2021-05

Physiological and Cognitive Performance in F-22 Pilots During Day and Night Flying

Combs, Elizabeth K.; Dahlman, Anna S.; Shattuck, Nita L.; Heissel, Jennifer A.; Whitaker, Lyn R.

Aerospace Medical Association

Combs, Elizabeth K., et al. "Physiological and cognitive performance in F-22 pilots during day and night flying." Aerospace medicine and human performance 92.5 (2021): 303-311.

http://hdl.handle.net/10945/69560

Physiological and Cognitive Performance in F-22 Pilots During Day and Night Flying

Elizabeth K. Combs; Anna S. Dahlman; Nita L. Shattuck; Jennifer A. Heissel; Lyn R. Whitaker

BACKGROUND: Many workers routinely transition between day and night shifts—including pilots, where night flights are commonly considered more stressful. The physiological toll from this transition is not fully understood, though fatigue is a factor in many aviation accidents. This research investigated the changes in physiological markers of stress and cognitive performance as F-22 pilots transitioned from day flying to night flying.

METHODS: There were 17 fully-qualified F-22 pilots who took part in a 2-wk data collection using salivary swabs, wrist-worn activity monitors, the National Aeronautics and Space Administration-Task Load Index (NASA-TLX) inventory, and a go/no-go (GNG) test.

RESULTS: No differences were found in comparing day and night flying on the GNG reaction time/accuracy, NASA-TLX scores, or sleep quantity. Cortisol levels were significantly higher than civilian levels in all experimental conditions and control days. Participants had higher than predicted cortisol levels postflight in the day-flying condition and lower than predicted cortisol levels postflight in the night-flying condition, relative to levels from control day patterns. We also found smaller changes in cortisol (pre- to postflight) in the day-flying condition for those with more F-22 experience. Finally, we found a negative correlation between Perceived Stress Scale scores and age of pilots ($r = -0.72$).

DISCUSSION: We hypothesized that the night-flying environment would be more stressful, but our results disputed this claim. Our results suggest day flying elicits more of a stress response; however, a larger sample size is required to verify results. Preliminary findings of potential stress adaptation may suggest stress adaptation in the F-22 community needs further investigation.

KEYWORDS: stress, performance, workload, cortisol, pilots.

Combs EK, Dahlman AS, Shattuck NL, Heissel JA, Whitaker LR. ***Physiological and cognitive performance in F-22 pilots during day and night flying*. Aerosp Med Hum Perform. 2021; 92(5):303–311.**

Many workers must move between day and night shifts. Here, we study a particularly stressful job with such a schedule: F-22 pilots. Sleepiness and fatigue are particularly important for this community; human error is a causal factor in upward of 70–80% of military aviation accidents.[25,30] Understanding how and why pilots make errors is a critical part of the mishap investigation process. Given the roles of the stress system and sleep in attention and focus, we examine whether shifts in flying times are associated with changes to sleep, perceived workload, stress system activation, and cognitive performance.

Recent informal interviews with high-performance aircrew suggest that night flights are significantly more stressful than day flights because of the lack of visual information available to the pilot. Simply stated, if pilots can see a hazard, they can avoid it. At night, they are less likely to perceive and avoid threats to their safety. Hence, their subjective stress levels could be significantly higher for night flights. Changes in sleep and stress levels may affect pilot ability to pay attention and focus.[4,14,18] However, physiological and cognitive comparisons between day and night flying to validate that claim have not been investigated directly.

Using nonflying days as a baseline, we examined whether sleep, workload, physiological stress, and performance differed between day-flying and night-flying days. The driving hypothesis

---

From the Naval Postgraduate School, Monterey, CA.

This manuscript was received for review in September 2019. It was accepted for publication in February 2021.

Address correspondence to: Jennifer A. Heissel, Naval Postgraduate School, 1 University Circle, Monterey, CA 93943, USA; jaheisse@nps.edu.

DOI: https://doi.org/10.3357/AMHP.5508.2021

was that a measurable difference existed in day vs. night flying. A secondary hypothesis was developed to examine evidence of a stress system adaptation within the F-22 community. The normal operating cycle for pilots flying F-22 aircraft shifts between day and night flight operations. Daytime sorties (training missions) usually commence at or around 07:00, which means that the pilots must be at the squadron roughly 2 h prior to takeoff time to prepare for the mission. Because crew rest protocol dictates a maximum 12-h duty day, aircrew typically return home at 17:00.

Night sorties usually require pilots to be on base between 13:00 and 14:00, with sorties commencing after sundown, typically 18:00 or later local time. Night flight rotations are required for aircrew to maintain night-vision goggle proficiency, night landing currency, and their Command Mission Ready rating. Night flying weeks typically occur for 2 wk once a quarter.

One of the primary outcomes in the present paper involves the response of the hypothalamic-pituitary-adrenal (HPA) axis and its main hormonal output, cortisol. Changes in cortisol are related to changes in attention and focus,[14,29,33] which matter for pilot performance. Basal cortisol levels follow a strong circadian rhythm throughout the day.[13] Cortisol is relatively high upon waking, increases substantially (50–60%) in the 30–40 min after waking (called the cortisol awakening response, or CAR), and subsequently declines across the day.[3,26]

The HPA axis periodically activates to respond to acute stress; such responses are considered adaptive.[7,8,31] In particular, the CAR increases in the presence of acute daily stressors,[2,5,11] while it is generally lower in the presence of traumatic stress, particularly when accompanied by posttraumatic stress symptoms or disorders.[27,34] Chronic stress is also associated with lower waking cortisol levels and slower drops in cortisol across the day.[1,9] At the same time, HPA activation also declines with repeated exposure to the same stressor, indicating habituation, although the relationship is complex.[12] In the case of F-22 pilots, we may then expect higher daily cortisol levels following more stressful piloting experiences, relative to less stressful flights or control cortisol levels. However, more experienced pilots may be less physiologically stressed than newer pilots due to habituation.

METHODS

Subjects

The research was designed as a prospective, quasi-experimental design within participants' observational study. Volunteer participants gave written informed consent and the study was conducted in accordance with IRB protocol NHRC.2018.0007. Participants collected cortisol samples on 2 control days and during 1 d in each experimental condition. There were 2 wk of sleep data collected.

Cognitive performance and workload data were collected 1 d during each experimental condition. Performance and salivary data were collected on the same day in each experimental condition. The first week was a normal flying week in which participants flew during daytime hours. The second week of data collection was a night week in which takeoff times were shifted to the evening hours. Participants were instructed to not alter their natural awakening time for the purpose of this study. Data was collected from February 26, 2018, to March 9, 2018, at the F-22 squadrons located at Hickam Air Force Base, HI, USA.

Volunteering to participate in the study were 17 male F-22 pilots. Inclusion criteria required the pilots to be actively flying during the 2-wk data collection period, not medically grounded, with no known disorder that affects HPA axis function. Inclusion criteria were self-reported and assessed during recruitment briefs. No access to medical records was required. Participants could take part in all or parts of the study depending on their comfort level.

Equipment

Wrist activity monitor. Participants were issued a wrist activity monitor (WAM), which was a nontransmitting activity monitor manufactured by the Philips Respironics company. WAMs were preprogrammed at the Naval Postgraduate School (NPS) for passive data collection following the recruitment brief and were worn continuously throughout the 2-wk data collection. Participants were instructed to wear the WAM on their nondominant wrist and only remove it for hygiene purposes. WAMs were returned at the conclusion of the data collection period. Two participants turned in their WAMs early due to changes in their flying schedule.

Data from the WAMs were downloaded to allow for calculation of total daily sleep, defined as the amount of sleep obtained in a 24-h period. In the actigraphic recordings, 1-min epochs were used. The 24-h period was calculated from midnight to midnight of each day and was separated into two categories: the average sleep for the day-flying week and for the night-flying week.

NASA Task Load Index. Perceived workload was measured using the computer NASA Task Load Index (NASA-TLX). The paper and computer NASA-TLX versions vary in sensitivity, with the computer-executed assessments incurring more workload.[24] We accepted this additional measured workload in the experimental design as it reduced computational error by researchers. Apple iPads were used to administer the NASA-TLX with preloaded study and participant information. Weightings of each of the six dimensions (mental demand, physical demand, temporal demand, performance, effort, and frustration level) and a workload score were automatically calculated by the NASA-TLX application.

Saliva Samples

Saliva samples for analysis of cortisol (sCort) were collected using the SalivaBio Oral Swab (Salimetrics, Carlsbad, CA, USA). Participants were instructed to refrain from eating or drinking for 30 min prior to salivary sample collection. Instructions and a list of items to avoid were provided to participants for their reference.

Participants collected saliva by placing the SalivaBio Oral Swab sublingually for 2 min and then placing the swab in the collection tube. Participants were instructed to not touch the swab either in the placement of the swab or in the removal of the swab to place in the tube. Participants were instructed to not chew on the swab during the sample collection. No stimulants were given to increase saliva production. Swabs were then transferred to collection tubes and either stored in home fridges (if self-collected) until they could be turned into NPS researchers or immediately in NPS-provided coolers if collected in the presence of researchers. All sample swabs were refrigerated for transport from Hawaii to California. Samples were then transferred to −80°C freezers at NPS for storage until analysis. Samples were frozen for approximately 1 mo prior to analysis.

Biochemical Analysis

The saliva samples were thawed overnight and the swabs were placed into sterile syringes. Saliva was extracted by pushing on the plunger to squeeze out liquid volume into 1.7-mL tubes. Tubes were refrozen in two −40°C freezers for approximately 20 d until analysis was completed. Saliva samples were thawed 24 h prior to biochemical analysis. When thawed, samples were vortexed and then centrifuged at 1500 rpms for 15 min prior to analysis.

Salivary cortisol concentrations were found using a Salimetrics Salivary Cortisol ELISA kit. The test required 25 μL of volume for analysis and had a sensitivity of $0.007\ \mu g \cdot dL^{-1}$ for the concentration range of $0.012–3.00\ \mu g \cdot dL^{-1}$. The serum-saliva correlation is 0.91.[28] Interassay variation was mediated by analyzing participants on the same microplate. All samples were run in duplicates when the volume allowed. Intra-assay coefficient of variation remained within an acceptable limit for all eight plates tested.

Go/No-Go

A computer Go/No-Go (GNG) test developed by the Naval Aviation Medial Research Unit at Dayton was used to evaluate changes in cognitive performance pre- and postflight in both day- and night-flying conditions. The GNG tests reaction time to one stimulus while refraining from responding to incorrect stimuli (response inhibition). The task is similar to decision making in the aircraft where "friend vs. foe" identification is required. The test included 180 trials, 20% of which were no-go stimuli. The interstimulus interval, or the time between trials, was 0.5–1.0 s. The random nature of the interstimulus interval prevented the participants from establishing a predictable response rhythm. Participants were instructed to press the space bar if a "go" indication was received and to refrain from pressing the space bar if a "no-go" indication was received. The indications were either a yellow or purple square in the middle of the laptop computer screen. The "go" and "no-go" indications were randomly assigned each time a participant signed into the computer. After a brief practice session, participants began the 5-min timed test. This procedure was the same for both the pre- and postflight data collection in both conditions.

Procedure

A recruitment brief was conducted in accordance with IRB protocol NHRC.2018.0007 and took place on February 23, 2018, during squadron administration briefs with approval by the Squadron Director of Operations. Participants were briefed on the background, purpose of the study, and various components of the study. Individual recruitments were conducted until February 28, 2018, to accommodate senior leadership schedules. Study volunteers were issued consent forms, the presurvey questionnaire, a preprogrammed WAM, salivary collection sampling materials, and instructions about the salivary collection process.

The prestudy questionnaire consisted of five validated stress and fatigue scales. Because research suggests that the interaction of home and job stress can negatively impact flying performance,[10] a Cohen Perceived Stress Scale (PSS) and Holmes-Rahe Stress Readjustment Rating Scale (SRRS) were administered to participants in addition to the Epworth Sleepiness Scale (ESS), Pittsburgh Sleep Quality Index (PSQI), and the Morningness-Eveningness Questionnaire-Self Assessment version (MEQ-SA).

Participants completed saliva sample collection during three conditions: nonflying, day flying, and night flying (**Fig. 1**). The nonflying condition was used as a control for establishing the participants' normal diurnal rhythm of sCort. Saliva sampling began February 26,2018 and consisted of five samples per day for 4 nonconsecutive days, of which 2 d were control days (consecutive), 1 d was a day-flight day, and 1 d was a night-flight day. All subjects completed their control and day-flight samples prior to night-flight samples. This adhered to the current schedule within the squadron.

The first sample for each day was to be taken immediately upon waking. The second was scheduled to be collected 30 min

Non-flying Day (Control)		Day Flight		Night Flight	
Cortisol Sample	Other Tests	Cortisol Sample	Other Tests	Cortisol Sample	Other Tests
Wake		Wake		Wake	
Wake+30 min (CAR)		Wake+30 min (CAR)		Wake+30 min (CAR)	
Before lunch		Before lunch		Before lunch	
		Pre-Flight	GNG		
		Post-Flight	GNG, NASA-TLX		
Afternoon		Afternoon		Afternoon	
Evening		Evening		Evening	
				Pre-Flight	GNG
				Post-Flight	GNG, NASA-TLX

Fig. 1. Schematic of approximate saliva sampling timing for the two nonflight (control) days, one day-flight day, and one night-flight day, as well as timing of the go-no go (GNG) and National Aeronautics and Space Administration-Task Load Index (NASA-TLX) tasks.

after waking to establish the CAR.[32] The three additional samples were to be taken at intervals throughout the day (approximately before lunch, afternoon, and evening). Participants were given flexibility in selecting the exact time for these sample collections to allow for flight times, simulator training, and additional constraints on individual schedules. Text message reminders by research personnel were sent at the request of participants preceding a sample collection day.

Participants were told not to alter their crew rest requirements or their normal sleeping schedule for the purpose of the study. Therefore, participants were required to collect several of their own salivary samples. The self-collection of participants' samples is common in salivary biomarker analytics and historically has high compliance rates.[15,22]

Two additional saliva samples were collected during flying days, one preflight and one postflight, for a total of seven sample schedules on those days. Pilots completed the preflight salivary sample within 30 min of planned engine start-up and the postflight sample immediately after returning to the aircrew flight equipment room. NPS researchers supervised the pre- and postflight salivary sampling.

In both day- and night-flying conditions, participants were instructed to take a GNG both before and after the scheduled flight. Participants elected to complete the GNG after the debrief but before donning their aircrew flight equipment (prior to their saliva samples). Participants completed the NASA-TLX assessment following the GNG only postflight. Participants completed the pair-wise comparisons each time they were administered the NASA-TLX.

Statistical Analysis

The sleep amounts measured by WAMs were compared for each day across both conditions (e.g., Monday of day-flying week was compared to Monday of the night-flying week) using a two-tailed paired sample *t*-test. Data from the WAMs were also used to establish an average awakening time in the absence of an awakening salivary sample for plotting the predicted sCort concentration curve.

NASA-TLX scores were calculated by the iPad NASA-TLX application. Scores were compared within subjects using a two-tailed paired sample *t*-test on Microsoft Excel software.

The CAR was calculated by subtracting the awakening samples from the awakening + 30 min sample (in $\mu g \cdot dL^{-1}$); this required participants to have provided both samples that day. The CAR value was compared across conditions using a two-tailed paired sample *t*-test.

Slopes of salivary cortisol levels were compared to civilian populations along with predicted sCort concentrations in relation to hours after awakening. We used the 50th percentile readings for men ages 31–40 from the CIRCORT database, as this group was most similar in age to our male pilots.[21] We fit a quadratic curve to cortisol levels as a function of hours since awakening using the database sCort levels. To see if observed cortisol levels of F-22 pilots differed from those predicted using the civilian-predicted fitted curve, we used a Wilcoxon sign test, where the test statistics, x, is the number of participants with more positive than negative residuals.

Because the exact timing of cortisol sampling differed across days, we used the cortisol observations from the control days to predict participants' "expected" diurnal patterns. Specifically, we fit sCort levels for the control (nonflying) samples as a quadratic function of hours since awakening on the available sample. We predicted the "expected" cortisol level for each pre- and postflight sample on flying days based on the individuals' hours since wake. We then compared the real observed value to this predicted value using a two-tailed *t*-test.

Reaction time and accuracy were automatically calculated by the GNG using MATLAB software. A two-tailed paired sample *t*-test was used to compare GNG results pre- and postflight in both day- and night-flying conditions. All statistical analyses were performed using JMP software (SAS, Cary, NC, USA) unless otherwise noted, and the level of statistical significance was set to $P < 0.05$ (two-tailed).

RESULTS

Participating in varying degrees in the study were 17 F-22 pilots. Participants were on average 36 yr old (SD = 6.29 yr), with the following averages ($\bar{X}$): total flight hours $\bar{X} = 1745$ (SD = 944 h); F-22 flight hours $\bar{X} = 576$ (SD = 347 h); and night-flying hours $\bar{X} = 204$ (SD = 158 h). Age was positively correlated with total flying hours ($r = 0.75$, $P = 0.004$), F-22 h ($r = 0.65$, $P = 0.016$), and night flying hours ($r = 0.67$, $P = 0.012$). Of the participants, 13 of 17 returned the MEQ-SA, ESS, PSQI, and PSS for a 76.5% return rate.

MEQ-SA test scores ranged between 16 and 86. A score of 41 or below indicates an evening type person, who tends to function better at night. Scores of 59 and above indicate morning types, who function better in the morning. A score in the range of 42–58 indicates an intermediate type, with neither morning nor evening characteristics. Four respondents (31%) were categorized as morning types, eight respondents (61%) were intermediate types, and one respondent (8%) was categorized as an evening type.

ESS scores fell into one of the following categories: lower normal daytime sleepiness (scores 0–5), higher normal daytime sleepiness (scores 6–10), mild excessive daytime sleepiness (scores 11–12), moderate excessive daytime sleepiness (scores 13–15), and severe excessive daytime sleepiness (scores 16–24). Seven participants' scores (54%) indicated normal daytime sleepiness levels. Four participants (30%) showed higher than normal daytime sleepiness. One individual showed moderately excessive daytime sleepiness (8%) and one individual showed excessive daytime sleepiness (8%).

The PSQI is rated on a 0–21 scale, with higher numbers indicating increasingly poor sleep quality. Scores greater than or equal to 5 indicate poor sleep quality. Six participants (46%) scored 5 or above, indicating poor sleep quality.

PSS scores fell into the following categories: low stress (score of 0–13), moderate stress (scores of 14–26), and high stress (scores of 27–40). There were 12 participants (92%) who indicated low stress. Only one participant (8%) indicated moderate stress.

Of the participants, 11 of 17 returned the SRRS, for a 64.7% return rate. An SRRS score of 150 or less indicates a low likelihood of developing a stress-related illness within the next 2 yr; all but one of the respondents fell within this range. One respondent scored 164, indicating a 50% chance of developing a health-related breakdown within the next 2 yr.

We evaluated the PSQI, ESS, SRRS, MEQ-SA, and PSS for a relationship with age. The PSS was the only one of the five tests that was significantly correlated with age. Scores on the PSS decreased with age ($r = -0.72$, $P = 0.005$).

There were 15 participants who wore WAMs during the day-flight week, and 14 who wore WAMs for all or a portion of the night-flying week. Participants slept on average 7.15 h during the day week (SD = 52.2 min) and 7.01 h during the night week (SD = 85.7 min). The difference in the expected number of hours slept between the 2 wk was not statistically different [$t(14) = -0.29$, $P = 0.77$]. The average sleep efficiency was 84% across the 2 wk and was positively correlated with the average minutes spent asleep ($r = 0.58$, $P = 0.021$). Postflight in the day-flight condition, 12 participants completed the NASA-TLX rating with an average workload score of 65.86 (SD = 14.74, $N = 12$). In the night condition, eight participants completed the NASA-TLX postflight with an average workload score of 66.33 (SD = 16.83, $N = 8$). Only six participants completed the NASA-TLX postflight in both the day- and night-flying conditions. There were no significant differences in workload ratings across the two conditions [$t(5) = 0.57$, $P = 0.59$].

A total of 17 pilots took part in the saliva collection with varying degrees of adherence to protocols. Only four participants completed the full sample collection (all control, day flying, and night flying samples) in the data collection period. There was no difference in the control to day-flight CAR responses [$t(6) = -0.85$, $P = 0.43$, $N = 7$], the control to night-flight CAR responses [$t(5) = -2.12$, $P = 0.087$, $N = 6$], nor the day- to night-flight CAR responses [$t(2) = -1.83$, $P = 0.209$, $N = 3$]. Noncompliance limited the number of observations available for this analysis. We compared participant control (nonflying) cortisol levels to the fitted CIRCORT curve. Pilots in the sample had higher than expected cortisol levels than what was predicted from the CIRCORT database curve. Cortisol levels tend to increase with age.[16] As a sensitivity analysis to verify that this pattern held, we used a more conservative sample excluding participants over the age of 40. The Wilcoxon tests showed that for both the total sample of 13 ($x = 13$, $P < 0.001$) and the age restricted sample of 9 ($x = 9$, $P = 0.002$), the observed median cortisol levels were significantly greater than predicted (**Fig. 2**). The results indicate that $x = n$, meaning that both sample sizes had more positive than negative residuals in the comparison.

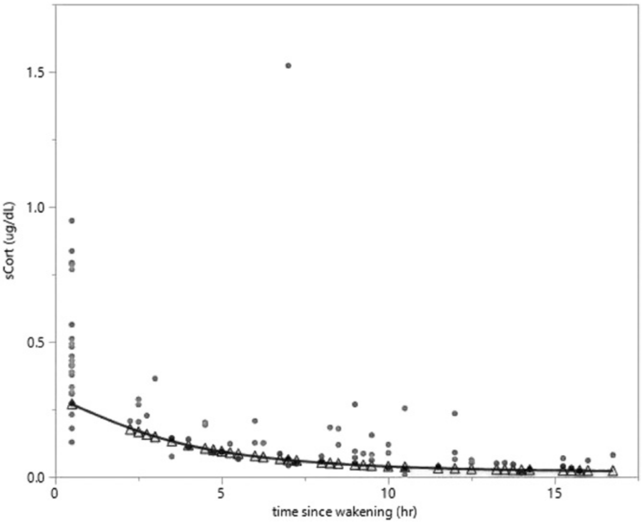

Fig. 2. Plot of actual and CIRCORT-predicted sCort levels ($\mu g \cdot dL^{-1}$) during control days by hours since awakening. Circles: sCort actual (control day); triangles: CIRCORT predicted sCort.

We compared actual cortisol levels to expected CIRCORT database sCort levels for person-specific pre- and postflight time since waking on the flying days. Actual sCort levels were significantly higher than CIRCORT predicted sCort levels in the preflight day-flying condition [$t(13) = 3.05$, $P = 0.009$], the postflight day-flying condition [$t(12) = 3.63$, $P = 0.003$], the preflight night-flying condition [$t(10) = 2.64$, $P = 0.025$], and the postflight night-flying condition [$t(10) = 3.78$, $P = 0.003$]. We fit sCort levels for the control (nonflying) samples as a quadratic function of hours since awakening using 83 data points from the nonflying days from 13 pilots. Using this curve, we predicted the typical sCort concentrations throughout the day, which had the expected downward slope (**Fig. 3**). We then compared this predicted level to the actual observed day- or night-flight cortisol level for both pre- and postflight using a two-tailed paired sample t-test.

Preflight actual and predicted sCort levels were not different in the day-flying condition ($N = 13$) [$t(13) = 0.96$, $P = 0.35$] or

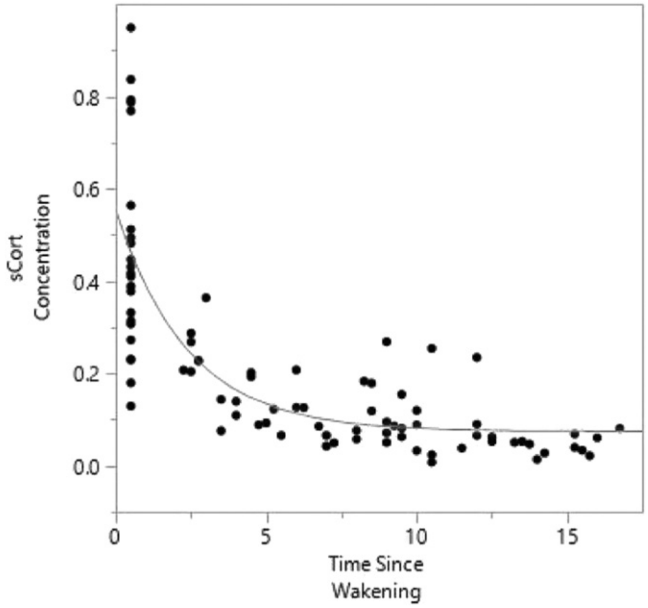

Fig. 3. Fitted diurnal curve for sample by hours since awakening.

night-flying condition ($N = 10$) [$t(10) = 0.25$, $P = 0.80$] relative to the slope predicted from the control condition. Postflight actual sCort levels were significantly higher in the day-flying condition [$t(12) = 2.19$, $P = 0.049$] and significantly lower in the night-flying condition [$t(9) = -2.30$, $P = 0.05$] relative to the slope predicted from the control condition (**Fig. 4**). The change in cortisol levels, preflight to postflight, was not significant in the day-flying condition [$t(11) = -1.33$, $P = 0.21$] or the night-flying condition [$t(10) = -1.19$, $P = 0.26$].

We evaluated flying hours and the PSQI, PSS, ESS, MEQ-SA, SRRS, and NASA-TLX for a relationship with the cortisol response change from pre- to postflight. A correlation existed only between the magnitude of the cortisol change (pre- to postflight) in the day-flying condition and F-22 h ($r = -0.89$, $P = 0.001$), such that individuals with fewer hours in the F-22 experienced a larger increase in cortisol from preflight to postflight. There were 14 individuals who completed the GNG with varying levels of compliance. We compared the conditions using a two-tailed paired sample *t*-test.

Eight participants completed both a pre- and postflight GNG test in the day-flying condition. There were no significant changes in no-go accuracy [$t(7) = -0.57$, $P = 0.58$], total test accuracy [$t(7) = -1.13$, $P = 0.29$], reaction time of the correct response in seconds [$t(7) = 0.38$, $P = 0.71$], or reaction time of an incorrect response [$t(7) = -0.19$, $P = 0.85$] from pre- to postflight on the day-flying days.

Seven participants participated in both a pre- and postflight night GNG test. There were no differences in accuracy of the inhibitory response [$t(6) = -0.63$, $P = 0.54$], total test accuracy [$t(6) = -0.82$, $P = 0.44$], reaction time of the correct response [$t(6) = -0.11$, $P = 0.91$], or reaction time of an incorrect response [$t(6) = 0.459$, $P = 0.66$] from pre- to postflight on the night-flying days.

Four participants completed the preflight GNG tests in both the day- and night-flying conditions. The preflight inhibitory accuracy was worse at night and suggestive of a change [$t(3) = -2.49$, $P = 0.09$], but did not meet the threshold for significance. Neither total test preflight accuracy [$t(3) = -1.93$, $P = 0.14$], preflight reaction time of a correct response [$t(3) = -0.318$, $P = 0.77$], nor preflight reaction time of an incorrect response [$t(3) = 0.34$, $P = 0.76$] changed from the day to the night postflying condition.

Seven participants completed both postflight GNG tests in day- and night-flying conditions. There were no significant changes in the postflight inhibitory accuracy [$t(6) = -1.37$, $P = 0.21$], total test accuracy [$t(6) = -1.47$, $P = 0.19$], reaction time of a correct response [$t(6) = -0.74$, $P = 0.48$], or reaction time of an incorrect response [$t(6) = -0.865$, $P = 0.42$] from the day to the night postflying condition. **Table I** shows the results.

DISCUSSION

F-22 pilots are high-functioning, resilient individuals tasked with operating the U.S. Air Force's fifth generation fighter in support of missions across the world. Pilots are required to transition from day operations to short-term night operations to ensure

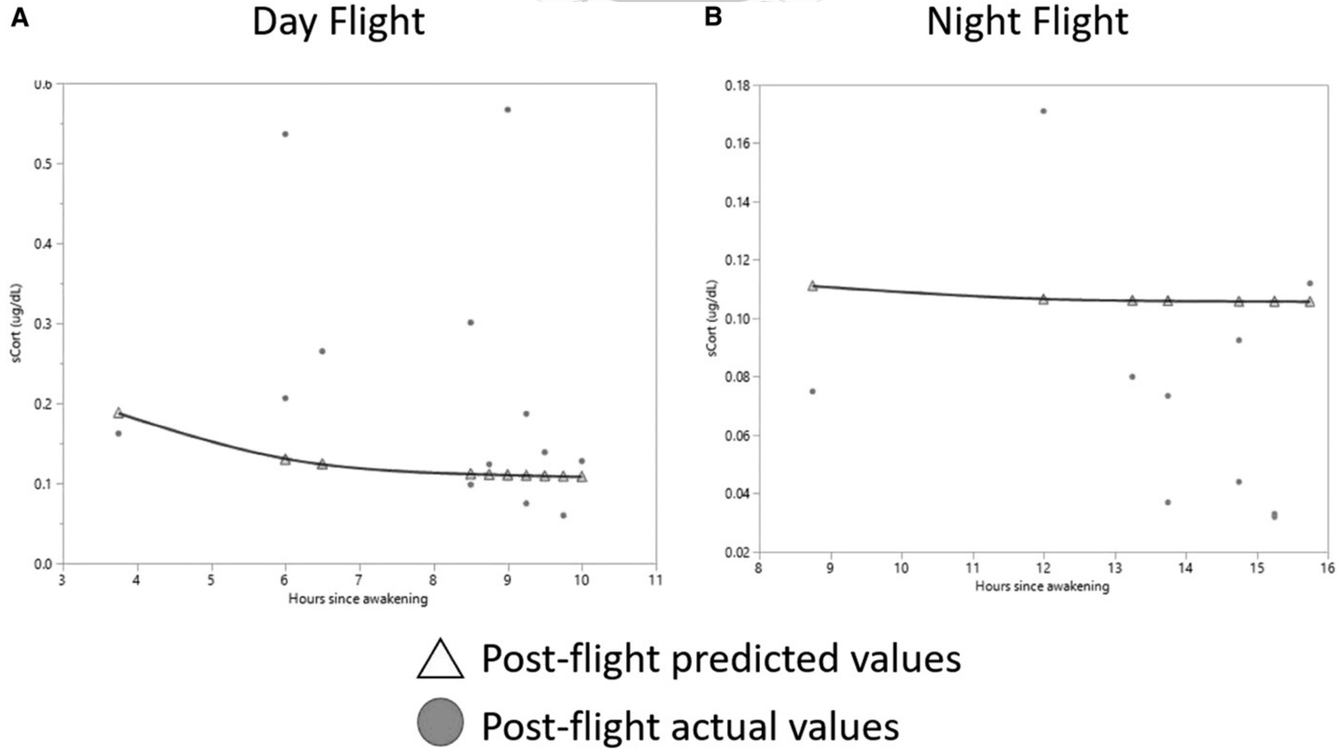

Fig. 4. Plot of actual and participant-developed curve predicted sCort levels ($\mu g \cdot dL^{-1}$) **A.** during the day postflight by hours since awakening and **B.** during the night postflight by hours since awakening.

Table I. Go/No-Go Testing Results by Condition.

CONDITION	MEAN NO-GO ACCURACY	MEAN TOTAL TEST ACCURACY	MEAN RT (s) (GO)	MEAN RT (s) (INCORRECT RESPONSES)
Preflight (d)* (*N* = 8)	0.9166 (SD = 0.088)	0.9825 (SD = 0.017)	0.3135 (SD = 0.025)	0.253 (SD = 0.031)
Postflight (d)* (*N* = 12)	0.8957 (SD = 0.072)	0.9768 (SD = 0.014)	0.3149 (SD = 0.03)	0.2631 (SD = 0.031)
Preflight (n)† (*N* = 7)	0.8134 (SD = 0.13)	0.9602 (SD = 0.029)	0.3049 (SD = 0.04)	0.2566 (SD = 0.028)
Postflight (n)† (*N* = 9)	0.7931 (SD = 0.175)	0.9555 (SD = 0.034)	0.2994 (SD = 0.033)	0.2595 (SD = 0.025)

* Day-flying condition; †night-flying condition.

that their training remains current; however, the impact on aviator performance from this short-term shift remains unknown. We quantified the stress burden on 17 volunteer participants as they made a short-term transition to night flights to maintain their readiness status. Performance, workload, fatigue, and stress patterns were monitored in the participants over 2 wk as the F-22 aviators transitioned from day- to night-flying operations.

Neither sleep nor subjective workload ratings differed across flying conditions. Participants had average workload scores of approximately 66 postflight in both the day- and night-flying condition. These results were similar to NASA-TLX scores of commercial airline pilots executing a landing with the loss of the autopilot in a simulator,[35] but significantly higher than F/A-18 pilots executing an instrument landing with multiple cautions and warnings in a simulator ($\bar{X} = 39.04$, SD = 7.86).[20]

Participants had higher cortisol levels than would be predicted from civilian men of similar ages on control days, day-flight days, and night-flight days. Day-flying appeared to cause a physiological stress response. In keeping with expected behavior, neither day- nor night-flight days showed elevated cortisol levels preflight and failed to support an anticipatory effect in the stress response. Participants' day-flight cortisol levels followed their control-day cortisol rhythms up until postflight, at which point they became elevated. Participants had higher than predicted cortisol levels postflight relative to what would be predicted by their control-day cortisol rhythms. Within the day-flight days, participants maintained the same level of cortisol pre- to postflight, which in itself is evidence of a stress response. Normal cortisol levels decrease throughout the waking day.[3,21] In the absence of a stress response, participant cortisol levels would have dropped postflight, following their normal diurnal rhythm. Similar analyses failed to support a stress system activation in the night-flying condition. Participants had significantly lower than predicted cortisol levels postflight at night-flying days.

The cortisol results were surprising, given the general perception that night-flying is more stressful to pilots. Several factors could have influenced these results. First, the sample shared their runway with an international airport that was commonly congested with commercial air traffic during the day. Because Hawaii is a tourist destination, sightseeing aircraft also used the airspace. It is possible that daytime air traffic may have contributed to a stress response. Less commercial air traffic at night may require less attention devoted to aircraft deconfliction.

Second, day flying requires a visual scan to "see and avoid." Radar assists in the identification of threats, but the visual system is the primary sensory input device in the daytime flight environment. Pressure on the pilot to visually identify objects in an already congested airspace may have contributed to activation of the stress response. At night, visual scans are limited to instrumentation cues. While some maneuvers need to be confirmed under night vision goggles, the majority of the visual information about the environment is digitally processed. Pilots are both assisted by and restricted to their instrument feedback. These factors may have contributed to the lower than expected cortisol levels in the night-flying condition.

Third, the physical requirements of day flying are demanding and also may have played a role in the cortisol spike. The F-22 is a 9-G_z capable aircraft. Several pilots reported experiencing high or repeated G forces during their day-flight sorties. G forces require a sustained muscle contraction, termed the anti-G straining maneuver, to maintain consciousness. Because night flying is G-limited, it does not require the same physical effort. The additional physical effort of day flying may have caused the higher than predicted cortisol levels postflight. Additionally, several pilots in the sample had indicators of poor sleep quality. Fatigue has been known to increase cortisol production. Still, there are a number of gaps in our understanding of why cortisol is elevated postflight in day flight and further investigation is warranted.

Two significant correlations suggest some form of stress adaptation within the F-22 sample. First, participant age was negatively correlated with their level of perceived stress on the PSS; that is, younger participants reported higher levels of stress. The PSS has generally shown a trend to decrease with age.[6] Notably, the F-22 pilots' distribution of PSS scores in this study is lower than the distribution averages of both men ($N = 968$, $\bar{X} = 15.52$, SD = 7.44) and of 35–44 yr olds ($N = 331$, $\bar{X} = 16.38$, SD = 7.07), in contrast to the elevated cortisol levels.

Second, we found a negative correlation between the magnitude of the cortisol change (from pre- to postflight) and the number of flight hours in the F-22, further suggesting an adaptation effect. As participants flew more hours in the F-22, the magnitude of their physiological stress response was smaller. Some experienced participants were nonreactors and maintained a steady cortisol decline even during flight. This suggests that with increased exposure to the stressor, in this case flying, the less stressed the individual would become.

These two points suggest potential adaptation, not burnout, within the F-22 sample.[17] This physiological adaptation falls in line with other research into repeated stress exposures.[12] One possible reason for the adaptation would be the sample's characteristics, or the pilot selection process that brings in individuals with high resilience. F-22 pilots are an elite group who are categorized as high-functioning, confident personalities who are often exposed to high-threat environments and show evidence of physiological adaptation to recurring stress like similar highly resilient populations.[19] We used a modified go/no-go test to measure reaction time and response accuracy in the participants. Despite differences in postflight cortisol across conditions, we found no difference as participants transitioned from day- to night-flying operations. On average, participants had an inhibition test accuracy of about 90% in both pre- and postflight in the day-flying condition. Reaction time of the correct response averaged approximately 0.31 s and reaction time of the incorrect response averaged approximately 0.26 s preflight and postflight in the day-flying condition. In the night-flying condition, participants scored on average approximately 80% inhibition accuracy, 0.30 s reaction time of the correct response, and approximately 0.25 s of the incorrect response in both pre- and postflight. Reaction times were similar to a civilian population using a 20% no-go test.[23] However, the F-22 pilots had a significantly higher average inhibition accuracy in both flying conditions (90% and 80%) than the 66% inhibition accuracy reported in the civilian population.[23] The authors acknowledge that the small sample size may have impacted the significance.

Many questions emerged from this study; continued research should focus on the impact of elevated daytime cortisol and whether it could pose a long-term health risk in this participant sample. If, indeed, further research confirms that job requirements are placing pilots' health at risk, we need to delve deeper. Are F-22 pilots required to accept stress as a normal part of their duty? Do these results hold up in less busy airspaces? Are these changes unique to military aviation? These questions and others offer several opportunities for further research. The results of the present research suggest that the full burden of F-22 flying on pilots is not yet understood. We hypothesized that the night-flying environment would be more stressful on the aviator. While more research is required to support the results found this in this study, it appeared the opposite is true: day-flying is more stressful. The authors acknowledge that the small sample size and location of the data collection may have contributed to the significance. More research on the topic is warranted. Understanding the stress burden to F-22 aviators during the transition between day- and night-flying operations is an important aspect of future mishap prevention efforts.

ACKNOWLEDGMENTS

The authors thank Rebecca Fosha for her help with data collection; Dr. Doug Granger; and Kara Blacker at Naval Aviation Medial Research Unit at Dayton for her assistance and loan of the GNG software. The views expressed in this document are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

Financial Disclosure Statement: This research was performed while the author A.S. Dahlman held an NRC Research Associateship award at the Naval Postgraduate School and author J. A. Heissel was supported by the Naval Postgraduate School's Research Initiation Program. Funding for the study itself was provided by the Naval Advanced Medical Development program. The authors have no competing interests to declare.

Authors and Affiliations: Elizabeth K. Combs, Maj., USAF, BSC, B.S., M.S., Flight Commander, Aerospace Physiology, 344 TRS, JBSA-Lackland, TX, USA; Anna S. Dahlman, M.Sc., Ph.D., Swedish National Road and Transport Research Institute, VTI, Gothenburg, Sweden; and Nita L. Shattuck, Ph.D., M.Ed., Jennifer A. Heissel, MPP, Ph.D., and Lyn R. Whitaker, Ph.D., M.S. Statistics, Associate Professor, Operations Research Department, Naval Postgraduate School, Monterey, CA, USA.

REFERENCES

1. Adam EK. Emotion-cortisol transactions occur over multiple time scales in development: implications for research on emotion and the development of emotional disorders. Monogr Soc Res Child Dev. 2012; 77(2):17–27.
2. Adam EK, Hawkley LC, Kudielka BM, Cacioppo JT. Day-to-day dynamics of experience—cortisol associations in a population-based sample of older adults. Proc Natl Acad Sci USA. 2006; 103(45):17058–17063.
3. Adam EK, Kumari M. Assessing salivary cortisol in large-scale, epidemiological research. Psychoneuroendocrinology. 2009; 34(10):1423–1436.
4. Barger LK, Ayas NT, Cade BE, Cronin JW, Rosner B, et al. Impact of extended- duration shifts on medical errors, adverse events, and attentional failures. PLoS Med. 2006; 3(12):e487.
5. Chida Y, Steptoe A. Cortisol awakening response and psychosocial factors: a systematic review and meta-analysis. Biol Psychol. 2009; 80(3):265–278.
6. Cohen S, Janicki-Deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. J Appl Soc Psychol. 2012; 42(6):1320–1334.
7. Del Giudice M, Ellis BJ, Shirtcliff EA. The Adaptive Calibration Model of stress responsivity. Neurosci Biobehav Rev. 2011; 35(7):1562–1592.
8. Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol Bull. 2004; 130(3):355–391.
9. Doane LD, Chen FR, Sladek MR, Van Lenten SA, Granger DA. Latent trait cortisol (LTC) levels: reliability, validity, and stability. Psychoneuroendocrinology. 2015; 55:21–35.
10. Fiedler ER, Rocco PD, Schroeder DJ, Nguyen KT. The relationship between aviators' home-based stress to work stress and self-perceived performance. Oklahoma City (OK): Federal Aviation Administration, Oklahoma City Aeronautical Center; 2000. Report No.: DOT/FAA/AM-00/32. [Accessed 2019 July 26]. Available from: https://apps.dtic.mil/docs/citations/ADA384889.
11. Fries E, Dettenborn L, Kirschbaum C. The cortisol awakening response (CAR): facts and future directions. Int J Psychophysiol. 2009; 72(1):67–73.
12. Grissom N, Bhatnagar S. Habituation to repeated stress: get used to it. Neurobiol Learn Mem. 2009; 92(2):215–224.
13. Gunnar M, Quevedo K. The neurobiology of stress and development. Annu Rev Psychol. 2007; 58(1):145–173.
14. Henckens MJAG, van Wingen GA, Joels M, Fernandez G. Time-dependent effects of cortisol on selective attention and emotional interference: a functional MRI study. Front Integr Neurosci. 2012; 6:66.
15. Kudielka BM, Broderick JE, Kirschbaum C. Compliance with saliva sampling protocols: electronic monitoring reveals invalid cortisol daytime profiles in noncompliant subjects. Psychosom Med. 2003; 65(2):313–319.
16. Larsson CA, Gullberg B, Rastam L, Lindlblad U. Salivary cortisol differs with age and sex and shows inverse associations with WHR in Swedish women: A cross-sectional study. BMC Endocr Disord. 2009; 9(1):16.
17. Lennartsson A-K, Sjörs A, Währborg P, Ljung T, Jonsdottir IH. Burnout and hypocortisolism–a matter of severity? A study on ACTH and cortisol responses to acute psychosocial stress. Front Psychiatry. 2015; 6:8.
18. Lim J, Dinges DF. Sleep deprivation and vigilant attention. Ann N Y Acad Sci. 2008; 1129(1):305–322.

19. Lü W, Wang Z, You X. Physiological responses to repeated stress in individuals with high and low trait resilience. Biol Psychol. 2016; 120: 46–52.
20. Mansikka H, Virtanen K, Harris D. Comparison of NASA-TLX scale, modified Cooper–Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks. Ergonomics. 2019; 62(2):246–254.
21. Miller R, Stalder T, Jarczok M, Almeida DM, Badrick E, et al. The CIRCORT database: reference ranges and seasonal changes in diurnal salivary cortisol derived from a meta-dataset comprised of 15 field studies. Psychoneuroendocrinology. 2016; 73:16–23. Erratum in: Psychoneuroendocrinology. 2017; 76:226–227.
22. Nater UM, Rohleder N, Schlotz W, Ehlert U, Kirschbaum C. Determinants of the diurnal course of salivary alpha-amylase. Psychoneuroendocrinology. 2007; 32(4):392–401.
23. Nieuwenhuis S, Yeung N, van den Wildenberg W, Ridderinkhof KR. Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. Cogn Affect Behav Neurosci. 2003; 3(1):17–26.
24. Noyes JM, Bruneau DPJ. A self-analysis of the NASA-TLX workload measure. Ergonomics. 2007; 50(4):514–519.
25. O'Hare D, Wiggins M, Batt R, Morrison D. Cognitive failure analysis for aircraft accident investigation. Ergonomics. 1994; 37(11):1855–1869.
26. Pruessner JC, Wolf OT, Hellhammer DH, Buske-Kirschbaum A, von Auer K, et al. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. Life Sci. 1997; 61(26):2539–2549.
27. Rohleder N, Joksimovic L, Wolf JM, Kirschbaum C. Hypocortisolism and increased glucocorticoid sensitivity of pro-inflammatory cytokine production in Bosnian war refugees with posttraumatic stress disorder. Biol Psychiatry. 2004; 55(7):745–751.
28. Salimetrics.Expanded range high sensitivity salivary cortisol enzyme immunoassay kit. State College (PA): Salimetrics; 2014.
29. Schoofs D, Preuß D, Wolf OT. Psychosocial stress induces working memory impairments in an n-back paradigm. Psychoneuroendocrinology. 2008; 33(5):643–653.
30. Shappell SA, Wiegmann DA. A human error approach to accident investigation: The taxonomy of unsafe operations. Int J Aviat Psychol. 1997; 7(4):269–291.
31. Shirtcliff EA, Peres JC, Dismukes AR, Lee Y, Phan JM. Hormones: commentary. Riding the physiological roller coaster: adaptive significance of cortisol stress reactivity to social contexts. J Pers Disord. 2014; 28(1):40–51.
32. Stalder T, Kirschbaum C, Kudielka BM, Adam EK, Pruessner JC, et al. Assessment of the cortisol awakening response: expert consensus guidelines. Psychoneuroendocrinology. 2016; 63:414–432.
33. Vedhara K, Hyde J, Gilchrist ID, Tytherleigh M, Plummer S. Acute stress, memory, attention and cortisol. Psychoneuroendocrinology. 2000; 25(6):535–549.
34. Wessa M, Rohleder N, Kirschbaum C, Flor H. Altered cortisol awakening response in posttraumatic stress disorder. Psychoneuroendocrinology. 2006; 31(2):209–215.
35. Zheng Y, Lu Y, Fu S. Flight crew workload evaluation based on the workload function distribution method. Aerosp Med Hum Perform. 2017; 88(5):481–486.