzer, A. J., Friend, M. A., Truong, D., Tuccio, W. A., & Aguiar M. (2018, January 9). Measuring and evaluating safety management system effectiveness using data envelopment alanysis. *Safety Science, 104* (2018), 55-69. https://www.doi.org/10.1015/j.ssci.2017.12.037

United States Department of Transportation Office of Inspector General. (2020, February 11). FAA has not effectively overseen Southwest Airlines' systems for managing safety risks. https://www.oig.dot.gov/sites/default/files/FAA%20Oversight%20of%20Southwest%20Airlines%20Final%20Report%5E02.11.2020.pdf

Wei, J., Gui, F., Chun-yang, L. & Han, W. (2020, April 21). Study on quantitative measurement result of safety culture. *Safety Science, 128* (2020). https://doi.org/10.1016/j.ssci.2020.104751