ARTICLE IN PRESS

Pergamon

Journal of Safety Research xx (2003) xxx–xxx

www.elsevier.com/locate/jsr

www.nsc.org

# Are pilots at risk of accidents due to fatigue?

Jeffrey H. Goode*

*Federal Aviation Administration, Office of Aviation Policy and Plans, Washington, DC 20591, USA*

Received 23 September 2002; received in revised form 31 January 2003; accepted 27 March 2003

**Abstract**

*Problem*: There is concern in the aviation community that pilot schedules can lead to fatigue and increased chance of an aviation accident. Yet despite this concern, there is little empirical analysis showing the relationship between pilot schedules and commercial aviation accidents. This study attempts to demonstrate an empirical relationship between pilot schedules and aviation accidents. *Method*: Data for human factors-related accidents and pilot work patterns were identified. The distribution of pilot work schedule parameters for the accidents was compared to that for all pilots using a chi-square test to determine if the proportions of accidents and length of duty exposure were the same. If the distributions are the same, then one could infer that pilot human factor accidents are not affected by work schedule parameters. *Results*: The proportion of accidents associated with pilots having longer duty periods is higher than the proportion of longer duty periods for all pilots. *Discussion*: There is a discernible pattern of increased probability of an accident as duty time increases for commercial aircraft pilots in the United States. *Impact on Industry*: The analysis suggests that establishing limits on duty time for commercial pilots would reduce risk. Such a rule is likely to be expensive and could substantially impact the commercial airlines. In return, there is likely to be a reduction in the risk of commercial aviation accidents due to pilot fatigue.


*Keywords:* Pilot flight and duty; Aviation accidents; Fatigue; FAA rulemaking; Risk analysis

## 1. Introduction and problem

There is a great deal of concern in the aviation community that pilot schedules can lead to fatigue and increased chance of an aviation accident. Since 1995, the Federal Aviation Administration (FAA) has been considering a proposed rule that will clarify and simplify flight and duty time limits and rest requirements to ensure that flight crews will receive an opportunity for adequate rest. Current rules have nothing to say about duty limits per se and instead focus on flight limitations and rest requirements. Current flight and rest limits vary based on the type of operation. Flight crewmembers conducting flights under Federal Aviation Regulations, Part 121 (Part 121 refers to the section of the Federal Aviation Regulations that control commercial aviation), for domestic operations are currently limited to 30 flight hours in any seven consecutive days. The seven consecutive day limit for flag operations is 32 flight hours, and there is no weekly limit for supplemental operations. In addition, Part 121 limits flight crewmembers engaged in scheduled domestic operations to 1,000 hours in any calendar year and limits flight crewmembers engaged in scheduled flag operations to 1,000 hours in any consecutive 12 calendar months. Operators are currently required to provide each crewmember a minimum of 24 consecutive hours of rest each week for all domestic, flag, and supplemental operations conducted under Part 121.

Most scientists believe that pilots should have the opportunity for 8 hours of sleep in a rest period. The current regulations do not ensure the opportunity for this amount of sleep. Pilots have filed reports with National Aeronautics and Space Administration's (NASA's) Aviation Safety Reporting System that also document the effects that work patterns have on pilot fatigue and performance (NASA, 1999). Over the 1994 to 1998 time period, there were 227 schedule-related fatigue incidents reported by pilots or approximately 45 per year (NASA, 1999).

Pilot schedule factors have rarely been cited as a cause or a factor in Part 121 aircraft accidents, but this may be due to the fact that there is no test for fatigue either before or after an accident. As such, the National Transportation Safety Board (NTSB) may have been reluctant to cite fatigue as a cause in accidents. The NTSB has explicitly

* Tel.: +1-202-267-9509; fax: +1-202-267-5370.
*E-mail address:* jeff.goode@faa.gov (J.H. Goode).

0022-4375/03/$ - see front matter © 2003 National Safety Council and Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0022-4375(03)00033-1

JSR-00367

recognized that "fatigue was a factor directly leading to this accident" in two aircraft accidents: the Guantanamo Bay (NTSB, 1993, p. 63) accident, and more recently the NTSB has determined fatigue to also be a factor in the 1998 Little Rock accident.

The scientific literature strongly supports the fact that rest is an important factor in human performance. In 1980, in response to a congressional request, the NASA Ames Research Center created a fatigue/jetlag program to study fatigue. In their technical memoranda in 1995, NASA concludes that the average sleep requirement is for 8 hours in a 24-hour period (NASA, 1996). The article by Rosekind, Neri, & Dinges (1997) is another example of a study in the scientific literature that states that most humans require about 8 hours of sleep per night. They argue that fatigue, sleep loss, and circadian disruption created by flight operations can degrade performance, alertness, and safety and that the scientific literature exists that provides important physiological information about the human operator that can be used to guide operations and policy. This article advocates that now is the time for aviation to meet the challenge of managing fatigue in flight operations. In addition, Battelle Memorial Institute did a study for the FAA that reviewed the scientific literature on fatigue. The study found that most researchers advocate an average sleep requirement for adults of 7.5–8 hours per day (Battelle Memorial Institute, 1998).

There have also been studies of pilot fatigue by the military. Neville, Bisson, French, Boll, and Storm (1994) studied airline crews that were exposed to extended work periods, reduced sleep, night work, and circadian dysrhythmia caused by shift work and time zone crossings during Desert Storm. Their research shows that recent sleep and flight histories are correlated with high subjective fatigue levels. They also found a tendency for fatigue to correspond with pilot error. Pilot fatigue can also be studied in aircraft simulators. A recent unpublished study at the Walter Reed Army Institute of Research (Escolas, Santiago, Holland, Kendall, & Russo, 2002) tested flight performance of eight pilots on a flight simulator based on time awake. The study examined pilot air refueling flight performance across 27 hours of continuous wakefulness. Preliminary analysis suggests that severe performance deficits occurred after one night of continuous wakefulness.

The scientific community recognizes that there is a complex relationship between pilot performance, how the performance is impacted by pilot schedules, and safety risk. For example, one can look at pilot work variables to see how they affect crewmember alertness, how alertness affects crew performance under differing workloads and operational environments, and how pilot work variables and alertness combine to affect safety performance that is measured in terms of accidents and incidents. Yet, there is little empirical analysis in the scientific community relating operator fatigue to aviation accidents.

## 2. Method

As part of a FAA rulemaking action on pilot flight and duty time, we conducted an assessment of pilot work practices and the risk of a Part 121 accident. Human factors-related accidents from the 1978 to 1999 time period were identified that involved, at a minimum, substantial damage to the aircraft or serious injuries to those on board. All turbulence-related accidents were excluded, as were accidents that did not have a 72-hour history of pilot activities prior to the accident. There were a total of 55 accidents for which the required data were available.

For this analysis, data were also obtained on pilot work patterns from 10 carriers covering 1 month of flight activity during 1999 (GRA, 2000). These data were used to create profiles of the work patterns of the pilot population. Data for nine carriers were provided by pilot labor unions. We also obtained data on actual pilot utilization from one major Part 121 air carrier that were added to data from the other carriers.

The data provided above were converted into one record for each pilot with a scheduled (or in one airline case, actual) line of flying for the month. Each pilot record tracked pilot activity for every hour in the entire month. The beginning and end of each flight segment were recorded for each pilot and put into a database. Parameters of interest were then calculated such as the length of each pilot's duty period, the amount of flight and duty time per day for each pilot, the amount of rest time, and the numbers of takeoffs and landings each day for each pilot. The analysis tracked these activities in local time as well as base time (defined as the time at the location where the pilot began a multiday trip).

Although some carriers provided data for both captains and first officers, other carriers provided data for captains only. The study used data only for captains to prevent weighing one carrier's responses more heavily than another in the measurement of exposure.

The distribution of pilot work schedule parameters for the accidents was compared to that for all pilots using a chi-square test to determine if the proportions of accidents and pilot duty time exposure were the same. The chi-square is a statistical test to measure the relationship between two different distributions. If the distributions are the same, then one could infer that pilot human factor accidents are not affected by work schedule parameters. However, the study found that there were differences between the two sets of data in some work schedule parameters examined.

To conduct the accident analysis, the exposure and accident data are structured as a $m \times 2$ contingency table to pose the question: Do the proportion of accidents and the proportion of exposure data, with respect to a given schedule-related factor, differ within defined subdivisions of the data? For the chi-square test, the degrees of freedom parameter is determined by the number of strata or subsets into which the data are disaggregated (Fleiss, 1981). In particular, the number of strata for a particular test is denoted by the value $m$, and for this test there would be $m-1$ degrees of freedom.

The chi-square test is one tailed, with critical values associated with each confidence level. The pilot exposure data and the schedule-related data for the set of accidents are categorized into $m$ collectively exhaustive strata.[1] Within each of these strata, a proportion of duty hours (or flight hours, or takeoffs, or whatever exposure parameter is under examination) will fall, with the sum of the proportions equaling 100%. Similarly, within each stratum, a proportion of the accidents will fall, with the accident proportions also summing to 100%. For this statistical test, the null hypothesis is that there is a unitary relationship between the distribution of exposure data and the distribution of accident data, in the sense that the relative frequency of accidents does not change from one exposure stratum to another. In this case, the proportion of exposure observations and the proportion of accidents should, under the null hypothesis, be about the same within each stratum; or, equivalently stated, the ratio of the accident proportion to the exposure proportion should be about one within each stratum.[2] Thus, the chi-square procedure tests the variability or divergence of these proportions from one another.[3] The results are reported below.

## 3. Results

The distribution of pilot work schedule parameters for the accidents was compared to that for all pilots (exposure data) using a chi-square test to determine if the proportions of accidents and exposure were the same. If the distributions are the same, then one could infer that pilot human factor accidents are not affected by work schedule parameters.

Table 1 shows the proportion of duty periods of various lengths for accidents and all pilots. As can be seen, the proportion of accidents associated with pilots having longer duty periods is higher than the proportion of longer duty periods for all pilots. For 10–12 hours of duty time, the proportion of accident pilots with this length of duty period is 1.7 times as large as for all pilots. For pilots with 13 or more hours of duty, the proportion of accident pilot duty periods is over five and a half times as high. The calculated chi-square of 14.89 is highly significant exceeding the 1% significance threshold, as shown in the bottom of the table.

Table 1
Captain duty hours and accidents by length of duty

| Hour in duty period | Captain's hours | Exposure proportion | Accidents | Accident proportion | Accident proportion relative to exposure proportion |
|---|---|---|---|---|---|
| 1–3 | 430,136 | 0.35 | 15 | 0.27 | 0.79 |
| 4–6 | 405,205 | 0.33 | 15 | 0.27 | 0.84 |
| 7–9 | 285,728 | 0.23 | 14 | 0.25 | 1.11 |
| 10–12 | 109,820 | 0.09 | 8 | 0.15 | 1.65 |
| 13 or more | 12,072 | 0.01 | 3 | 0.05 | 5.62 |
| Total | 1,242,961 | 1.00 | 55 | 1.00 | 1.00 |
| Calculated $\chi^2$ | | 14.89 | | 10% $\chi^2$ | 7.8 |
| Degrees of freedom | | 4 | | 5% $\chi^2$ | 9.5 |

As indicated in the Table, 20% of human factor accidents occurred to pilots who had been on duty for 10 or more hours, but only 10% of pilot duty hours occurred during that time. Similarly, 5% of human factor accidents occurred to pilots who had been on duty for 13 or more hours, where only 1% of pilot duty hours occur during that time. There is a discernible pattern of increased probability of an accident the greater the hours of duty time for pilots. The finding is highly significant (0.05% significance level).

## 4. Limitation of analysis

We needed to obtain data on pilot schedules as well as accident data. Data were obtained on pilot work patterns from 10 carriers covering 1 month of flight activity during 1999. These data were used to create profiles of the work patterns of the pilot population. On the other hand, due to the low number of commercial aviation accidents, data on accidents were collected over an extensive time period (1978–1999). This long period was necessary in order to have sufficient accident data with the requisite 72-hour history of pilot activities prior to the accident.

We believe that the two data sets can be compared. Both data sets represent activity that has occurred after deregulation of the aviation industry. We do not believe that pilot work patterns have changed dramatically over the 1978–1999 period. Pilot work patterns over this time span are similar not only due to postderegulation of the aviation industry (with the consequent airline emphasis on the hub and spoke system), but they are also similar because the FAA regulations governing pilot flight, duty, and rest time have not changed much over the period 1978–1999.

[1] Examples of such collections of strata include four 6-hour periods of the day (00:00–05:59, 06:00–11:59, 12:00–17:59, and 18:00–23:59) and 6 "hours of duty period" strata (duty period hours 1–3, hours 4–6, hours 7–9, hours 10–12, hours 13–15, and hours 16 or greater).

[2] For example, if about 65% of flight crew duty hours occur between the hours of 06:00 and 17:59 and 35% between 18:00 and 05:59, then under the null hypothesis the expected proportion of accidents occurring between 06:00 and 17:59 is 65%, with 35% expected to occur between 18:00 and 05:59. This null hypothesis would be tested using the chi-square test with one degree of freedom.

[3] Actually, the use of the chi-square test for this study was a bit different than how the chi-square is generally used. The test is usually used to assess goodness of fit between some data and a particular statistical distribution. We did not have that; we had two sets of data, and we were using the test to ask whether they were similarly distributed.

For our analysis, we compared the 1 month of flight activity data with 20 years of accident data arguing that the distributions of these data sets should be the same if length of duty has no impact on accidents. However, one may wish to be cautious of the empirical findings based on this assumption. If these distributions were not expected to be the same in the first place, then our findings are problematic. With this potential weakness in mind, our findings do strongly suggest that there is an increased probability of an accident as duty time increases, and therefore more stringent limitations on pilot duty time may be appropriate.

## 5. Discussion

While the scientific literature (e.g., Barth & Holding, 1976; Bougrine, Cabone, Mollard, Coblentz, & Speyer, 1995; Colquhoun, 1976; Hamilton, Wilkinson, & Edwards, 1972; Lille, Cheliout, Burnod, & Hazemann, 1979; Lyman & Orlady, 1980) has developed empirical relationships between work patterns and deteriorating performance, there are no direct measures of fatigue or its onset. For example, Bougrine et al. (1995) characterized fatigue as "a set of manifestations generated by intense and prolonged work extending beyond a certain limit" (p. 215). Identifying fatigue in the flight crew exposure data can be done only by inference, using a model comparing schedule-related factors and relative fatigue levels.

The risk analysis provides general support for regulatory proposals to govern duty time. Specifically, the proportion of accidents is higher for more lengthy duty periods than is the proportion of lengthy duty periods in the all-pilot group. This is illustrated in Table 1, where approximately 10% of pilot duty hours are in the 10th or greater hour of a duty period, while 20% of accidents that occur happen in the 10th or greater hour of the pilot's duty period. Similarly, 5% of the accidents occur when a pilot has been on duty for 13 or more hours where only 1% of pilot duty hours occur. These findings suggest that more stringent limitations on pilot duty time may be appropriate.

Using the data from Table 1, exposure as measured by duty hours worked and accidents can be plotted and this can be compared to the accident proportion relative to the exposure proportion. Fig. 1 illustrates this plot and shows that accidents are more prevalent as the length of the duty period increases.

## 6. Conclusion

There is a discernible pattern of increased probability of an accident the greater the hours of duty time for commercial aircraft pilots in the United States. Although the empirical analysis reported above notes that pilot scheduling was not a factor in all of these accidents, it does point to increased risk of accidents with increased duty time and cumulative duty time. The analysis does not indicate any discontinuity at a specific duty time such that it would point to exactly where risk increases significantly. Rather, the data show a relatively constant increase with increased length of work periods. In light of the above, the analysis suggests that establishing limits on duty time for commercial pilots would reduce risk.

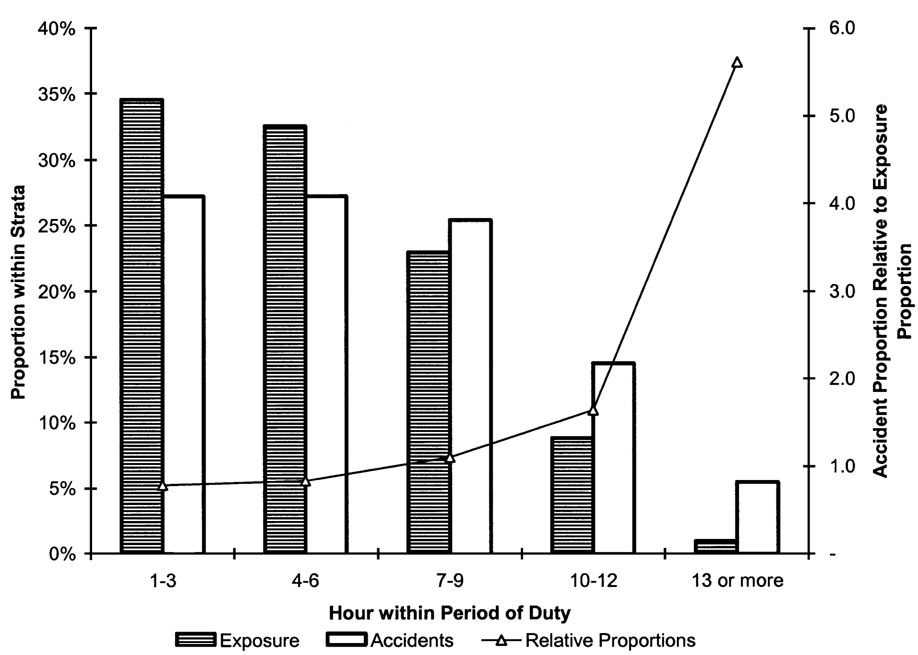


Fig. 1. Captains' duty hours and accidents by length of duty.

## Acknowledgements

The work was conducted as part of a FAA research effort to examine the impacts of a pilot flight and duty rulemaking. The author wishes to thank David Ballard for his contributions to the original research effort and Thomas Smith for his useful comments on this paper. The opinions of the author do not necessarily reflect the official position of the FAA.

## References

Barth, J., & Holding, D. H. (1976). Risk versus effort in the assessment of motor fatigue. *Journal of Motor Behavior*, *8*, 189–194.

Battelle Memorial Institute, JIL Information Systems. (1998). An overview of the scientific literature concerning fatigue, sleep, and the circadian cycle. Prepared for the *Office of the Chief Scientific and Technical Advisor for Human Factors, Federal Aviation Administration*, Washington, DC.

Bougrine, S., Cabon, P., Mollard, R., Coblentz, A., & Speyer, J. J. (1995). Coping with long-range flying. *Recommendations for crew rest and alertness* (p. 215). Blagnac: Airbus Industrie Ed.

Colquhoun, L. (1976). Psychological and psychophysiological aspects of work and fatigue. *Activitas Nervosa Superior*, *18*, 257–263.

Escolas, S., Santiago, S., Holland, D., Kendall, A., & Russo, M. (2002). *Pilot fatigue: The impact of extended wakefulness on simulated aerial refueling flight performance*. Washington, DC: Walter Reed Army Institute of Research.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. Hoboken, NJ: Wiley.

GRA (2000, October 30). *Flight and rest time safety and cost analyses (Phase 3).* Prepared for FAA Office of Aviation Policy and Plans under Work Order No. 1, Contract No. DTFA01-98-C-00096.

Hamilton, P., Wilkinson, R. T., & Edwards, R. S. (1972). A study of four days of partial sleep deprivation. In W. P. Colquhoun (Ed.), *Aspects of human efficiency: Diurnal rhythm and loss of sleep* (pp. 101–103). London: English Universities Press.

Lille, F., Cheliout, F., Burnod, Y., & Hazemann, P. (1979). Effects of aging and occupational activity on active wakefulness. *Gerontology*, *25*(6), 337–344.

Lyman, E. G., & Orlady, H. W. (1980). *Fatigue and associated performance decrements in air transport operations*. (NASA Contract NAS2-10060, NASA Contractor Report No. 166167). Mountain View, CA: Battelle Memorial Laboratories, Aviation Safety Reporting System.

NASA Aviation Safety Reporting System. (1999, June 30). *An analysis of Part 121/135 duty schedule-related fatigue incidents.* Prepared for the Deputy Associate Administrator for Regulation and Certification, Federal Aviation Administration. Washington, DC: NASA.

NASA Technical Memorandum. (1996). *Principles and guidelines for duty and rest scheduling in commercial aviation*. Washington, DC: NASA. Available: http://olias.arc.nasa.gov/zteam/fcp/pubs/p. and g.intro.html.

National Transportation Safety Board. (1993). *Aircraft accident report—uncontrolled collision with terrain, American International Airways flight 808, Douglas DC-8-61, N814CK, U.S. Naval Air Station, Guantanamo Bay, Cuba, August 18, 1993* (p. 63). NTSB/AAR-94/04. Washington, DC: Author.

Neville, K. J., Bisson, R. U., French, J., Boll, P. A., & Storm, W. F. (1994). Subjective fatigue of C-141 aircrews during operation desert storm. *Human Factors*, *36*(2), 339–349.

Rosekind, M. R., Neri, D. F., & Dinges, D. F. (1997). From laboratory to flightdeck: Promoting operational alertness. *Fatigue and Duty Limitations? An International Review. The Royal Aeronautical Society* (pp. 7.1–7.14). London: The Royal Aeronautical Society. Available: http://olias.arc.nasa.gov/zteam/fcp/pubs/raes.html.

Jeffrey Goode is an economist with the Federal Aviation Administration's Office of Aviation Policy and Plans. He is responsible for performing economic and regulatory impact analyses of aviation safety and efficiency matters. Currently he is working on a new rule that will attempt to reduce the effects of fatigue on pilot performance by regulating pilot working and rest periods. He received his PhD from the State University of New York at Stony Brook.