<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Measuring safety climate in the aviation industry: A review and recommendations for the future.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>O'Connor, Paul</td>
</tr>
<tr>
<td><strong>Publication Date</strong></td>
<td>2011</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Elsevier</td>
</tr>
<tr>
<td><strong>Item record</strong></td>
<td><a href="http://hdl.handle.net/10379/2539">http://hdl.handle.net/10379/2539</a></td>
</tr>
</tbody>
</table>

Some rights reserved. For more information, please see the item record link above.

Paul O’Connor¹, Angela O’Dea², Quinn Kennedy³, & Samuel E. Buttrey⁴

1. Corresponding Author: Centre for Innovation and Structural Change, J.E. Cairnes School of Business & Economics. National University of Ireland, Galway. Ireland Tel: +353 (0) 91 492897 Fax: +353 (0) 91 495524. Email: poc73@hotmail.com

2. Research Department, Naval Postgraduate School, Monterey, California 93943. Email: aodea69@hotmail.com

3. Operations Research Department, Naval Postgraduate School, 1411 Cunningham Rd, Monterey, California 93943, USA. Email: mqkenned@nps.edu

4. Operations Research Department, Naval Postgraduate School, 1411 Cunningham Rd, Monterey, California 93943, USA. Email: buttrey@nps.edu

ABSTRACT

This paper reviews 23 studies that have examined safety climate within commercial and military aviation. The safety climate factors identified in the aviation safety climate questionnaires were found to be consistent with the literature examining safety climate in non-aviation high reliability organisations. Therefore, it was concluded that the aviation safety climate tools had some construct validity (the extent to which the questionnaire measures what it is intended to measure). However, the majority of the studies made no attempt to establish the discriminate validity (the ability of the tool to differentiate between organizations or personnel with different levels of safety performance) of the tools. It is recommended that rather than constructing more aviation safety climate questionnaires, researchers should focus on establishing the construct and discriminate validity of existing measures by correlating safety climate with other metrics of safety performance. It is recognized that the accident rate in commercial aviation is too low to provide a sufficiently sensitive measure of safety performance. However, there are other measures of safety performance, collected as part of a company’s Aviation Safety Action Program or Voluntary Flight Operational Quality Assurance, which could be used to assess the discriminate validity of an aviation safety climate tool.

KEYWORDS: Safety climate, aviation, validity
1. Introduction

Traditionally, organizations have assessed their safety performance on the basis of “lagging indicators” of safety such as fatalities, or mishap rates. Lagging indicators show when a desired safety outcome has failed or has not been achieved (e.g., number of mishaps). However, as safety has improved and the frequency of mishaps has declined, mishap rates have ceased to be a useful metric of safety performance. Industries in which performance may be catastrophically impacted by failures in complex human technology systems are known as High Risk Industries (Shrivastava, 1986). Those organizations that succeed in avoiding catastrophes in high risk environments are known as High Reliability Organizations (HROs; Roberts & Rousseau, 1989). Given the low numbers of accidents that occur in HROS, these organizations have started to examine “leading indicators” of safety in an attempt to improve safety performance even further. The United Kingdom’s Health and Safety Executive (HSE, 2006) defined leading indicators of safety as measures of process or inputs essential to deliver the desired safety outcomes (e.g., safety climate surveys, hazard reports). Therefore, leading indicators of safety provide a more proactive method to gain insight into the safety performance of the organization and identify areas in which efforts should be made to improve safety.

One of the most commonly used leading indicators of safety in non-aviation HROs is safety climate. Zohar (1980) defined safety climate as a summary of perceptions that employees share about their work environment. Safety climate describes employees’ perceptions, attitudes, and beliefs about risk and safety (Mearns & Flin, 1999). It is a “snapshot” of the current manifestation of the safety culture in the organization. There has been an ongoing debate within the literature regarding the use of the terms “culture” and “climate,” and whether they represent the same or different concepts. The general consensus is that culture represents the more stable
and enduring traits of the organization, and has been likened to “personality.” Safety culture reflects fundamental values, norms, assumptions, and expectations, which, to some extent, reside in societal culture (Mearns & Flin, 1999). Climate, on the other hand, is thought to represent a more visible manifestation of the culture, which can be seen as its “mood state,” at a particular moment in time (Cox & Flin, 1998).

Wiegmann and colleagues (Wiegmann, Zhang, von Thaden, Sharma, & Gibbons, 2004) report that “few formally documented efforts have been made to assess safety culture within the aviation industry, with the notable exception of military aviation” (p. 117). This finding is surprising, given that the civilian aviation industry has been a leader in the development and utilization of a number of human-focused safety programs (e.g., crew resource management). In the last decade there has been an increase in aviation specific safety climate research such that there is now sufficient research to merit a literature review of this work. A key element missing from the literature is the extent to which aviation safety climate surveys actually measure what they are intended to measure and discriminate between groups varying in safety performance. The purpose of this paper is:

- to carry out a literature review of published aviation safety climate research;
- identify whether there is evidence of construct (the extent to which the questionnaire measures what it is intended to measure) and discriminate (the ability of the tool to differentiate between organizations or personnel with different levels of safety performance) validity of the questionnaires used; and
- make recommendations for what should be done to improve the validity of safety climate assessment in the aviation industry.
2. Literature Review

2.1 Search Methodology

A computerized search of the literature was conducted utilizing PsycINFO, Google Scholar, Medline, and Defense Technical Information Center. Keywords for the computerized search of the literature were: “aviation” with “safety climate”, or “safety culture”. The reference lists of published aviation safety climate studies were also examined. The search identified a total of 23 studies reporting a safety climate evaluation carried out in aviation (one study was published as both a report and journal article). A total of 48% of the studies were published in peer review journals, with the remainder consisting of theses (35%), reports (13%), and conference proceedings (4%). A total of 48% of the studies were conducted with U.S. military populations, with the remainder carried out in commercial aviation organizations. Examining the papers by occupational group, 35% were carried out with maintainers, 30% with pilots, 9% with ground handling personnel, 4% with cabin crew, and 17% were carried out with a mixture of occupational groups. The studies are summarized in Table 1, by occupational group, a detailed discussion of the studies is reported below.
Table 1. Summary of aviation safety climate studies.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Population</th>
<th>Questionnaire</th>
<th>Categories</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Civilian aviators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Evans et al. (2007); Australian Transportation Safety Bureau (2004) | 1,308 Australian pilots (26% response rate). | 30 item survey. | • Management commitment and communication  
• Safety training  
• Equipment and maintenance  
• Rules and procedures  
Communication (dropped following factor analysis)  
• Schedules (dropped following factor analysis) | The three factor model was established using EFA with half of the sample, and the model was supported using CFA with the remaining half of the sample.  
No significant differences were found between different groups of aviators (regular public transport, charter, and aerial work pilots). |
| Gibbons, et al. (2006) | 503 responses from pilots from a large U.S. airline (29% response rate). | 84 item survey based upon literature review. | • Organizational commitment  
• Operations personnel  
• Informal safety system  
• Formal safety system  
• Compliance | Factor structure was established through structural equation modeling.  
Need to account for the atypical structure of management–employee relationships in the airline industry. |
| **Cabin Crew** | | | | |
| Kao, Stewart, & Lee (2008) | 331 Taiwanese cabin crew (85% response rate). | 22 item questionnaire derived from previously used climate survey. | • Management commitment to safety  
• Cabin work environment  
• Rule compliance  
• Crew member involvement and participation  
• Accident investigation | Structural equation modeling confirmed the overall fit.  
High management commitment was significantly related to high crew member participation.  
Safe cabin work environment was significantly related to crew member's individual behavior.  
No relationship between management commitment and injury incidence. |
<table>
<thead>
<tr>
<th><strong>Civilian aviation ground handlers</strong></th>
<th><strong>Civilian Aviation Maintainers</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>166 ground handling personnel from three Spanish companies (45% response rate).</td>
<td>622 aviation maintainers from four European aviation maintenance organizations (78% response rate).</td>
</tr>
</tbody>
</table>
| • Company policies towards safety  
• Emphasis on productivity vs. safety  
• Specific strategies of prevention  
• Safety level perceived at the airport  
• Safety level perceived on the job | None delineated, only the mean climate score was analyzed. |
| • A factor analysis resulted in 6 factors explaining 69.8% of the variance.  
• Significant differences were found between the fuel company, airport authority, and the ground handling division.  
• No significant differences in position, level of education, whether working on the ramp or not. | Safety climate was found to discriminate between the different companies and agreed with the findings from interview and incident data. |
| 50 ground handlers from a single Swedish aviation ground handling company (75% response rate). | 109 item survey developed from interviews and observations of ground handlers. |
| • Working situation  
• Communication  
• Learning  
• Reporting  
• Justness  
• Flexibility  
• Attitudes towards safety  
• Safety-related behaviors  
• Risk perception | No factor analysis performed.  
Results revealed a good existing safety culture. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Methodology</th>
<th>Factors Identified</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gordon et al. (2007)</td>
<td>119 European Air Traffic Control (ATC) personnel (30% response rate).</td>
<td>59 items derived from a literature review and interviews with 52 ATC personnel.</td>
<td>• Communication/consultation  • Support from others  • Organizational support for safety  • Reporting  • Resources  • Organizational safety learning  • Responsibility for safety</td>
<td>Following an EFA, eight factors were identified.</td>
</tr>
<tr>
<td>Kelly &amp; Patankar (2003)</td>
<td>Used the Patankar (2003) sample in addition to a sample of 237 responses from another company.</td>
<td>Used the Patankar (2003) survey.</td>
<td>Same as Patankar (2003).</td>
<td>• Overall, found a positive safety culture.  • Differences between the companies were attributed to differences in age and experience.</td>
</tr>
<tr>
<td>Patankar (2003)</td>
<td>399 responses from both flight operations and maintenance personnel from a single U.S. company (55% response rate).</td>
<td>50 item survey derived from other safety climate surveys.</td>
<td>• Pride in company  • Professionalism  • Safety  • Supervisor trust and safety  • Effects of my stress  • Need to speak-up  • Safety compliance  • Hazard communication</td>
<td>• Factors were identified from factor analysis.  • Significant differences were found between groups for pride in company, safety opinions, and supervisor trust.</td>
</tr>
<tr>
<td>Study</td>
<td>Methodology</td>
<td>Findings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Block et al. (2007)       | Used 281 flight crew members responses from the Patankar (2003) study.    | • Purpose  
• Alignment  
• Control  
• Acceptable model fit obtained using structural equation modeling methodology.  
• The main drivers of safety outcomes were organizational affiliation and proactive management. |
| Gill & Shergill (2004)    | Developed a 52 item survey based upon the safety climate literature and a pilot study. | • Positive safety practices  
• Safety education  
• Implementation of safety policies and procedures  
• Individual’s safety responsibilities  
• Organizational dynamics & positive safety practices  
• Regulator’s role  
• Luck and safety  
• Management training & decision making  
• Factors were identified from a factor analysis.  
• Maintainers are committed to standards and procedures.  
• Pilots regard luck to be a significant contributor to safety.  
• Employers were not perceived to be placing much importance on safety management systems and safety culture. |
| U.S. Naval Aviators        | Aviators responded to CSAS and hospital workers to the Patient Safety Cultures in Healthcare Organizations (PSYCHO) survey. 23 items were common to both surveys. | None reported.  
The problematic response rate for hospital workers was up to 12 times greater than that among aviators on certain items. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size</th>
<th>Data Collection</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Desai et al (2006)     | 6,361 responses from 147 U.S. Naval squadrons. | Used the 61 item command safety assessment survey (CSAS) based upon the high reliability organization literature. | • Process auditing  
• Reward system  
• Quality assurance  
• Risk management  
• Command and control | Utilizing the mean safety climate score, positive associations were found between minor or intermediate severe accidents and future safety climate scores, although no effect was found for major accidents. |
| Adamshick (2007)       | 2,943 responses obtained from U.S. Navy strike-fighter aviators from 2001 to 2005. | Used the 61 item CSAS. | Same as Desai et al. (2006). | • PCA independently for each of the five theoretical factors of the CSAS. For all of the factors, except for quality assurance and reward systems, it was found that a two or more factors solution resulted in a better fit than a single factor model.  
• Perceived leadership factors were positively associated with safety climate differ between officers and enlisted. |
| Buttrey, et al. (2010) | 110,014 obtained from U.S Navy and Marine Corps aviators. | Used the 61 item CSAS. | Same as Desai et al. (2006). | • Carried out an EFA with half of the sample, followed by a CFA on the other half of the sample. The analysis did not result in a stable factor structure.  
• After screening the data to identify the respondents using an optimizing strategy (23,442 responses), and only including the 12 items for which there was some variance, it was possible to establish a stable two factor model (personnel leadership and integration of safety and operations). |
<table>
<thead>
<tr>
<th>U.S Naval Aviation Maintainers</th>
<th>268 responses from maintainers at three U.S. Naval aviation reserve squadrons.</th>
<th>67 item survey based upon a high reliability organization model and the CSAS.</th>
<th>Following a PCA, found, 25 out of the 67 items loaded on a single principle component. All of the factors were represented in the principle component. A 35-item version of the questionnaire was proposed.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baker (1998)</td>
<td>439 responses collected from maintainers at eight U.S Navy reserve squadrons.</td>
<td>Used the Baker (1998) 35 item revised survey.</td>
<td>Following a PCA, found a single principle components explaining a third of the variance. All of the factors were represented.</td>
</tr>
<tr>
<td>Goodrum (1999)</td>
<td>977 responses from maintainers at a U.S Marine Corps Air Wing.</td>
<td>Used the Baker (1998) revised survey.</td>
<td>Following a PCA, found a single principle components explaining a third of the variance. All of the factors were represented. Significant differences were found between communities based upon the type of aircraft.</td>
</tr>
<tr>
<td>Stanley (2000)</td>
<td></td>
<td></td>
<td>Demographics had little utility in predicting the scores of a given unit.</td>
</tr>
</tbody>
</table>
| Hernandez (2001) | 2,180 maintainer responses from 30 U.S. Naval aviation units. | Used the Baker (1998) revised survey. | Same as Baker (1998). | • Following a PCA, found a single principle components explaining a third of the variance. All of the factors were represented.  
• No difference in scores between internet and paper-and-pencil version of the questionnaire.  
• No difference in safety climate score between squadrons that had experienced a maintenance incident and those that had not. |
3. Summary of studies separated by occupational group

3.1 Commercial Pilots

Three studies reported a safety climate assessment using commercial aviation pilots (see Table 1). The Australian Transportation Safety Board (2004) and Evans, Glendon, and Creed (2007) report on the development of a safety climate questionnaire, designed to gain insight into pilots’ perceptions of workplace safety. The questionnaire consisted of six safety factors (see Table 1), each with five items. These factors were based upon previous safety climate research and input from aviation safety experts. Data from half of the sample were used in an exploratory factor analysis (EFA) that resulted in a three factor model of: management commitment and communication, safety training and equipment, and maintenance. A confirmatory factor analysis (CFA) on the remaining half of the sample showed the three factor model to be an adequate fit to the data. Finally, the responses from different types of pilots (regular public transport, charter, or aerial work such as emergency medical services or agriculture) were compared on each of the four identified safety climate factors. No significant differences between the groups were found. The Australian Transportation Safety Board (2004) concluded that this was due to a single professional safety climate for pilots as a group, regardless of the organization for whom they worked.

Gibbons, von Thaden, and Wiegmann (2006) developed a questionnaire designed to assess safety culture within the context of airline flight operations. The survey consisted of 84 items, grouped into five themes (see Table 2). The survey was designed by examining the content of safety climate questionnaires that have been used in other HROs. A total of 503 responses were received from a single company. After discarding 29 items and using CFA, the
analysis eventually resulted in structure of four broad factors (organizational commitment, operations personnel, informal safety system, and formal safety system), with three subfactors in each. The authors attribute their difficulty in establishing a stable factor structure with the analysis to issues in item writing (e.g., ambiguity, items that did not relate well to the target population). Another issue not mentioned in the paper is the relatively low ratio of responses to items (6.3 items for every response). No analysis of the revised questionnaire was reported.

3.2 Cabin Staff

Kao, Stewart, and Lee (2009) developed a 23-item questionnaire to assess the safety climate attitudes of Taiwanese cabin crews. The questionnaire was designed to assess the following safety climate themes: management commitment towards safety, cabin work environment, rule compliance, crewmember involvement and participation, accident investigation, and injury incidence. The items were based upon previous safety climate research. A total of 331 responses were obtained from cabin crews from four major Taiwanese airlines. Using a structural equation modeling approach, the researchers found an acceptable level of fit with the proposed factors. High management commitment to safety was significantly related to high crewmember participation in safety, and that safe cabin work environment was significantly related to crewmember’s individual behavior. However, the findings did not reveal a direct relationship between management commitment and injury incidence.

3.3 Ground Handlers

Diaz and Cabrera (1997) developed a 40-item safety climate questionnaire for aviation ground handlers, based upon the work of Zohar (1980). Ground handling is concerned with the
servicing of an aircraft while it is on the ground at an airport. Following a PCA on the data collected from 166 ground handling personnel from three different companies (the ground handling division of an airline, a fuel company, and the airport authority) at a Spanish airport, six factors were identified (see Table 1). Ratings were also obtained from 29 experts in ground handling operations on the level of safety in each company. It was found that the safety climate questionnaire responses from the three companies were consistent with the expert ratings of the levels of safety at the companies. Diaz and Cabrera (1997) concluded that the questionnaire was able to discriminate between organizations with different levels of safety.

Ek and Akselsson (2007) evaluated the safety culture in the ramp division of a ground handling company. A 109-item questionnaire was developed that addressed nine aspects of safety climate (see Table 1). Data were collected from 50 men employed by a single ground handling company. Acceptable levels of internal consistency were found for each factor. They concluded that the safety climate was good, but poorer than desired by managers.

3.4 Aviation Maintainers

As part of a larger research project, McDonald, Corrigan, Daly, and Cromie (2000) designed and utilized a safety climate questionnaire to survey aviation maintainers. The questionnaire was adapted from the one developed by Diaz and Cabrera (1997; described above). A 36-item questionnaire was designed based upon a factor analysis of 69 items (this analysis was not reported). A total of 622 responses were obtained from aviation maintainers from four companies. Significant differences in climate were found between different occupational groups. McDonald et al. (2000) reported that the data provided evidence of a strong professional subculture, which spanned all of the four companies that participated in the study. Further, this
subculture is relatively independent of the organization. Similar to the findings reported in the ATSB (2004) study described above, it was postulated that the subculture is likely to mediate between the organization’s safety management system and safety outcome.

3.5 Air Traffic Controllers

Gordon, Kirwan, Mearns, Kennedy and Jensen (2007) describe a pilot study of a climate survey designed for use by European air traffic controllers (ATC). The questionnaire consisted of 59 items of 13 elements designed around three themes (priority of safety, involvement in safety, and learning from safety). The items were selected based upon a literature review, 50 interviews with ATC personnel, and input from subject matter experts on the final items to be included. The questionnaire was piloted with 119 responses obtained. Following an EFA an eight factor questionnaire resulted (see Table 1 for a description of the factors). Gordon et al (2007) acknowledge that the sample was small, and they state that a larger validation study will be carried out.

3.6 Combined Aviation Occupational Groups

Four studies reported the evaluation of safety culture that included participants from a number of occupational groups. Patankar (2003) evaluated the safety climate of a stratified sample of 399 personnel (flight operations, maintenance, and other personnel) from a single aviation company using a common safety climate questionnaire. After a factor analysis (no details of this were reported), eight factors emerged (see Table 1). Significant differences were found between flight operations, maintenance, and “other” personnel with regard to the factors of pride in company, safety opinions, and supervisor trust. Patankar (2003) concluded that, overall,
the respondents were proud to work for the company, trusted management, and believed that safety is a result of collective efforts. Both flight and maintenance personnel had a high sense of personal responsibility for flight safety.

In a later study, the data collected by Patankar (2003; called company A) was compared to 237 responses collected at another company (called company B; Kelly & Patankar, 2004). It was found that, overall, there was a more positive safety climate at company A than company B. However, this finding was partially attributed to company A having older and more experienced pilots and mechanics than company B.

Block, Sabin, and Patankar (2007) reanalyzed the responses obtained from the 281 pilots from the Patankar (2003) sample. The purpose was to examine whether the data supported what Block et al. (2007) described as the purpose-alignment-control (PAC) model. A pair of experts recoded the Patankar (2003) survey items in accordance with the PAC model. The proposed factors were tested using a structural equation modeling methodology. The main drivers of safety outcomes were organizational affiliation (similar to ‘pride in company’ from Patankar, 2003) and proactive management (partially derived from the ‘safety opinion’ factor from Patankar, 2003). Organizational affiliation was directly influenced by communication, and proactive management was influenced by training effectiveness and relational supervision.

Gill and Shergill (2004) conducted a safety climate review across the New Zealand commercial aviation industry. The safety climate questionnaire they developed included questions designed to address two themes: organizations’ approach to safety management (26 items) and “safety management systems, and safety culture in organizations” (26 items). A factor analysis of 464 responses was run independently on each theme. The “safety management systems” theme was found to consist of four factors: positive safety practices; safety education;
implementation of safety policies and procedures; and individual’s safety responsibilities. The “safety culture in organizations” theme was also found to consist of four subfactors: organizational dynamics and positive safety practices; regulator’s role; luck and safety; and safety management, training, and decision making. The main findings from the study were that pilots believed luck and safety to be the most important factor in aviation safety, and employers were not perceived to be placing much importance on safety management systems and safety culture.

### 3.7 U.S. Naval Aviation

The U.S. Navy utilizes two different tools to assess safety climate in aviation. The Command Safety Assessment Survey (CSAS) is used to obtain feedback from aviators, and the Maintenance Climate Assessment Survey (MCAS) to obtain information from aviation maintainers. Because 39% of the studies utilized the CSAS and/or the MCAS, we explain the CSAS and MCAS and pertinent results in more detail, and use these surveys as an example of how methodological issues can impact the interpretation of the survey results.

The safety culture questionnaires were developed by researchers at the Naval Postgraduate School in Monterey, California (Desai, Roberts, & Ciavarelli, 2006). Both questionnaires are completed online. The questionnaires were based upon a conceptual model of Organizational Safety Effectiveness (MOSE) that identified five major areas relevant to organizations in managing risk and developing a climate to reduce accidents in HROs (Libuser, 1994; Roberts, 1990). The five MOSE areas are:

- Process auditing – a system of ongoing checks to monitor hazardous conditions.
- Reward system – expected social compensation or disciplinary action to reinforce or
correct behavior.

- Quality assurance – policies and procedures that promote high quality performance
- Risk management – how the organization perceives risk and takes corrective action.
- Command and control – policies, procedures, and communication processes used to mitigate risk.

On the basis of observations and interviews with maintainers, the MCAS has an additional sixth MOSE called “communication/functional relationships.” This theme is concerned with having an environment in which information is freely exchanged, quality assurance is seen as a positive influence, and maintenance workers are shielded from external pressures to complete a task (Harris, 2000). A description of the research that has been carried out using the MCAS data will be described first, followed by studies that have utilized the CSAS.

3.7.1 U.S. Naval Aviators

Adamshick (2007) analyzed the data of every Navy and Marine Corps Strike-Fighter aviator that completed the CSAS from 2001 until 2005 (2,943 responses). He carried out PCA independently for the items that make up each of the five theoretical factors of the CSAS. For all of the factors, except for quality assurance and reward systems (for Naval aviators only), it was found that a two or more factors solution resulted in a better fit to the theoretically-derived factors than a single factor model.

Gaba, Singer, Sinaiko, Bowen, & Ciavarelli (2003) compared the responses of health care respondents with those from Naval aviation. Aviators responded to CSAS and hospital workers to the Patient Safety Cultures in Healthcare Organizations (PSYCHO) survey. Both of these
instruments have partially overlapping items, with 23 items from the PSYCHO adopted directly from the CSAS. The survey included employees from 15 hospitals and Naval aviators from 226 squadrons. For each question a “problematic response” was defined as a response that suggested a lack of or antithesis to safety climate (Gaba et al., 2003). Overall, the problematic response rate for hospital workers was up to 12 times greater than that among aviators on certain items. These findings were true both for the aggregate of all health care respondents and, even more strikingly, for respondents from particularly hazardous health care arenas (e.g., emergency rooms and critical care) the number of problematic responses were 16 times greater than among aviators. This finding indicated that the aviators reported a more positive safety climate than the health care respondents.

Desai et al. (2006) measured the relationship between recent accidents and perceptions of safety climate, as measured by the CSAS, on a large, cross-sectional sample of respondents in several Naval aviation squadrons. The notion was to understand potential cognitive and behavioral changes following accidents. They hypothesized that safety climate would improve after an accident occurred. Moreover, the improvement would be greater following an extremely severe accident as compared to a minor accident. They postulated that after a major accident “managers may be motivated to direct more resources toward safety than are managers in groups with less severe accident records” (Desai et al., 2006: 642). As a result of the increase in investment in safety after an accident, the safety climate improves. Desai et al. (2006) also suggested that this argument is supported by cognitive research in that defensive attributions may increase in strength as the severity of accidents rises.

The study used the 6,361 responses from 147 Naval squadrons taking the online CSAS between July 2000 and December 2001. Aviation mishap information was collected from the
U.S. Naval Safety Centre (the number of mishaps used was not reported). The dependent variable was a safety climate perception construct developed by aggregating each individual’s responses to the CSAS. Six independent variables were constructed to measure accidents prior to survey administration. These mishap variables were recorded at the squadron group level of analysis. Desai et al. (2006) regressed the safety climate construct on several indicator variables tracking the occurrence of accidents, grouped by their severity, in periods roughly one year prior to survey measurement and two years prior to survey measurement. Analysis indicated positive associations between minor or intermediately severe accidents and future safety climate scores, although no effect was found for major accidents. These findings suggest a generally positive association between minor or intermediately severe accidents and perceived safety climate. This study suffers in that only limited information was obtained on the mishaps. Also, although the number of mishaps that occurred during the period of study were not reported, the number was likely to be fairly low. Finally, the rationale that the safety climate will improve after a mishap may be flawed. If the squadron personnel believe that the causes of the mishap have not been addressed, it may be that the safety climate may go down, rather than improve, as suggested by Desai et al.

Buttrey, O’Connor, and O’Dea (2010) attempted to establish the construct validity of the CSAS. They used 110,014 responses to the CSAS collected over eight years. Utilizing a combination of EFA and CFA, Buttrey et al. were unable to identify a stable factor structure for the 61 item CSAS. They attribute this finding to the effect of the non-constant variance of the data. The lack of a constant variance renders standard statistical tests invalid. The data was reduced by retaining only the 12 items which had substantial variance, and to those with a response time of greater than 10 minutes to complete the questionnaire (time to complete was
collected since 2006, and was used as a metric to discard respondents who were suspected of not giving the cognitive effort required to complete the questionnaire. Using a combination of EFA and CFA, with the 22,000 remaining respondents, a stable two factor structure (personnel leadership and integration of safety and operations) was established.

3.7.2. U.S. Naval Aviation Maintainers

A considerable amount of work examining the psychometric properties of the MCAS was carried out by Naval Postgraduate School Masters’ students in the late 1990s and early 2000s (see Table 1). The MCAS was developed by Baker (1998) directly from the CSAS. He carried out Principal Component Analysis (PCA) on 268 responses from the maintenance personnel of three reserve Naval squadrons. He found that 25 out of the 67 items loaded on a single principle component. However, as all of the six MOSEs were represented in this principle component, he concluded that there is no evidence against the construct validity of the questionnaire.

As a result of the analysis, Baker (1998) proposed a revision of the questionnaire consisting of 35 items. As can be seen from Table 1, using Baker’s revised survey with slightly different sub-populations, Oneto (1999), Goodrum (1999), Harris (2000), and Hernandez (2001) all drew similar conclusions regarding the factor structure. That is, a PCA resulted in one single principle component that explained a third of the variance, with almost all of the items from the questionnaire loading on this principle component. Baker, Oneto (1999), Goodrum (1999), Harris (2000), and Hernandez (2001) interpreted the finding of one MCAS principle component with all of the MOSE categories represented as evidence that the MCAS was theoretically sound. Harris (2000) and Hernandez (2001) also analyzed whether there was a relationship between MCAS score and aircraft-maintenance-related incident rate. Neither author reported a significant
relationship.

Most recently, Brittingham (2006) examined the MCAS responses from 126,058 maintainers collected between 2000 and 2005. After completing a PCA, she found that, prior to rotation, one principle component accounted for approximately 50% of the variance. She states that after varimax rotation, a second principle component emerged. The first principle component consisted of items concerned with overall command attention to safety, and the second related to workload and the availability of appropriate resources. Brittingham (2006) interpreted the failure to find the six MOSE components as individual factors in the PCA process to mean that “the MCAS was found to be an inadequate tool with questionable validity for gauging maintenance safety climate” (Brittingham, 2006: 31). This conclusion contradicts the prior conclusions (Baker, Oneto, 1999; Goodrum, 1999; Harris, 2000; Hernandez, 2001) that the MCAS was a valid tool because the single principle component represented all of the MCAS items.

It could be argued that both the interpretation of Brittingham (2006) and that of the earlier studies are flawed, due to the lack of a clear understanding of the methodology that was employed to identify the principle components. PCA is the method to use when the researcher is attempting to reduce a large number of variables to a smaller number of components (Stevens, 1996). PCA analyzes variance with the goal of extracting the maximum variance from a data set with a few orthogonal (i.e., uncorrelated) components (Tabachnick & Fidell, 1996). Because principle component scores are always uncorrelated by construction, unrotated PCA never accounts for correlations between the presumed factors underlying the observations. Furthermore principle components (or their coefficients) are never chosen with reference to a body of theory; they always arise automatically from the maximization of variance explained.
Another related issue, which may have accounted for the majority of items loading on a single principle component, is the large proportion of respondents responding positively to the items. To illustrate, Goodrum (1999) reported that all questions were answered positively, with a mean range of between 3.17 and 4.37 (on a 5-point scale). Hernandez (2001) reported a mean range between 3.18 and 4.15 for the items. Therefore, it would appear that there is limited variability in the responses to the items. This limited variability creates problems when carrying out a PCA because if all of the items have a similar lack of variability, then the PCA will tend to identify one principle component with a large number of items. Thus, it could be argued the PCA is an inappropriate method for identifying factors when variability in responses are limited, and researchers may want to consider other exploratory or confirmatory factor analysis techniques.

We now turn to studies employing the CSAS. In the next sections, the safety climate questionnaires that were used will be assessed to allow conclusions to be made about their construct and discriminate validity.

4. Assessing Construct Validity

Construct validity is concerned with the extent to which the questionnaire measures the underlying theoretical construct it intends to measure. The identification of a reliable factor structure, that is consistent with theory, helps the researcher substantiate claims regarding the construct validity of the questionnaire. The construct validity of the questionnaires will firstly be assessed by examining whether the factors identified by each are consistent with the broader safety climate literature, and then whether the factors identified from each questionnaire converge upon a common set of safety climate themes that are consistent across all of the questionnaires.
Of the ten unique safety climate questionnaires that were identified from the literature review, it was decided to remove the Ek and Akelsson (2007) from the analysis due to the small sample size, and lack of a factor analysis of the questionnaire. The factors reported in the remaining questionnaires were categorized into eight broad safety climate themes (see Table 2). We found that only three factors did not specifically fit within the above eight themes.
Table 2. Classification of questionnaire factors into common safety climate themes.

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Management/supervision</th>
<th>Operations personnel</th>
<th>Safety systems</th>
<th>Procedures/Rule</th>
<th>Communication</th>
<th>Resources</th>
<th>Training/education</th>
<th>Risk</th>
<th>Uncategorized factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evans et al (2007)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>Diaz &amp; Cabera (1997)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓(2)</td>
<td>None</td>
</tr>
<tr>
<td>Gordon et al (2007)</td>
<td>✓(2)*</td>
<td>✓(2)</td>
<td>✓</td>
<td>✓(3)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>Patankar (2003)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓(2)</td>
<td>Pride in company; effect of my stress</td>
</tr>
<tr>
<td>Gill &amp; Shergill (2004)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓(2)</td>
<td>Regulator’s role; luck and safety</td>
</tr>
<tr>
<td>CSAS</td>
<td>✓</td>
<td>✓(3)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>MCAS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓(2)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>None</td>
</tr>
</tbody>
</table>

*Numbers in brackets represent the number of factors from a particular questionnaire that were categorized as this theme when more than one applied.
The safety climate themes identified in Table 2 are broadly in agreement with a number of reviews of the safety climate literature (e.g. Flin, Mearns, O’Connor, & Bryden, 2000; Gadd & Collins, 2002; Guldenmund, 2000; Hale & Hovden, 1998; Shannon, Mayr, & Haynes, 1997). These themes are discussed below.

- **Management/supervision** - All of the aviation safety climate questionnaires had a factor concerned with management/supervision. Similarly, a factor concerned with management has been identified about 75% of the time in other safety climate research (Gadd & Collins, 2002; Flin et al., 2000). The importance of interactions between managers and workers has been clearly established through the research (e.g. Hale & Hovden, 1998). Specifically, management participation and involvement in work and safety activities, as well as frequent, informal communications between workers and management, are recognized as critical behaviors.

- **Safety systems** - A factor related to safety systems is identified in about two-thirds of safety climate studies (Gadd & Collins, 2002; Flin et al., 2000). Mearns, Flin, Gordon, and Fleming (1998) found that reporting systems, rules and procedures, and safety systems were among the key factors related to self-reported accident involvement.

- **Procedures/Rule** - Guldenmund (2000) identified procedural and rule compliance as one of the most frequently occurring themes in his review of safety climate research. Perceptions of safety rules, attitudes to rules and compliance, and violation of procedures are addressed.

- **Training/Education** - The workforce’s perception of the general level of workers’ qualifications, skills, and knowledge is the essence of this theme. Cooper and Phillips
(2004) demonstrated that workers’ perceptions of the significance of safety training were the most important safety climate factor predicting actual safety behavior

- **Risk**: Higher threat perception is positively related to safe behaviors. For example, Goldberg, Dar-El, and Rubin (1991) found that a high threat perception was related to readiness to participate in safety programs, the relationship was mediated by coworker support for safety.

Despite the broad agreement with the themes identified in other safety climate research carried out in HROs, three themes emerged as being particularly relevant to aviation. These themes are discussed below.

- **Communication**: Four of the aviation safety questionnaires had factors concerned with communication (see Table 2). The aviation industry consists of different occupational groups (e.g. air traffic control, maintenance, cabin personnel, pilots, dispatch) that are not co-located. This presents challenges to communication and means that personnel may not have the capacity for direct communication, and by implication, the ability to engage in informal and spontaneous interaction. As a result, safety communication may be more of a challenge in the aviation industry than other HROs in which personnel are co-located.

- **Resources**: Three aviation safety climate questionnaires had factors that were categorized as resources (see Table 2). This factor is concerned with the availability of resources for safety (e.g. money, time, equipment, etc.). This is unlikely to be an aviation specific issue. However, a resources factor was only included in the three most recently developed aviation safety climate questionnaires (see Table 2). Therefore, it is possible
that due to the current economic climate the availability of resources for improving safety has become more relevant than it was in the past.

- **Operations personnel** - this theme was concerned with the commitment of the operations personnel to safety. This theme has aspects of what Flin et al. (2000) categorized as ‘work pressure’ and ‘competence’ in their review of the safety climate literature. It is suggested that due to the different specialized occupational groups in the aviation industry, that it may be necessary to include a specific factor that address the safety commitment of operations personnel.

To summarize, all nine of the questionnaires reviewed consisted of safety climate factors that are in agreement with the broader literature on safety climate in HROs. Although there are themes that may be particularly relevant to safety climate in an aviation environment, these are not inconsistent with the safety climate literature. However, there was a lack of convergence on a specific set of safety themes that were consistent across all of the questionnaires reviewed. From Table 2 it can be seen that the only theme that is addressed in all nine questionnaires was management and supervision. The threat to discriminate validity of a lack of common safety climate themes is not confined to aviation safety climate measures, but does occur in safety climate questionnaires more generally (Flin et al, 2000; Gadd & Collins, 2002). Nevertheless, given the questionnaires in this review were all designed to assess safety climate within aviation, it might be expected that there would be greater convergence in the safety climate constructs assessed. Therefore, although there is evidence for some construct validity, the lack of a consistent set of common safety climate themes is an issue.

**5. Assessing Discriminate Validity.**
In addition to establishing the construct validity of a safety climate questionnaire, it is also necessary to determine the discriminate validity. If the tool is unable to differentiate between organizations or personnel with different levels of safety performance, then it is of limited usefulness. The discriminate validity can be assessed by correlating the data from the questionnaire with a criterion variable such as accidents, or other safety-related behaviour (Guldenmund, 2007). In recent years, a large number of research studies in HROs have sought to examine the contribution of safety climate to accidents. The challenge facing researchers has been to highlight measurable dimensions of safety climate that can be used to identify, in advance, the strengths and weaknesses within an organization that influence the likelihood of accidents occurring. A variety of different criteria are used upon which to base evaluations of organizational effectiveness in preventing accidents.

- **Company accident statistics.** Such studies have shown that the degree of safety program development and workers’ safety initiative were related to lower work accident and injury rates (Zohar, 2000; Donald & Canter, 1994; Mearns, Rundmo, Flin, Gordon, & Fleming, 2004).

- **High and low accident rate plants.** Other studies have compared high- and low-accident-rate plants (or evaluated plants with outstanding safety records) as their criteria upon which to base judgments of effectiveness. For example, management’s commitment to safety was found to be greater in low-accident-rate plants than in the high-accident-rate plants (Cohen, Smith, & Cohen, 1975; Smith, Cohen, Cohen, & Cleveland, 1978; Cohen & Cleveland, 1983).

- **Self report.** Self-reported safety behavior and safety attitudes are an alternative to relying on mishaps data to evaluate the effectiveness of an organization’s safety program. For
example, Thompson, Hilton, and Witt (1998) suggested that minor workplace accidents often go unreported, yet these events may be the best indicators of improving (or worsening) safety conditions that might eventually lead to serious injury.

As can be seen from Table 1, only four studies attempted to establish the discriminate validity of the aviation safety climate survey. Kao et al. (2008) found no relationship between management commitment and self-reported injury incidents. Neither Harris (2000), nor Hernandez (2001) found a significant link between MCAS responses and maintenance incidents in U.S. Navy squadrons. The only study in which evidence of discriminate validity was found was in Diaz and Cabrera’s (1997) assessment of the safety climate of ramp personnel. As discussed in the literature review, the safety climate measures were found to be in agreement with expert ratings of the three company’s level of safety. This finding is encouraging. However, no statistical assessment was carried out to measure the strength of this link. Further, no attempt was made to link the safety climate responses with actual safety performance measures such as mishap rates, hazard reports, etc. This, although the questionnaires would appear to have construct validity, there is insufficient evidence for discriminate validity.

6. Recommendations

It is recommended that rather than constructing more aviation safety climate questionnaires, researchers (and sponsors) should focus on establishing the discriminate validity of existing measures by correlating safety climate with other measures of safety performance. Similar to other HROs, the accident rate in commercial aviation is so low that it is not a useful metric of safety performance. To illustrate, for U.S. commercial aviation, the accident rate was
0.2 per 100,000 flight hours from 2000 until 2009 (NTSB, 2009), compared to 1.5 major accidents per 100,000 flight hours in U.S. Naval aviation over the same period (Naval Safety Center, 2009). As the CSAS has been used by the U.S. Navy to collect safety climate information for close to a decade, there may be a sufficient numbers of accidents such that they can be used as a metric to evaluate the discriminate validity of the CSAS. In commercial aviation other metrics of safety performance should be used to evaluate the discriminate validity of safety climate questionnaires.

For example, many commercial aviation companies have an Aviation Safety Action Program (ASAP) to encourage employees to voluntarily report safety information (see FAA, 2002, for a discussion of these programs). In addition, to ASAPs, many companies also have a voluntary flight operational quality assurance (FOQA) program. FOQA uses quick access recorders to identify deviations for flight parameters specified in the standard operating procedures (Civil Aviation Authority; CAA, 2003). This information can be used to identify inadequate procedures, ineffective training and briefing, poor team skills, fuel inefficiency and environmental impact, aerodynamic inefficiency, power plant deterioration, and systems deficiencies (Holtom, 2000). To assess the discriminate validity of a safety climate questionnaire in commercial aviation, it will be necessary to obtain safety performance information, and questionnaire responses, from a number of companies. This level of access, and co-operation, will undoubtedly be challenging. Nevertheless, collaboration between rival companies with the goal of improving safety climate has been achieved in other domains, such as the offshore oil and gas industry (Mearns, Whitaker, Flin, Gordon, & O’Connor, 2003). Pooling safety climate data across companies provides a larger sample size for analysis, and allows the discriminate validity to be evaluated.
7. Conclusion

Although there is evidence that the aviation questionnaires that were reviewed in this paper have some construct validity to the extent to which factors identified were consistent with theory, there was a lack of convergence upon a common set of safety climate constructs that were consistent across all of the questionnaires. Support for the discriminate validity of these measures was also found to be lacking. In the absence of evidence supporting a link between safety climate and other measures of performance it will be difficult to convince the aviation industry of the utility of the survey as an accident prevention tool. Nevertheless, it is important to note that it is unlikely that a strong relationship between safety climate and other measures of safety performance will be established (see Clarke, 2006 for a discussion).

The aviation industry has an advantage over many other high reliability industries in that it has international safety regulatory bodies to which all member states must comply (e.g. the European Aviation Safety Agency, EASA). Moreover, EASA and other aviation regulatory bodies such as the Federal Aviation Authority and the Civil Air Navigation Services Organization have recognized the importance of safety culture and are engaged in research exploring, measuring and enhancing safety climate (EUROCONTROL, 2008). We argue that given the involvement of these regulatory bodies, along with the other areas of standardization within the aviation industry, it may be possible to access a large pool of safety climate and safety performance data to allow a valid aviation safety climate tool to be developed that can be shown to have utility in preventing accidents.

Acknowledgements
This work was funded by the Defense Safety Oversight Council (DSOC). All opinions stated in this paper are those of the authors and do not necessarily represent the opinion or position of the U.S. Navy, the Naval Postgraduate School, or the National University of Ireland, Galway.
References


