ILL Number: 42408902

OCLC

Borrower: BUF

Receive Date: 5/5/2008 08:10:48 AM

Lending String: CIN,CIN,PIT,*EEM,ORE

Patron: Jui-Feng, Lin

Journal Title: Contemporary ergonomics ; proceedings of the Ergonomics Society's Annual Conference.

Volume: **Issue:**
Month/Year: 2007 **Pages:** 41-46

Odyssey
Charge
Maxcost: $30.00IFM

Article Author: Ergonomics Society (Great Britain). Conference.

Article Title: ; Fatigue in aviation inspection; laboratory and validation studies

Imprint: London ; Philadelphia ; Taylor & Francis

Notes:

Shipping Address:
SUNY AT BUFFALO
ILL
234 LOCKWOOD LIBRARY
BUFFALO, NY 14260

Fax: (716)645-6298
Ariel:
Odyssey: 128.205.243.243

FATIGUE IN AVIATION INSPECTION: LABORATORY AND VALIDATION STUDIES

Colin G. Drury, Brian D. Green & Jui-Feng Lin

Department of Industrial and Systems Engineering, University at Buffalo: SUNY, Buffalo, NY 14221 USA

A number of highly-visible accidents resulting in failures of the inspection system have led to investigation of issues of fatigue in a complex inspection task. Two experiments were performed investigating Shift (0300 vs. 0900 start time), Period (1 vs. 2 hour task) and Time on Task (each 20 min interval) with potential countermeasures of Breaks (None vs. 3 minute break after each 20 min inspection) and Lights (20 vs. 400 Lux). Experiment 1 used participants from the local community and tested under simulated conditions, while Experiment 2 used aviation inspectors in hangars with a similar design. Performance and sleep/fatigue were measured. The two experiments showed a number of similar findings such as the differential effects of breaks for different shifts and periods, but there were differences. The main difference was that a vigilance decrement was found with aviation inspectors but not with community participants.

Introduction

Inspection of airframes and engines of commercial airliners is vital to public safety as all fatigue programs for aircraft are based on the ability of inspection systems to detect defects with a known reliability before they become dangerous (Goranson and Rogers (1983)). Inspection reliability is the result of tasks performed by an often complex human machine system so that all aspects of human/system interaction need to be understood if accurate estimates of reliability are to be obtained. Recently, we have completed a series of studies leading to Good Practices Guides for the principle technologies of aviation Non-Destructive Inspection (NDI), starting from the technique of Fluorescent Penetrant Inspection (FPI), see Drury (2003). A number of highly-visible accidents as a result of failures of the inspection system, such as engine failures at Pensacola and Sioux City have led to a number of initiatives to examine aspects of inspector performance. Because in both of these incidents the inspectors were working under normal conditions, the National Transportation Board asked the Federal Aviation Administration in the USA to investigate issues of fatigue in an FPI task. The current project is the result of that request.

This three year project this project began by examining the literature of fatigue, inspection and vigilance and the findings showed very little similarity between the complex inspection tasks performed in the hangar and the classic laboratory vigilance studies (Drury, Saran and Schultz, 2003). Temporal effects likely to impact inspection were classified as Week Scale (shift work, cumulative fatigue), Day scale (circadian rhythms), Hour scale (vigilance) and Minute scale (sequential effects in repetitive tasks), concluding that the last of these was of only minor importance. As noted by Horowitz, Cade, Wolfe, and Cziesler (2003), vigilance decrements may only be found in tasks having no search element, whereas most inspection tasks do contain search. A survey of 40 NDI inspectors at two airlines showed a rather older population (mean age 47 years) performing NDI tasks over about 2-hour intervals for a 40-hour week, with about 10 min breaks between inspection intervals. This set the stage for developing an FPI simulation task using high-quality digital photographs of six views of each of 63 engine blades. Cracks were inserted using Photoshop as were excess penetrant patches typical of the task of FPI. Inspectors could view each blade from each viewpoint, swab off excess penetrant patches with a swab tool using the mouse, and discover whether a patch in fact concealed a crack. Cracks could sometimes be seen on multiple faces, as in normal FPI. Finally, participants would write up a short description of each crack found using a dialog box similar to the Non-Routine Repair report used on the hangar floor. This simulation was tested with NDI inspectors from one airline who agreed that conclusions from using the simulation would be valid. This simulation was used in two experiments detailed in this paper to provide findings and recommendations to the FAA concerning fatigue and hours of work for NDI inspectors. The original aim was to use the initial 2^5 factorial experiment to determine which interactions were significant, and then explore these in more detail in subsequent parametric experiments. However after Experiment 1 there was little need for experiments with more levels of the significant factors but a great need to confirm the findings on a population of aviation inspectors.

Methodology of experiment 1: industrial participants

This experiment is described more fully in Drury, Green and Henry (2006) so that only a brief version will be given here. The simulation was based on a detailed Hierarchical Task Analysis of FPI (Drury, 2003) using photographs of each of the six faces of 63 different blades. Eighty participants with a mean age of 46 years were recruited from industrially-experienced members of the public. The experiment was designed to test known effects from the vigilance and fatigue literature, comprising Shift (0300 vs. 0900 start time), Period (1 vs. 2 hour task), Defect Rate (7% vs. 15% of blades) and Time on Task (each 20 min interval). Additional factors represented potential interventions: Breaks (None vs. 3 minute break after each 20 min inspection) and Lights (20 vs. 400 Lux). We measured performance (probabilities of detection, PoD, and false alarm, PoFA, time per blade), sleep in the prior 24 hours using an ActiWatch, and rated fatigue and sleepiness from validated scales. Participants had a training session lasting about 3 hours, with two subsequent sessions, one starting at 0300 and the other at 0900. The design was between-participants except for Shift and Time on Task.

Methodology of experiment 2: aviation inspector participants

This experiment was designed to validate the findings of Experiment 1 on a population of aviation inspectors so as to ensure face validity for the subsequent recommendations to the industry. It repeated Experiment 1 with the exception of factors that could not be controlled in a maintenance hangar setting. These were Lights, because we had to take the rooms offered to us by airline partners, and Shifts where scheduling limitations meant that inspectors could not be tested at both 0300 and 0900. The Lights factor was omitted and the Shifts factor became a between-participants factor instead of within-participants. Also we had to relax the 0300 condition as some inspectors on night shifts were only available at 0000 instead of 0300. Four sites were used in the experiment, two at airlines and two at third-party maintenance contractors. All 70 participants performed inspection, and, although some had little specific experience with FPI, all were familiar with the visual inspection task which is the part of FPI tested by the simulation. One site did not run a night shift so all inspectors there were run at 0900. Experiment 2 was thus a 2^3 between-participants factorial experiment for Shifts, Period and Breaks as in Experiment 1 and Time on Task again a within-participants factor at either 3 or 6 levels depending upon the 1 or 2 Hours factor. The same measurements as Experiment 1 were taken with the exception of the ActiWatch data where we did not have access to inspectors for the 24 hours prior to the session. No separate training session was provided, just a familiarization with the FPI program. Most inspectors found the program to be a valid simulation of FPI except for one site where lighting had recently been changed so that the predominant colour of the blades was no longer what the inspectors had been used to.

Comparison of experiments

There were a total of 7 performance measures, 5 self-report scales (of which some such as TLX and SSS had multiple components) and many individual difference variables for each experiment so that the data have only been summarized here. Each variable was subjected to ANOVA analysis using the design appropriate to the experiment. Because of the different designs, Experiments 1 and 2 could not be compared exactly. In addition, with half of the participants only working for a single hour (3×20 min Time on Task data points), no single analysis would be correct. A number of different analyses of subsets of the data were possible and the one reported here is for each hour considered as a separate unit. We have shown a number of the results graphically.

Significant performance effects for Experiment 1 included the factors of Shift, Beaks and Period. For the two-hour participants, breaks at 20 minute intervals helped performance (primarily PoFA) and speed when used during the day at 0900, but had the opposite effect at night (0300). This interaction was eliminated or perhaps slightly reversed for the one-hour participants. Higher throughput was found when the lighting conditions matched the outside lighting. There were almost no correlations between any performance measures and any sleep measures. There was no overall Time on Task effect on PoD but an interaction with Shift, showing a slight decrease in PoD over 2 hours at night (with an end-spurt in the last 20 min) but no effect in the day. These findings suggest that care be taken in extrapolating findings from standard vigilance tasks.

For Experiment 2, there were significant main effects of Time on Task (PoD and speed) Breaks (PoFA) and Period (speed), and interactions between Shift and Breaks for speed and PoFA. The interaction of Shift and Breaks was very similar to Experiment 1, with breaks helping in the day but not at night, primarily for the Hour Period. As in Experiment 1, there were again very few (1 out of 16) significant correlations between performance and sleep measures. The main interest lies in the different Time on Task effects in the two experiments, shown for the performance variables in the composite Figure 1. This figure shows that there was a decrease in PoD as a function on Time on Task for Experiment 2 but not for Experiment 1, whereas for PoFA there was a decrease over time in both experiments. This would give a classic vigilance decrement result for Experiment 2, but not for Experiment 1. Note also that performance in Experiment 2 was better overall than in Experiment 1, with higher PoD and about the same PoFA rates. Note that no statistical test has been performed between the two experiments. In terms of speed, while both experiments start at about the same rate of 12 blades per 20 min (36 per hour) the participants in Experiment 2 became much faster, up to 60 blades per hour for the 2 hour Period and 75 per hour for the 1 hour Period.

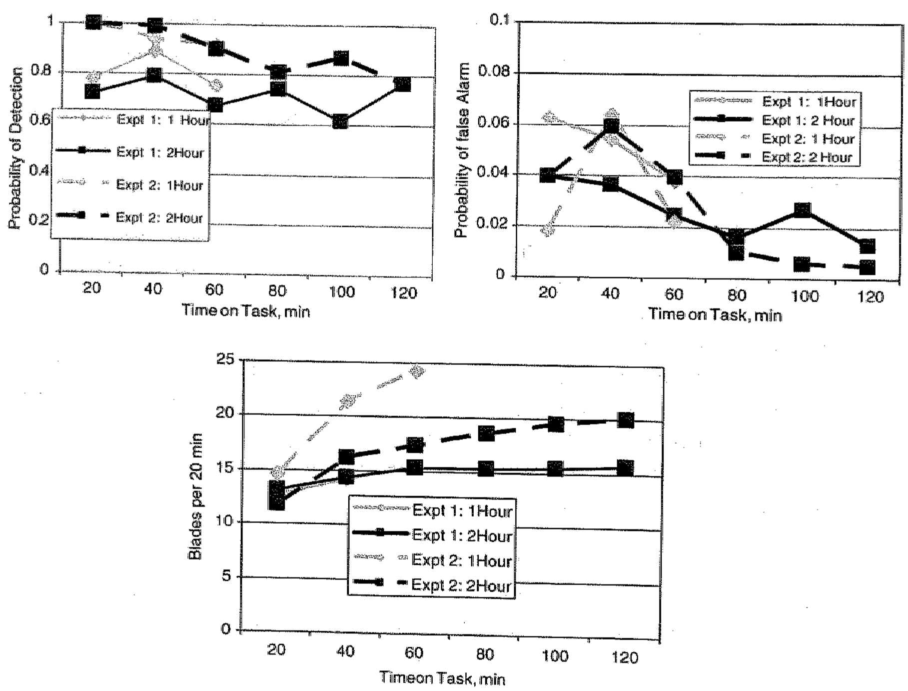

Figure 1. Effects of time on task and hours spent inspecting on three aspects of performance for experiments 1 and 2.

Discussion and Conclusions

What we have not shown here for reasons of space are the many internal consistencies in the data. For example, when each crack was analyzed separately, the PoD was the expected logistic function of crack length and crack contrast, cf. Spencer (2001). The validity of the simulation was again verified by most participants.

There were similarities between the two experiments, but also differences. We can be reasonably sure of our conclusion that breaks help performance, but on the day shift and for the 2 hour period only. At night during a 2 hour period they actually impair performance. The other conclusion from Experiment 1 on lighting matching the external light was not re-tested due to on-site limitations. We again found very few correlations between sleep and performance measures. With those results, it appears that complex NDI inspection has little in common with classical vigilance tasks, but the Time on Task data from Experiment 2 do not support such a conclusion. Here we found a typical vigilance decrement, with decreases in detections accompanied by decrease in False Alarms as the inspection period progressed. It may still be appropriate to declare, as we did after the first experiment that 'These findings suggest that care be taken in extrapolating findings from standard vigilance tasks' as quoted above, but we must be more circumspect in our interpretation.

The main difference between the experiments was one of face validity. Experiment 1 used participants recruited via the local newspaper in Buffalo, NY, attracting a variety of people with different levels of industrial experience. Some were indeed inspectors but many were not, and some number had retired or been laid off from their jobs. They performed the task in our laboratory in calm conditions but with a tape recording of noise from an airline inspection shop playing at correct volume. The inspectors in Experiment 2 were all currently employed by their airlines or repair stations, and performed in an area of the hangar selected by the company. This was typically a training room or break room. The mean ages were similar between the two samples (46 vs. 43 years), but all of the inspectors were male as is typical of the industry while in experiment 1 there were 22% females. Any of these differences could account for the combination of similar and different results, but the practical fact remains that the industry is more likely to place weight on face-valid findings.

In terms of implications for ergonomics theories (not a main aim of either study), the addition of Experiment 2 does suggest that we remain very careful when setting up off-line simulations. We took great care in this study to ensure that the simulation was developed in cooperation with inspectors, and that Experiment 1 was run under close-to-target conditions, but the results of Experiment 2 were still surprising. We would add to our recommendations concerning breaks, shifts and lighting one further recommendation to avoid a second continuous hour of inspection.

References

Drury, C.G. (2003). Human factors and inspection: best practices. *Proceedings of International Ergonomics Association*, Seoul, Korea, August 2003.

Goranson U.F. and Rogers J.T. (1983). Elements of damage tolerance verification, *12th Symposium of International Commercial Aeronautical Fatigue*, Toulouse, France.

Drury, C.G., Green, B.D. and Henry, E.L. (2006) Environmental effects in fluorescent penetrant inspection, Proceedings of the IEA Congress, Maastricht.

Saran M., Schultz J. and Drury C.G. Temporal effects in aircraft inspection: what price vigilance research. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, New Orleans, LA, September 20–24. 2004.

Horowitz, T.S., Cade, B., Wolfe, J.M. and Czeisler, C. A. (2003). Searching night and day: a dissociation of effects of circadian phase and time awake on visual selective attention and vigilance. *Psychological Science*, 14(6), 549–557.

Spencer F.W. (2001) Estimating probability of detection curves from regression data. Materials Evaluation, 59(7), 866–870.