Effects of Instructor and Evaluator Behaviors and Management Styles on Stress and Pilot Performance in the Aviation Training and Evaluation Environment

Kevin O'Donnell
Department of Aviation
Farmingdale State College (SUNY)
2350 Broadhollow Road, Farmingdale, NY 11735
United States of America

Abstract

The purpose of this study was to determine whether improved instructor/evaluator management styles and behaviors are being incorporated into the training and evaluation environment, in order to reduce non-flight related stress which can degrade pilot performance. A literature review was conducted to evaluate modern aircraft cockpit management principles and related models of occupational stress, to determine whether improved understanding in these areas have been incorporated into instructor/evaluator training. Factors that induce stress in a flight environment can be organized into three categories based on their source. The first category, consisting of operational stressors such as poor weather and mechanical malfunctions, is addressed with cockpit resource management training. The second category, consisting of occupational stressors where substandard performance can impact employment, is addressed with intervention to better match training with the demands of the job. A third category, in the training and evaluation environment, is the anxiety introduced by instructors and evaluators who are perceived as intimidating due to personality traits or records of below-normal grading. Findings indicate that although effective management styles can reduce stress and improve pilot performance in the operational environment, these management styles have not been incorporated by operators for instructor/evaluators in the training environment.

Keywords: Management styles, Occupational Stress, Behavioral marker system, Cockpit resource management, Inter-rate reliability

1. Introduction

There are many factors that can induce stress in a flight environment which may affect a pilot's performance. One category of stressors occurring in the flight environment are operational factors such as poor weather, mechanical malfunctions, and fatigue. These factors cannot be avoided and pilots learn to cope with them through training and experience. A second category of stressors occurs in the aviation training and evaluation environment where a pilot's performance has the added occupational stress of potentially impacting a professional pilot's livelihood. Substandard performance during training or evaluation may prevent a pilot from being hired, result in removal from operations to receive additional training, or cause termination of employment. This form of stress is self-induced but may be mitigated by increasing confidence through the improvement of knowledge and skills to a level that matches the demands of the various tasks that are introduced. A third category of stress is the anxiety that may be introduced by the potentially intimidating attitudes and behaviors of the instructor or evaluator. This form of stress is not part of the flight environment and is therefore deemed nonessential stress. Although much attention has been given to improving pilot interactions and Captain management styles in an operational environment, little attention has been given to the effect of instructor/evaluator cockpit management styles and their potential for inducing additional but nonessential stress in a training or evaluation environment. Variations in pilot examiner behaviors are considered to be significant ("News – FAA," 2018), and there are substantial differences in flight instructor behaviors as well (Crow et al., 2011).

This paper will first examine the development and progression of cockpit/crew resource management (CRM) training and evaluation for pilots (the subjects that instructor/evaluators assess). Following this will be a literature review of occupational stress. Although several models of occupational stress have been published, this paper will focus on those whose designs appear to fit the aviation training and evaluation environment. If the development of non-flight related stress can be predicted, based on specific instructor/evaluator management styles and behaviors that are undesirable, then intervention can be initiated to mitigate or eliminate them (Flin & Martin, 2014). The paper will continue with topics that are addressed in current instructor/evaluator training. Finally, this paper will summarize topics, based on lessons learned from CRM and occupational stress, that can be incorporated into operator procedures and training for instructor/evaluators.

2. Literature Review

2.1.1 Single-Pilot Background

Historically, pilots have been trained, evaluated, and certified in a single-pilot environment. Pilots learned to operate in the flight environment by relying on their own skills, knowledge and decision-making capabilities. In addition, instructors and evaluators in a single-pilot environment frequently do not have experience or training in a multi-pilot crew environment. Therefore, single-pilot environment instructors and evaluators frequently develop their own attitudes, behaviors, and management styles without any standardization. Many times, these instructor/evaluator management styles and behaviors induce additional stress beyond that of the operational flight environment. Pilots can perceive instructor/evaluator behaviors such as body language, tone of voice, questioning of their responses, etc. as intimidating, thus raising their level of stress. These single pilots, many of whom become single-pilot instructors and evaluators themselves, are initially exposed to a flight training and evaluation environment that consists of a wide range of non-standard instructor/evaluator management styles and behaviors where pilots pass or fail based solely on their own capabilities. When these pilots, instructors, and evaluators are then placed in a multi-pilot crew environment, they bring their single-pilot management styles, attitudes, and behaviors with them, which can cause problems.

2.1.2 Crew Resource Management Principles and Training

The potential impact of poor captain management styles, behaviors, and attitudes was recognized in the 1970s after several aircraft accidents which were determined to have human error as a causative factor. An employee of the National Aeronautics and Space Administration (NASA) coined the term Cockpit Resources Management (CRM) (Ciuica & Mihai, 2015). The focus of CRM, later modified to Crew or Company Resource Management, was on improving information management, decision making, communication, and leadership in the flight environment (Gregorich, Helmreich, & Wilhelm, 1990). Much of the early attention on the role of the Captain in the cockpit, a predominantly male environment, focused on male gender identification as a predictor of problem solving, leadership style, authoritarianism, and management style (Chamberlin, 1993). Chamberlin's study found that men's ability to integrate cognitive and emotional activities is a necessity for effective problem solving. Carretta et al. (2014) also found that, in addition to cognitive ability, personality traits are a major predictor of successful Air Force pilot training performance. Therefore, we would expect this positive characteristic of effective personality traits to extend to the multi-crew cockpit. Miller (2006) found that Captains who feel they are the experts, with a need to outperform others, tend to leave their crew behind in terms of information and communication, which increases risk during demanding flight situations. Cockpit Task Management has also become more complex in recent years due to the increased sophistication of technology in the cockpit (Larkins & Parrett, 2009). Although advanced technology increases situational awareness, Larkins and Parrett explain that this also requires a reprioritization of tasks and increased emphasis on CRM skills.

2.1.3 Crew Resource Management Behavioral Marker System

Having verified the importance of effective CRM attitudes and behaviors in today's multi-crew operational flight environment, we now turn our attention to the training and evaluation environment where CRM principles and management styles are trained and evaluated, since the goal is to train and qualify pilots for the multi-crew operational environment. Although the role of the instructor/evaluator is to assess the performance of a pilot, the instructor/evaluator frequently occupies a crewmember position in flight (as pilot in command) and in the simulator, and therefore should actively support CRM principles during the mission. In a training and evaluation environment most instructor/evaluators assess pilot performance using the documentary method, which utilizes extensive note-taking for judging observable actions as well as the related underlying abilities such as decision making (Roth, 2015).

The first step in training and evaluating flight crew use of CRM principles is to identify and organize key attitudes and behaviors into a Behavioral Marker System (BMS) (Flin & Martin, 2001). Although several marker systems have been developed most of them identify CRM skills as nontechnical skills, which are usually divided into two categories: cognitive (including workload management and decision-making) and interpersonal (such as crew coordination and communication). Flin and Martin also identified a third category of skills (technical skills, as opposed to nontechnical CRM skills), which are usually regarded as individual pilot abilities related to procedures, system knowledge, and maneuvers. Hence, we see the training and evaluation of technical skills as a continuing reversion to a single-pilot mind-set even though pilots now operate in a multi-crew environment. We should thus not be surprised to see instructors and evaluators also reverting to nonstandard management styles and behaviors that were developed in the past single-pilot environment. Many times the instructor/evaluator will ask the second (non-evaluated) pilot to withhold checklist reminders, altitude callouts, and deviations from approach paths, which is contrary to CRM training and principles. In regards to technical skills, another study has also measured pilot variance, whereby a pilot stays within tolerances for heading and altitude, but constantly fluctuates within these limits as opposed to stabilized flight (McClernon & Miller, 2007), when not alerted to this behavior by the other pilot. Flin and Martin mention that airlines utilizing an Advanced Qualification Program (AQP) usually combine technical skills and nontechnical skills into their Behavioral Marker System. Flin and Martin also state that this method of combining technical and nontechnical skills into a single Behavioral Marker System (BMS) has been adopted by several airlines which responded to their survey.

One of the first tools for evaluating the application of CRM principles was the Line/LOS Checklist (LLC) version 4.4, as revised by Helmreich, Butler, Taggart, and Wilhelm, in 1997. This checklist allowed observable and measurable attributes to be rated on a four-point scale. The purpose of evaluating CRM attitudes and behaviors was to ascertain whether the knowledge of CRM principles acquired during training was actually transferring to pilot application of CRM skills in the operational flight environment. It is important to note that the rating of CRM skills is applied as an assessment of full crew performance, rather than an assessment of individual pilot CRM skills. However, Flin and Martin (2001) state that there are efforts to assess individual pilots on CRM (nontechnical) skills.

2.2 Occupational Stress in the Aviation Environment

Having been introduced to modern standards for fostering the training and evaluation of CRM principles in pilots, in order to improve crew performance in the multi-crew operational flight environment, we will now focus on the possible impact of undesirable cockpit management styles and additional stress when CRM principles are not fostered by the instructor/evaluator in the training and evaluation environment. According to Vine et al. (2015), researchers have found that high levels of stress are a major cause of pilot errors. Vine states that, before his research, attempts to draw a direct relationship between stress and performance had been mostly unsuccessful. Sarmah (2014), states that occupational stress is a component of the total stress that is related to the work environment, and is defined as the physical, emotional, and physiological responses occurring when the capabilities of the individual may not match the demands of the job. Sarmah states that the concept of occupational stress has evolved to a psychological concept with three components: cause – medium – effect. The cause consists of a factor or factors in the environment that generate stress in an individual (the medium). The effect is the response of the individual to the causative factor or factors in the environment.

More specifically, the cause of stress is a demand, or multiple demands, placed on the flight crew by factors in the environment which require an effective response. The Person – Role Fit model of occupational stress (French, 1973) is one that appears to be appropriate for the aviation training and evaluation environment. According to this model of stress, the match between a person's abilities and the demands of the environment is the major factor in determining the stress placed on an individual, and the greater the gap in this fit, the greater the stress (Sarmah, 2014). As explained by Sarmah, a person's abilities and available resources must meet or exceed the demands of the environment, in order to minimize the stress.

Another model that appears to be appropriate for the aviation training and evaluation environment is the National Institute for Occupational Safety and Health (NIOSH) model (Hurrell & McLaney, 1988). In this model of occupational stress, the emphasis is on the medium's (individual's) perception of the stress factor or factors, in determining the individual's response. That is, individual differences determine how stress factors are subjectively perceived, which determines how the individual responds. According to Sarmah (2014), stress cannot be measured but the effects of stress (responses to stress) can be measured. Sarmah also points out that background stress must be accounted for in an overall assessment of occupational stress.

As mentioned earlier, pilot performance in a training or evaluation situation may be affected by the added occupational stress self-induced by the potential impact of poor performance on a professional pilot's livelihood. Substandard performance may prevent hiring or result in termination of employment. Airline concerns in this area are evidenced by replies to a questionnaire distributed by Flin and Martin (2014). Flin and Martin stated some airlines felt that if CRM assessments (as well as technical assessments) are jeopardy evaluations (pass or fail), then it is desirable to provide an opportunity to resolve the situation with additional training, possibly without a re-evaluation. Flin and Martin also stated some pilots replied that instead of using the evaluation system to find and correct errors, it should be expected that human errors will occur and therefore the emphasis should be on learning to trap (catch) and minimize (correct or mitigate) errors, rather than preventing them.

2.3 Current Practices in Instructor/Evaluator Training.

Having reviewed the purpose of CRM principles which is to improve safety in the operational flight environment, and having reviewed how occupational stress can impact the performance of pilots and flight crews, we now turn our attention to current practices in instructor/evaluator training. Recent attention on instructor/evaluator training since the advent of CRM has focused on inter-rater reliability (IRR), in order to standardize and improve the consistency of measuring pilot performance among instructor/evaluators, to include both technical and nontechnical skills (Holt, Hansberger, & Boehm-Davis, 2002). As mentioned by Flin and Martin (2001), many studies have emphasized that reliability or consistency among instructor/evaluators, when rating pilots, is essential in order to properly validate the results of training and evaluation assessments (Butler, 1991). When pilot ratings become independent of the rater, that is, when raters can be interchanged without appreciably affecting the rating, then a high degree of inter-rater reliability has been achieved. Although great strides have been made for inter-rater reliability in the airline industry, this training has received little attention in general aviation training. As mentioned by Flin and Martin (2001), one study has attempted to identify various reasons for rating differences or biases, in order to mitigate or eliminate them (Hamman & Holt, 1997). Since instructor/evaluators have a high workload during a training or evaluation session, the elements of a BMS assessment tool must be simple and easily measured, in order to be useful (Flin & Martin).

Although instructors and evaluators are trained to not provide leading questions that would assist and improve a pilot's chances of determining a correct course of action for a given scenario, they receive no significant instruction in management style or behavior that would minimize or eliminate the introduction of additional non-flight related stress into the training or evaluation environment. Crow et al. (2011), state that training for flight instructors has changed very little over the years (except for inter-rater reliability in the airline industry), and that little research has been done to investigate and improve flight instructor training. Crow et al. also state that training for flight instructors is based predominantly on role-playing, where the actual instructor acts as a student and the instructor trainee acts as the instructor. Crow et al. point out that instructor behavior changes when they move from role-playing in the training environment to the instruction of actual students.

There are many indications that significant stress during training and evaluation sessions can be induced by factors beyond the flight environment, such as anxiety due to the potential impact on livelihood and the management styles/behaviors of instructor/evaluators. As shown by McClernon and Miller (2011), the impact of stress on performance can be significant in that, without specific training to perform under stress, stress will increase anxiety and reduce performance.

Hamilton (2006) had copilots assess the desirable attributes of Captains using ten descriptors, which were analyzed with multi-dimensional scaling (MDS). Hamilton found that copilots preferred to fly with Captains whom they trusted or were enthusiastic about flying with, rather than Captains with technical expertise, because the preferred Captains fostered team building. Although this study has yet to be repeated for instructor/evaluators, it could be expected that results would also demonstrate preferences for or against instructor/evaluators based on interpersonal skills. Vandermark (1988) administered four personality tests to Captains and copilots and found that personality and attitude characteristics of Captains are strongly correlated with capabilities in Cockpit Resource Management. Vandermark stressed that although personality may not change, attitude change is a basic goal of Cockpit Resource Management training. Again, it could be expected that personality and attitude characteristics of instructor/evaluators would produce similar results in training and evaluation sessions, but studies of this nature have not yet been conducted.

3. Conclusion

The above discussion makes it clear that CRM principles have improved pilot performance in the multi-crew operational flight environment, while factors that increase stress beyond a pilot's abilities to cope with it can decrease pilot performance. One perspective for analyzing pilot stress is to separate it into three categories. Factors in the operational flight environment that increases stress, such as poor weather and mechanical problems, are to be expected and pilots must train to cope with them. Stress in the training and evaluation environment caused by thoughts about the impact of poor performance on livelihood are more difficult to deal with, but are self-imposed and not part of the flight environment. Finally, stress imposed by instructor/evaluator management styles and behaviors are not part of the flight environment and should be reduced or eliminated, just as undesirable Captain management styles, attitudes, and behaviors are being mitigated with CRM training.

After assessing current literature on the topic, there appears to be no significant research on the relationship between instructor/evaluator cockpit management styles and their potential for inducing additional stress, which could impact pilot performance. Without guidance to standardize instructor/evaluator management styles and behaviors during the conduct of training and evaluation sessions, the sessions become open to those management styles that are contrary to CRM principles.

4. Suggestions for Future Research

Further studies could be conducted to measure the effect of instructor and evaluator management styles and behaviors on pilot performance. Independent variables could be established for instructor/evaluator management styles and behaviors, such as three levels of feedback (positive, negative, and neutral). Dependent variables of pilot performance on several tasks, with varying degrees of difficulty, could then be measured.

The survey developed by Hamilton (2006) could be administered to Captains and copilots for assessment of the desirable attributes of instructors and evaluators, using the ten descriptors that he analyzed with multi-dimensional scaling (MDS).

Operators could formalize rules or guidelines for the conduct of their instructors and evaluators during training and evaluation sessions, such as allowing the non-evaluated second pilot to continue standard inputs of checklist reminders, altitude callouts, and deviations from approach paths (in accordance with CRM principles).

Finally, operators might track and evaluate the overall satisfactory/unsatisfactory ratings and scores administered by instructors and evaluators during training and evaluation sessions. Although inter-rater reliability may indicate consistent results would be produced for the various elements of the BMS, this tracking could help detect non-standard instructor/evaluator management styles and behaviors that may be affecting pilot performance by inducing additional stress.

References

Butler, R. (1991). Lessons from cross-fleet/cross-airline observations: Evaluating the impact of CRM/LOFT training. In R. Jensen (Ed.), Proceedings of the 6th Symposium of Aviation Psychology (pp. 326–331). Columbus, OH: Ohio State University.

Carretta, T. R., Teachout, M. S., Ree, M. J., Barto, E. L., King, R. E., & Michaels, C. F. (2014). Consistency of the relations of cognitive ability and personality traits to pilot training performance. *The International Journal of Aviation Psychology, 24*(4), 247-264. doi: 10.1080/10508414.2014.949200

Chamberlin, W. C. (1993). Gender role conflict as a predictor of problem solving, leadership style, authoritarian attributes and cockpit management attitudes (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 9414417).

Ciuica, O., & Mihai, E. (2015). Human error – cockpit vulnerabilities. *Proceedings of the Scientific Conference AFASES 2015, 1,* 43-47. Retrieved from http://www.afahc.ro /ro/afases/2015/afases_2015/air_force/Ciuica_Mihai.pdf

Crow, B., Niemczyk, M., Andrews, D., & Fitzgerald, P. (2011). Role Playing in Flight Instructor Training: How effective is it? *International Journal of Applied Aviation Studies, 11*(1), 1-12. Retrieved from https://asu.pure.elsevier.com/en/publications/role-playing-in-flight-instructor-training-how-effective-is-it

Flin, R. & Martin, L. (2001). Behavioral markers for crew resource management: A review of current practice. *The International Journal of Aviation Psychology 11*(1), 95-118. doi: 10.1207/S15327108IJAP1101_6

French, J.R. (1973). Person-role fit. *Occupational Mental Health, 3*, 15-20. Retrieved from https://eric.ed.gov/?id=EJ080430

Gregorich, S. E., Helmreich, R. L., & Wilhelm, J. A. (1990). The structure of cockpit management attitudes. *Journal of Applied Psychology, 75*(6), 682-690. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11536518

Hamilton, R. A. (2006). Identifying effective leader behavior and its relationship to the exercise of power in the cockpit: a multidimensional scaling study. *International Journal of Applied Aviation Studies, 6*(1), 183-192. Retrieved from https://www.academy.jccbi .gov/ama-800/Spring_2006.pdf

Hamman, W., & Holt, R. (1997). Line operational evaluation (LOE): Air carrier scenario based evaluation. In D. Smith (Ed.), Proceedings of the Human Factors and Ergonomics Society 41st annual meeting (pp. 907–911). Albuquerque, NM: Human Factors and Ergonomics Society.

Helmreich, R., Butler, R., Taggart, W., & Wilhelm, J. (1997). The NASA/University of Texas/Federal Aviation Administration Line/LOS Checklist: A behavioral-based checklist for CRM skills assessment (Version 4.4) [Computer software]. Austin, TX: NASA/University of Texas/Federal Aviation Administration Aerospace Group.

Holt, R. W., Hansberger, J. T., & Boehm-Davis, D. A. (2002). Improving rater calibration in aviation: A case study. *International Journal of Aviation Psychology*, *12*(3), 305-330. Retrieved from http://www.tandfonline.com/toc/hiap20/12/3?nav=tocList

Hurrell, J. J. & McLaney, M. A. (1988). Exposure to job stress - A new psychometric instrument. *Scandinavian Journal of Work Environment and Health, 14*, 27-28.Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/3393871

Larkins, A. & Parrett, M. (2009, May/June). Cockpit task management. *Mobility Forum, 18*(3), 4-6. Retrieved from https://proxy.cecybrary.com/login?url=https://search-proquest-com .proxy.cecybrary.com/docview/213800988?accountid=26967

McClernon, C. K., & Miller, J. C. (2011). Variance as a measure of performance in an aviation context. *International Journal of Aviation Psychology, 21*(4), 397-412. doi: 10.1080/10508414.2011.606765

Miller, D. (2006, July/August). Cockpit/crew resource management and flight discipline. *Mobility Forum, 15*(4), 3. Retrieved from https://proxy.cecybrary.com/login?url=https:// search-proquest-com.proxy.cecybrary.com/docview/213798316?accountid=26967

News - FAA Designated Pilot Examiner (DPE) program under watch (by FSANA and others) (2018). *Flight School Association of North America*. Retrieved from www.fsana.com /news-faa-designated-pilot-examiner-dpe-program-under-watch-fsana-and-others

Roth, W-M. (2015). Flight examiners' methods of ascertaining pilot proficiency. *International Journal of Aviation Psychology, 25*(3-4), 209-226. doi: 10.1080/10508414.2015. 1162642

Sarmah, P. (2014). A critical study of the theories and models relating to occupational stress. *Humanities and Social Sciences Review, 3*(5), 481-493. Retrieved from www.universitypublications.net/hssr/0305/html/V4Z114.xml

Vandermark, M. J. (1988). Airline pilot personality and attitude characteristics: A peer-survey based comparison of the exceptional versus modal airline aviator (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 8809156)

Vine, S. J., Uiga, L., Lavric, A., Moore, L. J., Krasimira, T-A., & Wilson, M. R. (2015). Individual reactions to stress predict performance during a critical aviation incident. *Anxiety, Stress & Coping, 28*(4), 467-477. doi: 10.1080/10615806.2014.986722