2007

The Effects of Sleep Deprivation on Flight Performance, Instrument Scanning, and Physiological Arousal in United States Air Force Pilots

Fred H. Previc

Nadia Lopez

William R. Ercoline

Repository Citation

Previc, F. H., Lopez, N., & Ercoline, W. R. (2007). The Effects of Sleep Deprivation on Flight Performance, Instrument Scanning, and Physiological Arousal in United States Air Force Pilots. *2007 International Symposium on Aviation Psychology*, 540-545.
https://corescholar.libraries.wright.edu/isap_2007/43

THE EFFECTS OF SLEEP DEPRIVATION ON FLIGHT PERFORMANCE, INSTRUMENT SCANNING, AND PHYSIOLOGICAL AROUSAL IN UNITED STATES AIR FORCE PILOTS

Fred H. Previc
Southwest Research Institute, San Antonio, Texas

Nadia Lopez
Air Force Research Laboratory, Brooks City-Base, Texas

William R. Ercoline
Wyle Laboratories, San Antonio, Texas

The effects of 34 hr of continuous wakefulness on flight performance, instrument scanning, subjective fatigue and EEG activity were measured. Ten fixed-wing pilots flew a series of 10 simulator profiles, and root-mean-square error was calculated for various flight parameters. Ocular scanning patterns were obtained by means of infrared tracking. The results showed that flying errors peaked after about 24-28 hr of continuous wakefulness, in line with peaks in subjective- and EEG-measured fatigue. Instrument scanning was very consistent across pilots but was mostly unaffected by the sleep deprivation.

Introduction

Fatigue due to sleep deprivation is considered a major risk to flight safety, with surveys suggesting that up to half of all pilots have actually “dozed off” while flying (Caldwell & Gilreath, 2002). Fatigue degrades not only basic cognitive abilities but also flight performance, including the ability to maintain designated flight parameters (Caldwell, Caldwell, Brown & Smith, 2004; Caldwell, Caldwell, & Darlington, 2003; LeDuc et al., 1999). Whether the risk to flight safety is due primarily to reduced general cognitive capacity or to flying-specific factors (e.g., stick control or instrument scanning) has yet to be determined.

One possible correlate of the decrements in flight performance is a reduced quality of instrument scanning and other oculomotor functions. Basic eye-movement parameters such as increasing blink rate, decreasing pupil diameter, decreasing saccadic velocity, and increasing saccadic distance have been shown to be correlated with extended wakefulness or time-on-task (Caldwell et al., 2004; De Gennaro, Ferrara, Urbani, & Bertini, 2000; Lavine, Sibert, Goturk, & Dickens, 2002 Morad, Lemberg, Yofe, & Dagan, 2000; Stern, Boyer & Schroeder, 1994). No previous study has investigated changes in pilot instrument scanning behavior with extended wakefulness, although there have been several studies of pilot scanning behavior under normal wakefulness (e.g., Bellenkes, Wickens, & Kramer, 1997; Itoh, Hayashi, Tsukui, & Saito, 1990).

The major purpose of this study was to investigate in experienced United States Air Force (USAF) pilots the effects of over 30 hr of extended wakefulness on flight performance and basic eye-movement parameters and instrument scanning. Another objective was to relate changes in flight performance and instrument scanning to changes in subjective and objective fatigue/arousal, the latter being measured by the scalp-recorded electroencephalogram (EEG).

Methods

Participants

Ten USAF plots served in this study, all of whom were compensated for their off-duty participation. Eight of the 10 (all male) were active-duty pilots, while the remaining two were reserve officers. The average age of the pilots was 34.2 years (ranging from 23 to 46 years), with half of the pilots over 35 years and half 30 years or under. Their average flight experience was 2,806 hr (ranging from 207 hr to 5,800 hr).

All pilots had normal visual and vestibular function. None of them had a history of sleep problems or seizures, was currently taking any psychoactive medication, or was a habitual smoker or caffeine drinker. They refrained from caffeine, alcohol, and other mild stimulants or sedatives while monitored at home on the night before the sleep-deprivation period as well as during the 34 hr of continuous wakefulness in the laboratory.

Procedures

Flight Profile. This study was conducted in the Gyroflight Sustained Operations Simulator (GSOS), a four-axis flight simulator with additional spatial disorientation-producing capabilities. The GSOS possesses motion capabilities in pitch (up to +/-25 deg), roll (up to +/-25 deg), and yaw (up to 360 deg of sustained yaw) and also possesses limited heave (up to +/- 12 cm). It also has a three-channel high-resolution out-the-window visual display that spans 28 deg vertically by 120 deg horizontally. The GSOS aeromodel replicates the T-6 aircraft, with which most of the pilots were familiar, and its reconfigurable instrument panel was also designed to depict as closely as possible that of the T-6 aircraft.

The GSOS flight profile consisted of seven major segments: 1) takeoff at 360 deg and climb to 8,000 ft; 2) right climbing turn to 10,000 ft and 235 deg; 3) wings-level climb to 12,000 ft; 4) right level turn to 180 deg; 5) wings-level descent to 7,500 ft; 6) left descending turn to 4,000 ft and 90 deg; and 7) visual descent and landing at 360 deg. The flight, which required ~19 min to complete, simulated a transition from a dusk takeoff to a nighttime landing and was performed mostly in Instrument Meteorological Conditions (IMC). The exceptions to IMC were during a brief period after takeoff, during a small section of the wings-level climb while pilots searched for traffic, and during the turn to final approach followed by the visual approach and landing. On each segment, the pilot was commanded to maintain a specified set of control or performance parameters, including airspeed (all segments), heading (Segments 1, 3 and 5), vertical velocity (Segments 2, 3, 5 and 6), bank (Segments 2, 4 and 6), and longitudinal bearing and glide slope (Segment 7). On odd-numbered flights, the pilot flew as described above, while on even-numbered flights the pilot flew a mirror profile, beginning with a climb to the left followed by a wings-level climb at 125 deg rather than to the right at 235 deg. On four of the flights (1, 4, 7 and 10), seven spatial disorientation conflicts were thrown in during various segments of the flight, but these were not shown to influence any of the flight performance measures analyzed in this study. The GSOS profile was designed to be semi-automated and required the operator to directly instruct the pilot only during gross flight errors such as the wrong turning direction, wrong course, or wrong heading.

Root-mean-square error (RMSe) was used as the measure of flight performance for each instrument. Each segment had three measures except Segment 1, which only had airspeed and heading. In order to compute composite RMSe values for different flight segments and parameters as well as for the entire flight, all RMSe values were divided by the baseline value (the RMSe in Flight 1) and then converted to log units before averaging.

Eye-movement Recordings. Eye movements were recorded using an Eye-Trac 6000 (Applied Science Laboratories, Cambridge, MA) helmet-mounted infrared system. Head position was measured using a six degree-of-freedom magnetic system (Ascension, Burlington, VT). The position of the eye in the orbit was determined by the relative angle of the pupil and corneal reflectance, as sampled at 60 Hz using an infrared camera. Together, the head and eye signals helped to determine gaze with an error of <0.5 deg.

Many eye-movement records were later discarded because the pupil or corneal images were lost for more than 15% of the samples during flight, there was excessive ocular drift during the flight, or there was excessive blinking. In the end, 71 of 100 records were retained for the scanning analyses and 64 of 100 records were used in the blink and pupil diameter analyses, including at least two each from the early and late flights of nine pilots. (One pilot had excessive lid drooping during all but one of his later flights and was eliminated from the eye-movement analysis altogether).

The eye-movement data were collected during each of the seven segments and corresponded mostly with the periods in which flight performance data were gathered. There were a total of 22 measures obtained from the eye-movement recordings. Five of these were basic measures—average pupil diameter, average blink rate, mean fixation duration, average saccade length, and percentage of dwell times greater than 2 s. There were also 17 measures related to the pilot's instrument scan. These included the percentage of dwells on each of five flight instruments—the electronic attitude director indicator, or EADI, airspeed indicator, altimeter, horizontal situation indicator, or HSI (also known as the heading indicator), and the vertical velocity indicator (VVI)—as well as the percentage of dwells off the instrument panel altogether. There were also 10 measures of transitioning to and from the five flight displays as well as a measure of transitioning to and from the instrument panel as a whole. In determining the dwell and transition patterns for the five major flight displays, their outlines on the instrument panel space were mapped to the calibration space for the eye-tracker and superimposed on the scan pattern from each flight.

EEG Recording. EEGs were recorded using a GRASS-Telefactor Instruments Aurora recording system (West Warick, RI) running TWin™ collection and analysis software. EEGs were recorded from gold-cup electrodes at two sites (C_z and P_z) and were referenced to linked-mastoid electrodes, while an additional ground lead was attached to the scalp. EEGs were recorded with cutoff filters at 1 Hz and 70 Hz and were digitized at 200 Hz. Preflight EEGs were recorded in the GSOS control area just before the pilot's gaze was calibrated. A total of three 3 s epochs, each selected because they were free of any obvious muscle or blink artifacts, were selected from each preflight EEG record. EEGs from each electrode site and condition were analyzed separately for their Fourier amplitude in each of three bands: delta (1.5–3.0 Hz), theta (3.0–8.0 Hz), and alpha (8.0–13.0 Hz).

Overall Schedule

Pilots arrived the night before the beginning of the continuous wakefulness period for initial training on two versions of the flight profile and on various cognitive tests. After their monitored sleep, pilots arrived back in the laboratory at either 0730 or 0830 and flew the GSOS profile for a final practice flight and received additional training on the cognitive tests. At mid-morning on Day 1 of continuous wakefulness, each pilot had EEG electrodes attached and, after a period of rest and lunch, began the experiment. The pilots were run in tandem, with pilot #1's first flight beginning at 1200 and pilot #2's first flight beginning at 1300. Successive flights were run at 3-hr intervals

Immediately after completing a flight, each pilot went to a testing area where he completed the various cognitive tests and two subjective fatigue surveys: 1) the Profile of Mood States (POMS), a 65-question survey (McNair, Lorr, & Droppelman, 1981) and 2) the Visual Analog Scale (VAS), a computerized scale involving ratings of various dimensions with a line and pointer (Penetar et al., 1993). The "fatigue-inertia" dimension on the POMS and the "sleepy" dimension on the VAS were considered the two most direct measures of subjective fatigue.

Results

Subjective Fatigue Ratings

The ratings on the POMS fatigue dimension and the VAS sleepiness scales are shown in Figure 1. The two measures paralleled each other fairly well, although their correlation was somewhat modest ($r = +.58$). Pilots reported little subjective fatigue/sleepiness over the first four flights, a large increase in fatigue/sleepiness over the next two flights (i.e., in the early morning hours), and continued high subjective fatigue over the final four flights. Both measures varied highly significantly ($p < .001$) across sessions.

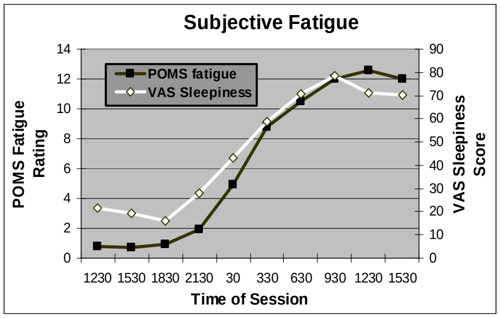

Figure 1. Changes in POMS and VAS scores across flight session.

Flight Performance

The changes in the composite RMSe during the continuous wakefulness interval are shown in Figure 2. A slight improvement in flight performance occurred over the first five flights, which was followed by a sharp increase in RMSe in the early morning hours that leveled off slightly during the daytime hours on Day 2. The early-morning increase in RMSe was sharper and slightly delayed relative to the rise in subjective fatigue. Despite the decrease in precision flying performance during the final five flights, gross flight errors requiring GSOS operator intervention were fairly rare and occurred no more than three times in any session (summed across all pilots), except for the first session in which six flight corrections were made. Overall, an approximate 25% decrement in performance occurred from the peak performance in Session 4 (1230 of Day 1) to the maximum deficit that occurred in Session 8 (0930 of Day 2). However, this effect masked large individual differences, with three pilots showing a slight reduction in error from the five early to five late flights and two pilots (both of whom showed at least one documented in-flight micro-sleep) showing decrements of over 30%. Age of the pilot did not significantly interact with flight performance across sessions.

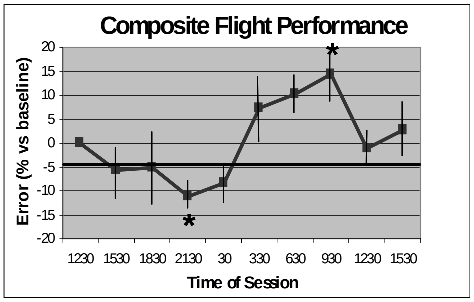

Figure 2. Composite flight performance across sessions. Error bars represent the standard deviation of percentage change from baseline flight (#1) across pilots. Asterisks denote significant deviation from baseline.

A repeated-measures ANOVA, using log RMSe, was performed for the composite flight performance score across the 10 flight sessions. The effect of session was only marginally significant (*p*=.07), partly because not all flight parameters comprising the composite score were significant. In addition to the analysis of composite flight performance, four individual flight parameters that were measured in at least three segments (airspeed, heading, vertical velocity, and bank) were subjected to individual ANOVAs. Vertical velocity varied significantly across the 10 flight sessions (*p*=.005), as did airspeed (*p*=.033), but performance differences across sessions for heading and bank were nonsignificant.

Eye Movement Analyses

As shown in Figure 3, pilots overall were strikingly consistent in their scanning behavior, For examples, the dwell percentage on the EADI across pilots ranged from 43% to 57% (mean of 48.8%) and the mean percentage of transitioning to and from the EADI as a percentage of all instrument transitions ranged from 61% to 76% (mean of 69.8%). The dwell-time percentages on each of the five instruments as well as the off-panel locus were also consistent with the type of maneuver being performed. The percentage of overall dwell time away from the instrument panel was greatest for the visual approach, second-most for the takeoff, and third-most for the segment in which pilots briefly looked for traffic. Dwelling on the EADI, meanwhile, was greatest in the three turning maneuvers (Segments 2, 4 and 6), and VVI dwell percentages were greatest in the four climbs and descents in which a specific vertical velocity speed was mandated (Segments 2, 3, 5 and 6).

Eye-movement measures were collapsed across the five early flights and the five late ones. The differences between the early and late averages were then analyzed by a set of paired *t*-tests, one for each of the 22 eye-movement parameters. Overall, the instrument scanning patterns for each pilot were remarkably similar for the early and late flights, as even slight variations in the pattern of scanning across individual pilots were similarly evidenced in both the early and late flights (Figure 3). The average amount of time spent on the five major flight instruments (the EADI, airspeed indicator, altimeter, HSI, and VVI) never differed by more than 12% in any individual case from the early to late flights, and only 2 of the 10 transitional probabilities among the five instruments differed by more than 10%. Of the 22 measures, only two turned out to significantly differ from the early to late flights: percentage of dwell time on the HSI (*p* = .04), and transitioning between the HSI and the EADI (*p*=.01). However, the actual differences were very slight in these measures, with the HSI dwell percentage increasing only slightly from 11.87% to 12.48% and the HSI-EADI transition percentage increasing from 20.25% to 20.86%. No dwell or transition measure achieved statistical significance for the VVI and airspeed indicator—the flight instruments associated with the largest and second-largest changes in RMSe across sessions. However, the two numerically greatest decrements in transitioning from the early to late flights both involved the VVI—a 32.68% reduction in airspeed-VVI transitioning and a 30.1% drop in HSI-VVI transitioning.

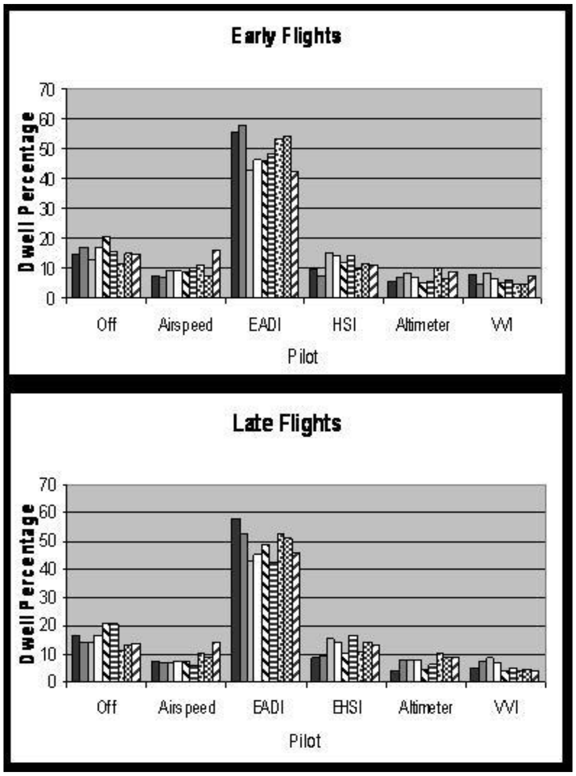

Figure 3. Dwell time on specific instruments for early and late flights and each of nine pilots.

It should be noted that changes in none of the basic eye-movement parameters of interest (blink rate, pupil diameter, average saccade length, average fixation time, and percentage of long dwells) even approached significance (*p* values of .31, .16, .22, .70, and .48, respectively). In the case of blink rate, the lack of significance was less due to the overall increase (18%) in the blink frequency in the later flights than to the enormous variability in blink rate change (six of the nine pilots showed increases of 50% or more in their blink rate, including one who had an increase of 163%, whereas two pilots showed very little change and one pilot even showed a decrease of >50%).

EEG Analyses

Generally, alpha decreased as the period of wakefulness increased, whereas delta and theta activity increased The effect of flight session was significant only for theta and alpha at P_z (p=.04 for both). Across flights, the correlation between P_z alpha and theta was -.49.

Discussion

The results of this study demonstrate that over 30 hr of continuous wakefulness significantly degrades flying precision in a simulator. This effect proved highly variable across pilots and cannot be explained by a comparable deterioration in instrument scanning. In fact, instrument scanning was remarkably unaffected by the pilot's sleep deprivation status, which otherwise produced a large increase in subjective fatigue, fatigue-related changes in EEG activity, degradation of flight performance, and decreases in vigilance and cognitive capability.

The flight performance deficits began in the early morning hours and peaked at 0930 before waning by noon on the second day. The magnitude of the flight performance nadir relative to the first flight was not large (~15%), although the decrement was ~25% when compared to the peak of flight performance on the fourth flight. The increased flight error in the early morning hours of the second day was paralleled by an early-morning decline in EEG alpha activity and early-morning peaks in subjective fatigue (as measured by the VAS) and EEG theta activity, all of which exhibited their maximum changes at 0630 or 0930. The subsequent rebound in flight performance suggests that flight performance was affected not only by sleep deprivation but perhaps by the circadian cycle, although subjective fatigue ratings and EEG alpha-to-theta did not rebound nearly as much as flight performance. It should be noted that both the time of maximal flight performance deficit and the subsequent rebound were generally consistent with the results of Caldwell et al. (2003) for helicopter pilots and Caldwell et al. (2004) for F-117 pilots.

The decline in flight performance masked very large individual differences among pilots. For example, the mean percentage change in RMSe was 4.8% from the five early to the five late flights, but two pilots showed changes of >30% while three pilots actually showed improvement from the early to later flights. The existence of individual differences in "fatigue resistance" has long been noted in the literature and has been attributed mostly to recent sleep history, personality factors such as introversion, or even baseline arousal (Caldwell et al., 2005). The segments in which most of the micro-sleeps were observed and which seemed most problematic in terms of flight error were those with long wings-level climbs and descents. The fact that bank control was much less severely affected by sleep deprivation than was vertical velocity—which was commanded during the wings-level climbs and descents—indicates that pilots may have been able to increase their arousal level somewhat while turning.

Contrary to the previous literature, we did not observe changes in any basic eye movement parameter, even though our average blink rate of 0.35/s, our average pupil diameter of 4.79 mm, and our average fixation duration of 435 ms were all in the ranges reported previously. The failure to find significant changes in basic ocular measures may be partly attributable to the large variability across pilots in their early versus later oculomotor behavior. For example, the blink rate changes with sleep deprivation ranged from a decline of 52% to an increase of 163%, with seven of the nine pilots showing increases.

Our basic ocular scanning results appear to be highly consistent with those of other studies (for example, our pilots looked ~50% of the time at the EADI, as opposed to ~60% in Itoh et al.'s 1990 study of commercial airline pilots), remarkably consistent across pilots, and sensitive to the expected demands of particular flight segments, with fixation on the EADI greater during turning and fixation on the VVI greater during climbs and descents. By contrast, the resilience of instrument scanning during the extended wakefulness period proved highly surprising, given the altered scanning reported by others during sleep deprivation (e.g., De Gennaro et al., 2000), the large increase in subjective fatigue and the significant decline in flying precision. The maintenance of normal scanning even after extended wakefulness suggests that instrument scanning even in pilots recently graduated from USAF undergraduate pilot training is a highly practiced behavior that is largely resistant to fatigue. Hence, the deterioration in flight performance during fatigue, either generally or for specific instrument parameters such as airspeed and vertical velocity, is more likely due to poorer information processing and decision-making rather than to a change in scanning behavior per se.

Acknowledgments

We thank Joe Fischer and Phil Tessier of General Dynamics, Terrell Scoggins of AFRL/HEP, Josh Borah of Applied Science Laboratories, and Col. Dave Tubbs of the Advanced Instrument School at Randolph Air Force Base for their assistance. This study was funded by the Air Force Research Laboratory.

References

Bellenkes, A. W., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine*, *68*, 569-579.

Caldwell, J. A., Caldwell, J. L., & Darlington, K. K. (2003). Utility of dextroamphetamine for attenuating the impact of sleep deprivation in pilots. *Aviation, Space, and Environmental Medicine*, *74*, 1125-1134.

Caldwell, J. A., Caldwell, J. L., Brown, D. L., & Smith, J. K. (2004). The effects of 37 hours of continuous wakefulness on the physiological arousal, cognitive performance, self-reported mood, and simulator flight performance of F-117A pilots. *Military Psychology*, *16*, 163-181.

Caldwell, J. A., & Gilreath, S. R. (2002). A survey of aircrew fatigue in a sample of U.S. Army aviation personnel. *Aviation, Space and Environmental Medicine*, *73*, 472-480.

Caldwell, J. A., Mu, Q., Smith, J. K., Mishory, A., Caldwell, J. L., Peters, G., et al. (2005). Are individual differences in fatigue vulnerability related to baseline differences in cortical activation? *Behavioral Neuroscience*, *119*, 694-707.

De Gennaro, L., Ferrara, M., Urbani, L., & Bertini M. (2000). Oculomotor impairment after 1 night of total sleep deprivation: A dissociation between measures of speed and accuracy. *Clinical Neurophysiology*, *111*, 1771-1778.

Itoh, Y., Hayashi, Y., Tsukui, I., & Saito, S. (1990). The ergonomic evaluation of eye movement and mental workload in aircraft pilots. *Ergonomics*, *33*, 719-733.

Lavine, R. A., Sibert, J. L., Gokturk, M., & Dickens, B. (2002). Eye-tracking measures and human performance in a vigilance task. *Aviation, Space and Environmental Medicine*, *73*, 367-372.

LeDuc, P. A., Riley, D., Hoffman, S. M., Brock, M.E., Norman, D., Johnson, P. A., et al. (1999). *The effects of sleep deprivation on spatial disorientation* (USAARL Report No. 2000-09). Ft. Rucker, AL: United States Army Aeromedical Research Laboratory.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1981). *Manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Service.

Morad, Y., Lemberg, H., Yofe, N., & Dagan, Y. (2000). Pupillography as an objective indicator of fatigue. *Current Eye Research*, *21*, 535-542.

Penetar, D., McCann, U., Thorne, D., Kamimori, G., Galinski, C., Sing, H., Thomas, M., & Belenky, G. (1993). Caffeine reversal of sleep deprivation effects on alertness and mood. *Psychopharmacology*, *112*, 359-365.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human Factors*, *36*, 285-297.