Video Article

Collecting Sleep, Circadian, Fatigue, and Performance Data in Complex Operational Environments

Lucia Arsintescu[1], Kenji H. Kato[2], Cassie J. Hilditch[1], Kevin B. Gregory[1], Erin Flynn-Evans[3]

[1]San Jose State University Research Foundation

[2]ASRS Research and Technology Solutions

[3]Fatigue Countermeasures Laboratory, NASA Ames Research Center

Correspondence to: Erin Flynn-Evans at erin.e.flynn-evans@nasa.gov

Abstract

Sleep loss and circadian misalignment contribute to a meaningful proportion of operational accidents and incidents. Countermeasures and work scheduling designs aimed at mitigating fatigue are typically evaluated in controlled laboratory environments, but the effectiveness of translating such strategies to operational environments can be challenging to assess. This manuscript summarizes an approach for collecting sleep, circadian, fatigue, and performance data in a complex operational environment. We studied 44 airline pilots over 34 days while they flew a fixed schedule, which included a baseline data collection with 5 days of mid-morning flights, four early flights, four high-workload mid-day flights, and four late flights that landed after midnight. Each work block was separated by 3–4 days of rest. To assess sleep, participants wore a wrist-worn research-validated activity monitor continuously and completed daily sleep diaries. To assess the circadian phase, pilots were asked to collect all urine produced in four or eight hourly bins during the 24 h after each duty block for the assessment of 6-sulfatoxymelatonin (aMT6s), which is a biomarker of the circadian rhythm. To assess subjective fatigue and objective performance, participants were provided with a touch-screen device used to complete the Samn-Perelli Fatigue Scale and Psychomotor Vigilance Task (PVT) during and after each flight, and at waketime, mid-day, and bedtime. Using these methods, it was found that sleep duration was reduced during early starts and late finishes relative to baseline. Circadian phase shifted according to duty schedule, but there was a wide range in the aMT6s peak between individuals on each schedule. PVT performance was worse on the early, high-workload, and late schedules relative to baseline. Overall, the combination of these methods was practical and effective for assessing the influence of sleep loss and circadian phase on fatigue and performance in a complex operational environment.

Video Link

The video component of this article can be found at https://www.jove.com/video/59851/

Introduction

Fatigue, arising from sleep loss and circadian misalignment, is a serious threat to safety in occupations that require 24 h operations, irregular schedules, and extended work hours[1,2]. Laboratory research has been instrumental in characterizing how changes in sleep duration and timing influence subsequent alertness and performance[3,4,5]. These studies form the basis for fatigue risk management recommendations and work scheduling practices in operational environments[6].

In this manuscript, a field study of aviation operations is used to demonstrate an approach for collecting sleep, circadian, fatigue, and performance data in complex operational settings[7]. We studied 44 airline pilots over 34 days while they flew a schedule that included periods of mid-morning flights, early flights, high-workload mid-day flights, and late flights that landed after midnight. Each work block was separated by 3–4 days of rest. Pilots collected objective and subjective data over the entirety of the study period including both flight duty and rest days.

Given the differences between laboratory and real-world environments, implementation of strategies and countermeasures developed in the laboratory do not always translate into operations as expected. Individual differences, a broad range of operational work schedules, irregular and unpredictable operations, organizational practices and culture, and labor agreements are some of the factors that can complicate the application of science into practical operational use. As a result, it is important to evaluate the impact of such interventions using consistent and reliable methods to assess sleep, circadian rhythms, fatigue or alertness, and performance. The level of monitoring and data collecting must be kept proportional to the anticipated levels of fatigue and associated risks to safety within an operation[8]. Furthermore, in any safety-sensitive setting, maintaining safe operations is paramount to the investigatory protocol.

The gold standard method for assessing sleep duration and quality is through polysomnography (PSG), which involves measuring brain activity, heart rate, eye movement, and muscle activity through a collection of electrodes and sensors placed on the scalp, face, and chest. Although robust, PSG is not practical for collecting sleep information in most operational environments. Many wearable devices have been developed to estimate sleep timing, duration, and quality, but few have been validated[9,10]. The combination of wrist-worn actigraphy and daily sleep

diaries have been widely used to estimate sleep in field studies across a range of occupations[11,12,13,14] and have been validated against PSG, showing concordance for sleep duration[15]. In addition, using actigraphy and sleep diaries for field studies places a low burden of effort on study participants, because most actigraphy devices are worn on the non-dominant wrist and only removed for showering or swimming, much like a wristwatch. Likewise, a well-designed sleep diary, presented on a phone or touch-screen device, can typically be completed by participants in less than two min.

The sleep-wake cycle is coordinated by the circadian pacemaker located in the suprachiasmatic nuclei of the hypothalamus[16]. This pacemaker also synchronizes many other aspects of biological function such as body temperature and hormonal rhythms (e.g., melatonin and cortisol). The endogenous circadian rhythm is near to, but not exactly, 24 h; therefore, it must be reset each day to allow for stable synchronization (i.e., entrainment) to the 24 h day. The primary resetting agent of the circadian pacemaker is light. In operational environments that require non-standard schedules and 24 h operations, circadian misalignment can occur, in which the circadian drive to sleep coincides with scheduled work[11]. It is possible to determine when the circadian pacemaker is promoting sleep and wake by measuring the peak timing (i.e., circadian phase) of the rhythms of biological signals that are controlled by the circadian rhythm.

It is important to measure the circadian phase following implementation of countermeasures in order to better understand whether such techniques are successful in aligning the circadian pacemaker with the imposed work schedule. Many of the outputs of the circadian system used to determine phase in laboratory settings are prone to masking, making them unsuitable for use in a field environment. For example, circadian changes in body temperature are difficult to detect in free-living individuals who may engage in activities such as exercise that alters their body temperature. Melatonin is acutely suppressed by light exposure, making collection of melatonin in blood or saliva impossible in situations where light cannot be controlled. However, 6-sulfatoxymelatonin (aMT6s), the major metabolite of melatonin, is excreted in urine and is less affected by the masking effects of light, making it an ideal candidate for measuring the circadian phase in operational environments[17,18].

In addition to measuring changes in physiology, it is also important to measure the impact of work schedule changes on subjective fatigue or alertness. While there are several scales available to measure different aspects of alertness and fatigue, the most commonly used in aviation are the 7-point Samn-Perelli Fatigue Scale (SP)[19] and 9-point Karolinska Sleepiness Scale (KSS)[20]. The SP is also commonly used in field studies of shift workers across a wide range of occupations[21,22,23,24]. The KSS has been validated against objective measures of sleepiness such as electroencephalography (EEG) and slow rolling eye movements[20,25], as well as performance[25]. This scale is commonly used in studies both in the laboratory and the field[24,26]. There may be other subjective scales that are appropriate for different shiftwork or occupational environments. It is important to choose a scale that has been validated and ideally has meaningful thresholds for levels of "acceptable" alertness. For example, KSS scores over 7 are associated with high levels of physiological signs of sleepiness and impaired driving performance[25,27], while Samn-Perelli ratings relate directly to flying duties[28]. For the study described in this manuscript, the Samn-Perelli was used, because it was originally developed as a subjective fatigue measure in a study population consisting of pilots.[28]

Although measuring sleep and circadian phase is an important component in evaluating an intervention, a primary outcome of interest in field studies is typically objective performance. There are a variety of tests that have been developed to evaluate cognitive performance, but the most sensitive and reliable test for measuring the effects of sleep loss and circadian misalignment is the Psychomotor Vigilance Task (PVT). The original PVT (PVT-192) is a simple reaction time test, where an individual is presented with a stimulus and is instructed to respond to the stimulus by pressing a button as quickly as possible[29]. The PVT has been validated under conditions of acute and chronic sleep loss and circadian misalignment[4,5,30]. The duration of the task can be varied based on the design of the study[31,32]; although, the traditional 10 min duration is preferred in laboratory studies[33,34], while a 5 min duration PVT is typically more feasible in field studies where operational demands can interfere with the administration of the test[35].

In addition, the PVT shows little to no learning effects and is simple to use, making it a practical test for deploying in field environments where study participants may not be observed during testing[36]. The ubiquity of touch-screen devices allows for easy deployment of the PVT, but researchers should be cautious when implementing the PVT, because there are numerous aspects of touch-screen devices that can introduce error into the collection of PVT data[37,38]. For example, different hardware and software combinations have different system latencies, and other applications running in the background can introduce unknown error into the recorded reaction times. As a result, it is important to collect PVT data using a validated PVT, with consistent hardware and software, with WiFi, and with all other applications turned off. In addition, given that it is not practical to observe study participants during tests in operational environments, it is critical that participants are trained to complete each PVT with the device in the same orientation, using the same finger[38,39].

Each of these elements of data collection is important and these tools have been used in other operational studies in the past[40,41,42,43]. However, in addition to the challenges described above, it can be difficult to achieve compliance with study procedures when participants are required to independently complete tasks, especially when such tasks include a time-sensitive component. A final element that is important in data collection in operational environments is the organization of information in a way that makes it easy for individuals to complete tasks on time. The NASA PVT+ application for touchscreen devices can be customized to present tasks to participants in sequence, guiding them through study procedures. For example, in the study presented here, airline pilots are provided with touchscreen devices pre-loaded with an application that is used to complete sleep diaries every morning and evening. The devices are also used to complete PVT tests and fatigue ratings, among other tasks, in the morning, at the top-of-descent (TOD) of each flight, post-flight, and in the evening before bed. This presentation of information allowed pilots to complete study procedures with minimal inconvenience to their work-related tasks.

It can be very difficult to collect data among pilots, as the nature of the job requires them to travel long distances and work in confined spaces (i.e., cockpits) with many distractions and often unpredictable workloads. Despite these challenges, it is critical to collect data in this population, because pilot fatigue is a threat to safe aviation operations[40,44,45]. The high intensity of airline operations is conducive to degradation of crew performance and increases the risk of fatigue-related incidents[46,47,48,49,50]. Using the combination of methods described above, we measured sleep, circadian rhythms, fatigue and performance among 44 short-haul airline pilots over 34 days. During the study, pilots flew a fixed schedule that included a baseline data collection with 5 days of mid-morning flights, four early flights, four high-workload mid-day flights, and four late flights landing after midnight. Each work block was separated by 3–4 days of rest. These findings demonstrate how comprehensive data collection, including measures of sleep, circadian rhythms, fatigue, and performance, can be used in operational environments.

In this case, the purpose of the study was to evaluate sleep, circadian rhythms, fatigue, and performance by duty start time as follows. 1) Baseline: during the first duty block, all pilots worked 5 days that each included two flights of about 2 h each, starting in the mid-morning, to allow for an adequate nighttime sleep episode. This block was followed by 4 rest days. 2) Early starts: during the early duty block, all pilots worked 5 days that each included two flights of about 2 h, each starting between approximately 5:00 AM and 8:00 AM. This block was followed by 3 rest days. 3) High-workload shifts mid-day: during the mid-day duty block, all pilots worked 5 days, which each included 2–4 flights of ~2–6 h each, starting at approximately mid-day. This block was followed by 3 rest days. 4) Late finishes: during the late duty block, all pilots worked 5 days, which included two flights of about 3 h each, starting in the late afternoon around 4:00 PM and ending around midnight. This block was followed by 3 rest days.

Protocol

This study was approved by the Institutional Review Board (IRB) of NASA Ames Research Center, and all subjects provided written informed consent. All study procedures conformed to those in the protocol approved by the NASA IRB (protocol number HRI-319).

1. Participant Selection and Preparation for the Experiment

1. **Determining the study schedule**
   1. Include a baseline data collection period in order to assess outcomes in the absence of intervention.

2. **Identifying when to collect data during the experiment**
   1. When assessing fatigue, it is critical that outcome measures include more than one data point due to time-of-day changes in fatigue and performance.
   2. Collect performance data in conjunction with operationally critical tasks when possible. In the case of airline pilots, it is useful to collect a PVT and fatigue rating at the TOD of a flight, which is in the final phase of cruise, just before the critical task of landing occurs.
      NOTE: More inflight PVT data may be required to assess outcomes for longer flights or in other operational environments. A researcher may be interested, for example, in the change in fatigue across a rest break, which requires measures to be taken pre- and post-break.

3. **Recruitment procedures may vary depending on study objectives. Ensure that participants are employed by a single organization and that they represent the typical population of the chosen organization; therefore, no additional screening criteria needs to be applied.**
   NOTE: In laboratory environments, participants are typically screened out of participation if they have chronic health conditions or score out of normal range on sleep disorder questionnaires. In operational environments, participants may have chronic conditions and undiagnosed sleep disorders, but these individuals should generally be included in studies in order to gauge the effectiveness of interventions across a broad spectrum of workers. Individuals should be invited to participate in the study through e-mail or in-person presentation to the population of interest.
   1. Have participants engage with study personnel directly and assure them that employment decisions will not be made based on individual data.
   2. Take any additional precautions that may be necessary to protect the confidentiality of participants from their employer, such as obtaining a certificate of confidentiality from the National Institute of Health or a letter from company management assuring volunteers that their employment will not be affected by their participation in the study.
   3. Have participants undergo informed consent confirming that participation in research is strictly voluntary. Ensure that interested volunteers can follow the study work schedule and discourage them from trading work shifts during the study.

4. **Invite participants who volunteered in the study to attend a 30–60 min training/briefing session.**
   1. Provide participants with a data collection kit that includes the activity monitor, the touchscreen device with the appropriate application installed, and urine collection supplies. Have participants complete background questionnaires [e.g., Pittsburgh Sleep Quality Index (PSQI)[51], Fatigue Severity Scale (FSS)[52], Epworth Sleepiness Scale (ESS)[53], Checklist of Individual Strength (CIS)[54], Morningness/Eveningness Questionnaire (MEQ)[55]] to estimate the prevalence of baseline sleep issues in the population. These questionnaires may also be used as covariates in data analysis.
   2. Review all aspects of the study with the participants and train them on procedures for completing study questionnaires and tests. Have participants complete one sleep diary and all daily questionnaires in front of the study staff to ensure that the participant understands how to complete the questions and use the application. Train the participants on the appropriate procedures for completing the PVT as described in section 6 below.
   3. Provide participants with a research validated activity monitor. Instruct participants to wear the activity monitor at all times, only removing it when it may be immersed in water. Ask participants to note the time of activity monitor removals in the application.
   4. Provide participants with urine collection materials and orient them to urine collection procedures. If a participant is unwilling or unable to collect urine samples, then this element of the study can be included as an optional sub-study if approved by the ethical review board.

2. Experimental Design

1. **Work schedule: Ensure that all individuals follow the same schedule (or balanced schedule in the case of a randomized experiment). In addition, include a baseline or placebo condition in the experiment to interpret findings in the absence of intervention or schedule manipulation.**
   NOTE: It will likely be necessary to work closely with the partner organization to implement the experimental schedule. Other considerations may be necessary depending on the population to be studied. In the case of our study, in which we evaluated short-haul airline pilots, we

designed a schedule that allowed them to return home each day to ensure they had a consistent sleep environment. The final schedule for evaluating outcomes of interest should be similar to the one diagrammed in **Figure 1**.

1. Data collection protocol on non-work days:
   1. Instruct participants to complete the sleep diary each day within 30 min of waking up and going to bed.
   2. Ask participants to wear their activity monitor continuously.
   3. Instruct participants to complete a Samn-Perelli fatigue scale (SP) and a PVT 3x a day: morning (1–2 h after waking), mid-day (8–9 h after waking), and evening (1–2 h before going to bed).

2. Data collection protocol on work days:
   1. Instruct participants to complete the sleep diary each day within 30 min of waking up and going to bed.

2. Data collection schedule: ensure that the data collection schedule includes collection of PVT data at operationally relevant times and at multiple timepoints per day (at a minimum, morning, mid-day, and evening) in order to assess changes in performance associated with sleep loss or circadian misalignment. In addition, verify that the data collection schedule does not overly burden the participant with unnecessary data collection.
   NOTE: The balance between collecting enough information to evaluate interventions and not over-burdening the participant is critical to maintaining study compliance and minimizing withdrawal from the study.

3. Actigraphy Collection Procedures

1. Select an activity monitor that has been validated against PSG in a laboratory environment (**Figure 2**).
   NOTE: Some activity monitors include additional features, such as temperature and heart rate, but as described previously, these features can be influenced by an individual's activities. Two useful features available on some devices are event markers and light sensors. Lighting information can be useful in interpreting circadian phase information and event markers can be used to mark time-in-bed and activity monitor removals, but these features are not required for the collection of sleep information.
2. Instruct participants to wear the activity monitor snugly and securely fastened on the wrist of the non-dominant hand during the entire experimental period. If the activity monitor is not securely fastened on the wrist, activity counts could be compromised. Demonstrate appropriate fitting during the pre-study briefing session.
3. Instruct participants to wear the activity monitor during exercise but remove it before swimming and showering. Activity monitor removals should be noted in the daily sleep diary to distinguish inactivity due to activity monitor removals from naps.
4. If the activity monitor includes an event marker feature, ask the participant to hit the marker whenever the activity monitor is removed. In addition, instruct the participant to press the event marker when going to bed and waking up for every sleep episode. This will enhance information obtained in the sleep diary and assist in analysis.

4. App-based Questionnaire, Sleep Diary, and PVT Collection

1. Select the application for daily data collection.
   NOTE: App-based questionnaires should be cross-checked against validated paper versions to ensure that the app-based versions faithfully reflect the original instruments. In particular, questionnaires that include visual analog scales should be checked to confirm that app developers did not transform outcomes to Likert-type scales. Similarly, all questions and responses should be checked to confirm that original language and response options are fully included and that the visualization of the questions and responses are not compromised by the screen size of device being used.
2. Questionnaires: have participants complete baseline questionnaires (outlined in section 1.4.1) and demographic information prior to engaging in any study procedures to aid in the interpretation of study outcomes.
3. **Sleep diaries: ask participants to complete the sleep diary before and after the individual's main sleep episode. It is desirable to provide participants with an application-based sleep diary rather than paper because 1) it is easy for the participant to complete and 2) application-based sleep diaries are time stamped, which minimizes the possibility that an individual will complete the sleep diary retroactively.**
   1. Waking sleep diary: upon waking from the main sleep episode, have participants indicate 1) wake up time, 2) number and duration of awakenings, and 3) sleep quality.
   2. Bedtime sleep diary: just prior to going to bed, instruct participants to document the duration and timing of any naps and anticipated bedtimes.
4. PVT: instruct participants to take the PVT at pre-determined times throughout the study using precise methodology as described below (in section 6).
5. Present study information clearly and concisely, so that participants are presented with tasks in a manner that allows them to quickly determine what they need to do.
   NOTE: We use the NASA PVT+ app for our studies. The NASA PVT+ is an application for touchscreen devices developed at NASA Ames Research Center. The application is customized for each experiment to include or exclude various questionnaires as needed. The app presents each activity that a participant needs to complete in sequence (see **Figure 3**). For the present study, the main page of the app displays three main links: "Study enrollment", "Rest day" and "Duty day". The "Study enrollment" link includes the following questionnaires to be completed during the training day: Demographics, MEQ, CIS, and ESS. After completion, the link is not visible to the participants. The "Rest day" link includes the questionnaires to be completed during days off: morning sleep diary; SP and PVT for morning, afternoon, and evening; and evening sleep diary, presented in that order. The "Duty day" link displays three main links: Morning (sleep diary); Work Tasks; and Evening (sleep diary). The Work Tasks link is comprised of three links: Pre-flight (SP, PVT, commute time); Record at TOD, which displays the number of flights selected by the participant to complete the tests for that particular flight (SP, PVT); and post-flight (SP, PVT).

5. Urine Collection Procedure

1. **Use urine collection to measure aMT6s production to estimate circadian phase (adapted for urine collection in pilots from procedures developed by Lockley[18] and Hull[56]).**
   1. Instruct participants to collect urine samples following any intervention that is expected to shift circadian phase. Provide participants with the urine kit, urine log, and instructions on the training day. The urine kit (**Figure 4**) includes a urinal hat or urinal collection container, several pipettes, five labeled urine collection tubes per 24 h collection, two extra tubes and white sticker labels, clean biohazard zip-lock bags, shipping materials, an ice pack, a urine collection log (**Figure 5**), and a copy of instructions for reference during each collection block (outlined in section 5.2).
   2. General overview instructions: inform participants that they need to collect all urine produced over a 24 h period. The first collection block starts from the time the participant wakes up on the first urine collection day and continues in 4 h blocks during the day and an 8 h block overnight. In total, each 24 h collection includes five samples.
   3. Provide participants with the instructions below and review with them every step in the procedure during pre-study training. Ensure that participants are provided with a pre-paid, addressed shipping label.

2. Use the following urine collection instructions for participants:
   When you wake up on the first day go to the bathroom in the toilet as usual. You will begin collecting your urine after you go to the bathroom for the first time.
   At the start of each collection block record the date, the start time and the approximate end time of the collection block on the collection log (e.g., 07:00–11:00 AM). You might urinate more than once during a 4 h collection block (or during the 8 h overnight collection block). In this case, you will urinate in the same urine collection container.. For example, during the 07:00–11:00 AM collection window you may urinate at 8:00 AM and 10:55 AM. Both of these urine voids will mix together in the same collection container. At the end of the collection block you will take a sample. Every time you urinate, record the exact clock time you urinate on the collection log. Every time you urinate, all of your urine should go into the collection container.
   Empty your bladder just prior to taking a sample. For example, if the collection block ends at 11:00 AM, try to urinate in the collection container just before 11:00 AM, then take the sample. At the end of the collection block, record the total volume of the urine collected using the markings on the container.
   Take a new pipette and transfer a small amount of the urine from the collection container into a small tube. Fill the tube with enough urine so that it is at least half-full. Cap the tube. Do not fill the tube completely, because urine expands when it is frozen and may break the tube if overfilled. After you fill the tube, then you can throw the pipette away. Each sample tube is labeled with a number and arranged in numeric order. Use the sample tubes in numerical order (i.e., use 1 first, then 2, etc.).
   Place the small tube into a zip-lock biohazard bag. Place the large zip-lock bag in the shipping box with a cold pack. After taking the sample, record the tube number and clock time that you took the sample on the collection log, then discard the remaining urine in the toilet. Rinse the collection container with water (do not use anything other than water for rinsing the collection container). Repeat the process for the next three daytime collection blocks (e.g., 11:00 AM–3:00 PM, 3:00 PM–7:00 PM, and 7:00 PM–11:00 PM).
   For the overnight collection block, collect your fourth daytime sample just before you go to bed. The collection container should be empty when you go to bed. The overnight collection block will include all of the urine you urinate during the night, as well as the urine from the first time you urinate in the morning. When you wake up in the morning, you will urinate into the collection container. Any urine from the night will be mixed in with this first urine you produce in the morning. If you do not wake during the night, then the overnight sample may only include your first morning urination. After you have urinated for the first time on Day 2, the overnight collection block is complete.
   Please collect all of the urine you produce in a 24-hour period. If you accidentally urinate in the toilet, then please note this on your urine collection log and let us know that you missed a collection."

6. PVT Administration Methods

NOTE: As described in the introduction, NASA-PVT is a 5 min sustained-attention, reaction time test that measures the speed at which individuals respond to a visual stimulus. The duration of the test can be changed based on study design. There are numerous PVT designs that have been developed, including those that illuminate a target[35,57] or checkboard pattern[39]. The NASA-PVT was designed to mimic the laboratory PVT-192 device in which the target is in the form of a milliseconds counter.

1. Read the following instructions to each participant to ensure that each receives the same training: "Please hold the device in the landscape position each time and hover each of your thumbs over the device within a few millimeters of the screen the entire time you are taking this test. During the test, tap the screen using the thumb of your dominant hand (that is, the hand you typically write with) as soon as you see the red numbers scrolling in the box. You must use your thumb from your dominant hand to respond to the stimuli in all tests. The numbers in the display show how fast you responded each time. The smaller the number, the better you did. Try to do your best and get the lowest number you possibly can each time. If you tap on the screen too early (before the numbers appear) you will see an error message ('FS') indicating a false start. If you tap using your non-dominant thumb, then you will see the message 'ERR', indicating an error. Avoid 'FS' and 'ERR.' If you forget to lift your thumb, the text screen will remind you after a short time." **Figure 6** displays the screen of the Demo NASA-PVT showing the correct position of the touchscreen device while taking the PVT and location of the thumb.
2. Instruct participants to keep the touchscreen device in Airplane mode, with WiFi off at all times.
   NOTE: This is especially important for the PVT task where the accuracy of the internal stopwatch is influenced by connectivity functions, thus influencing the reaction time[38].
3. Instruct participants to initiate the PVT at a time that is free from distractions. If distractions occur, have participants note the number of distractions within the application following the test.
   NOTE: Given the demands of operational environments, it is possible that participants will not be able to complete a scheduled PVT test. In this case, participants should be instructed to take the PVT as soon as possible after the missed test. It is also important to inform participants that they should maintain at least 30 min between PVT test sessions.

Representative Results

Using the methods described, we were able to collect over 700 days of data and over 3,000 PVTs and fatigue ratings among 44 short-haul pilots[7]. The goal of this study was to characterize changes in sleep, circadian phase, fatigue ratings, and performance among short-haul pilots by work start-time and workload during daytime flights.

To account for the within-subjects study design, all condition effects were evaluated for sleep and performance outcomes using repeated measures analysis of variance with unstructured covariances, using participant as a repeated factor. To evaluate whether sleep and performance outcomes varied by day on a given schedule, linear mixed effects models were applied to the changes in sleep and performance by day. To account for individual differences in adaptation to a given schedule, the intercept and slope were allowed to vary by individual.

The first objective addressed with these methods was to examine the impact of duty start time on sleep. Sleep duration, bedtime, wake time, and sleep quality were calculated using the sleep diary and actigraphy. An example of the actogram derived from activity monitor is illustrated in **Figure 7**. It was demonstrated that sleep timing and duration varied significantly as a function of work start time using mixed-effects regression analysis. **Table 1** displays the bedtime, wake time, sleep duration and sleep quality by schedule type as reported by participants in the sleep diary. Participants went to bed on average at around 23:10 (SD = 1:41) on baseline block. The bedtime for early duty schedule block differed significantly from baseline ($p < 0.01$) with participants reporting earlier bedtimes. The bedtimes for mid-day and late duty schedules also differed significantly from baseline ($p < 0.01$), with participants reporting later bedtimes. Participants went to bed significantly later ($p < 0.01$) on rest days compared to baseline.

Figure 8 displays the mean-actigraphy derived sleep duration by day for each schedule type. Participants obtained significantly less sleep ($p < 0.01$) on early starts compared to baseline. The sleep duration on the other schedule types were not different from baseline. Sleep latency and sleep efficiency obtained from actigraphy were not significantly different from baseline for any of the schedule types. Wake after sleep onset (WASO) was significantly different for early starts compared to baseline ($p < 0.05$), with pilots reporting being more awake during the early starts. There were no differences between the baseline and other schedule types. There were no significant differences between rest days and the baseline.

The second objective addressed with these methods was to examine the impact of duty start time on circadian phase as measured by aMT6s. The peak timing (acrophase) of the aMT6s rhythm is a reliable marker of circadian phase[58]. **Figure 9** shows an example of the circadian rhythm of aMT6s over 24 h for one individual, while **Figure 10** shows the aMT6s acrophase for each individual who participated in the urine collection procedures by study block. Consistent with the findings on sleep, it was found that mean circadian phase was significantly shifted according to work start time. It is important to note the missing data collection information in **Figure 10**. Some individuals had difficulty with the urine collection procedures for some of the blocks or they forgot to log the timing of their sample collection. In these cases, it was not possible to generate reliable estimates of circadian phase from the aMT6s concentration and as a result some data are missing. In situations where collection of circadian phase information is important, it may be prudent to call participants prior to each urine collection to ensure the procedures are properly followed.

The third objective addressed with these methods was to examine the impact of duty start time on self-reported fatigue as measured by SP, and objective performance as measured by the PVT. Consistent with our findings with sleep, using mixed-effects regression analysis, we found that both fatigue (**Table 2**) and PVT reaction times (**Figure 11**) were worse during early starts, high workload mid-day shifts, and late finishes, relative to our baseline data collection ($p < 0.001$ SP; $p < 0.01$ PVT RT). Participants showed a significant increase in lapses for each schedule type compared to baseline ($p < 0.01$ early; $p < 0.05$ mid-day; $p < 0.01$ late). Performance on rest days was similar to that of the baseline. These results are also described in **Table 3**.
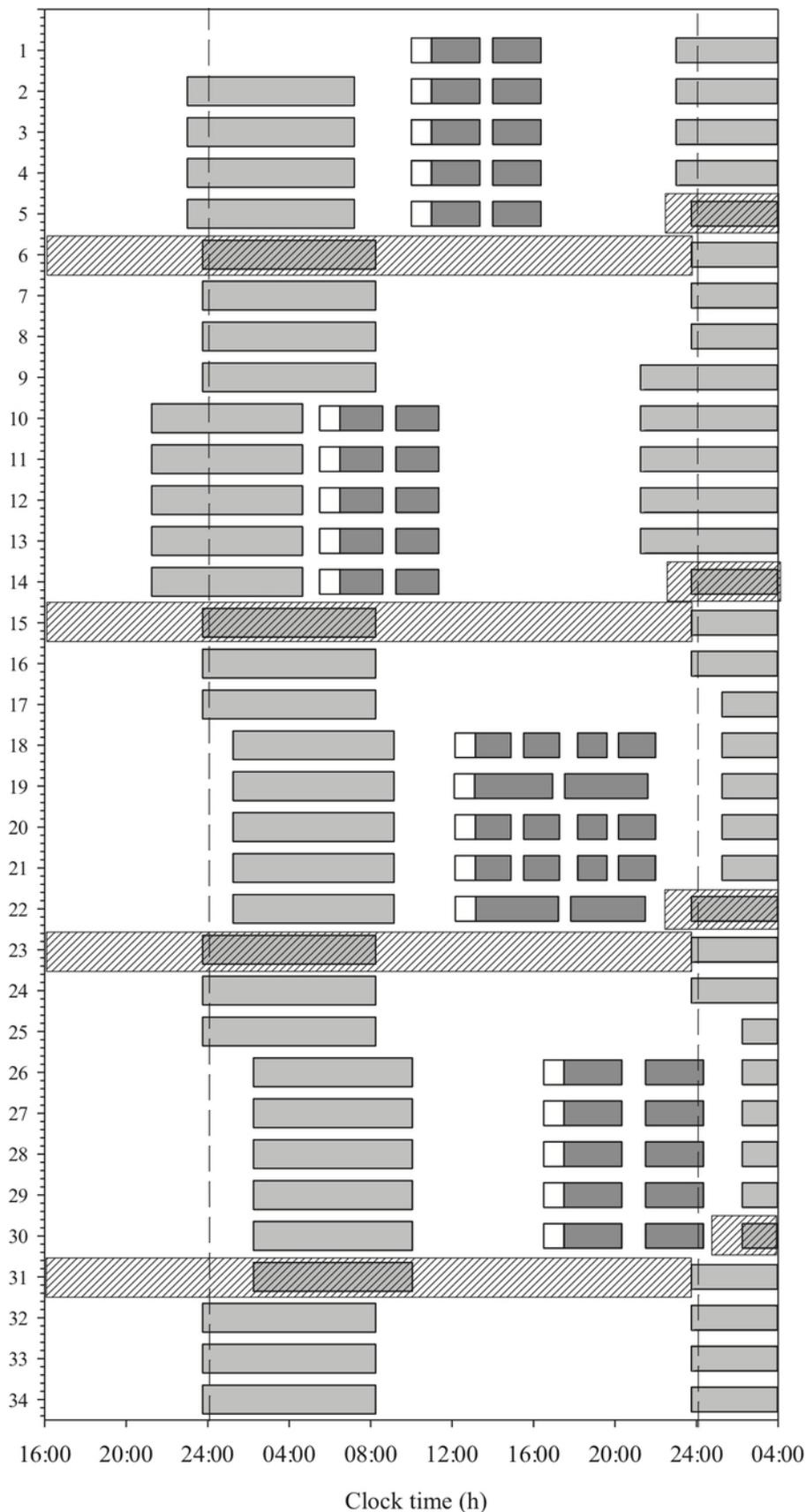
Figure 1: Study protocol by time of day for each day of the study. The dark gray bars represent the flight periods including the pre-flight report time (open bars), and the light gray bars represent the sleep periods. Days 1–5 represent the baseline duty block, days 10–14 represent

the early duty starts, days 18–22 represent the mid-day duty starts, and days 26–30 represent the late starts. The shaded bars represent the first rest day post duty block when urine is collected. This figure is reproduced from Flynn-Evans et al.[7]. Please click here to view a larger version of this figure.

Figure 2: The activity monitor/accelerometer device worn on the wrist of the non-dominant hand. Please click here to view a larger version of this figure.
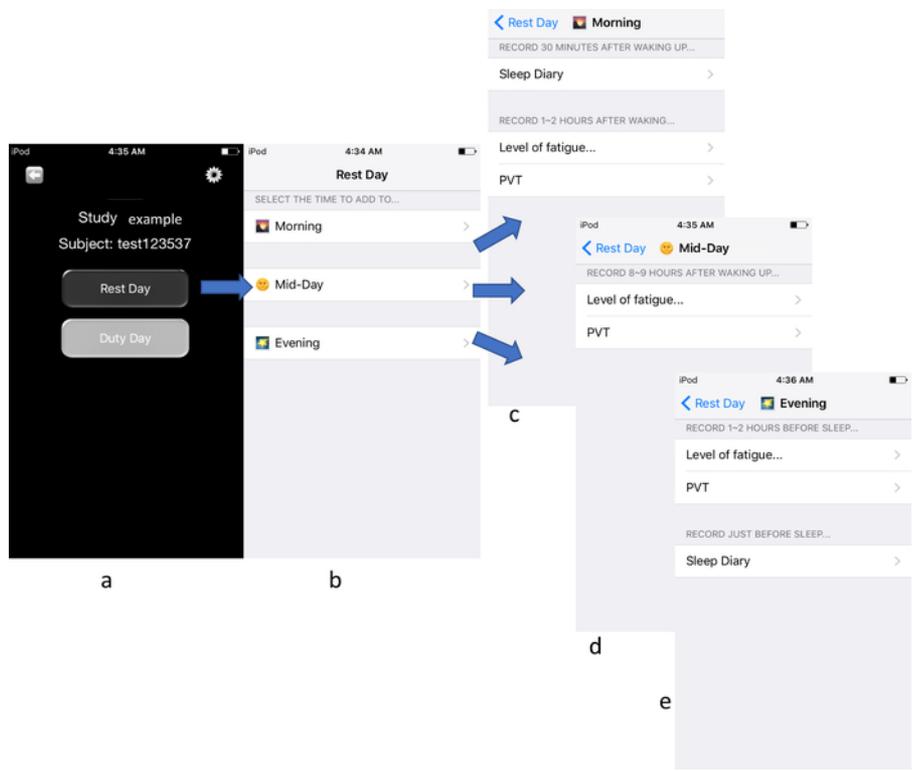
Figure 3: Example of tests taken during rest days using the touchscreen application. From left to right: (**A**) the main page of the app displays two links; (**B**) the rest day displays three links: morning, mid-day, evening; (**C**) the morning link displays the tests taken in the morning; (**D**) the mid-day link displays the tests taken in the afternoon, and (**E**) the evening link displays the test taken in the evening. Please click here to view a larger version of this figure.
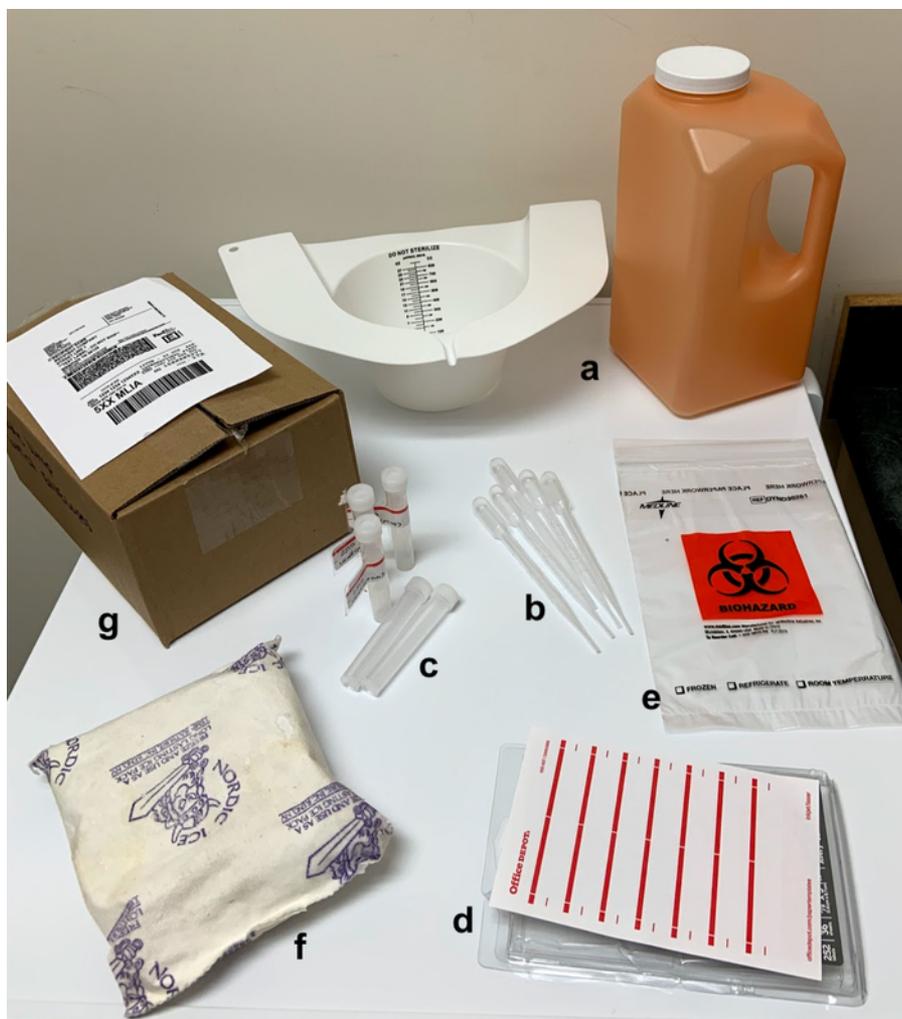
Figure 4: Urine kit. The kit contains (**A**) a urinal hat or urinal collection container, (**B**) pipettes, (**C**) urine collection tube, (**D**) white sticker labels, (**E**) a bio-hazard bag, (**F**) ice pack, and (**G**) shipping materials. Please click here to view a larger version of this figure.

Urine Collection Log Sheet

Date (dd/mm/yy)	Collection Block	Clock Time of Collection Block	Clock Time When Your Urinated	Total Urine Volume (ml)	Clock Time Sample Taken From Collection Container, Added to Sample Tube	Sample tube #
Please do not collect your first void when you wake on the first day. Urinate in the toilet and begin collection the next time you go to the bathroom.						
Date (dd/mm/yy)	#1	Start: End:				
Date (dd/mm/yy)	#2	Start: End:				
Date (dd/mm/yy)	#3	Start: End:				
Date (dd/mm/yy)	#4	Start: End:				
Date (dd/mm/yy)	#5	Start: End:				

EXAMPLE Urine Collection Log Sheet

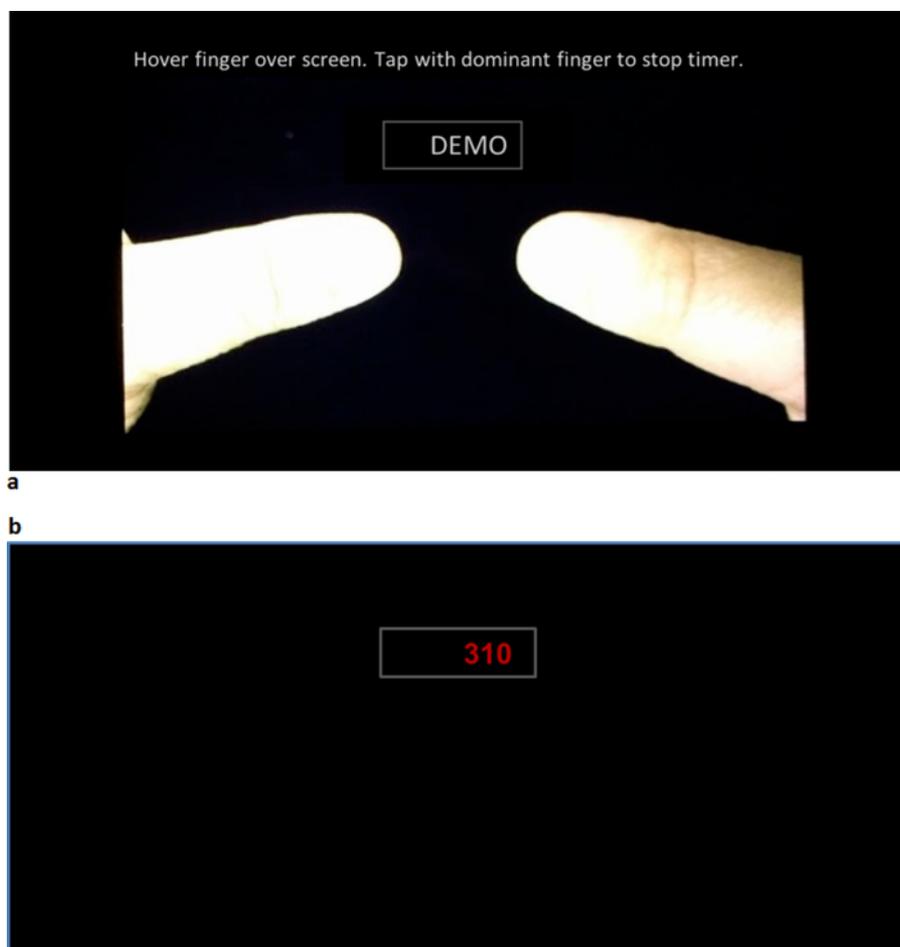| Date (dd/mm/yy)	Collection Block	Clock Time of Collection Block	Clock Time When Your Urinated	Total Urine Volume (ml)	Clock Time Sample Taken From Collection Container, Added to Sample Tube	Sample tube #
Please do not collect your first void when you wake on the first day. Urinate in the toilet and begin collection the next time you						
09/15/2019	#1	Start: 0700 End: 1100	0745 1050	550	1055	21501
09/15/2019	#2	Start: 1100 End: 1500	1230 1330 1455	600	1500	21502
09/15/2019	#3	Start: 1500 End: 1900	1855	400	1900	21503
09/15/2019	#4	Start: 1900 End: 2300	2115 2300	500	2305	21504
09/16/2019	#5	Start: 2300 End: 0700	0300 0700	850	705	21505

Figure 5: Example of the urine collection log.

Figure 6: Psychomotor Vigilance Task (PVT). (**A**) The touchscreen device is oriented in landscape position and the thumbs are displayed on the screen at the beginning of the test. (**B**) the reaction times are displayed on a rectangular box in the upper middle part of the screen. Please click here to view a larger version of this figure.
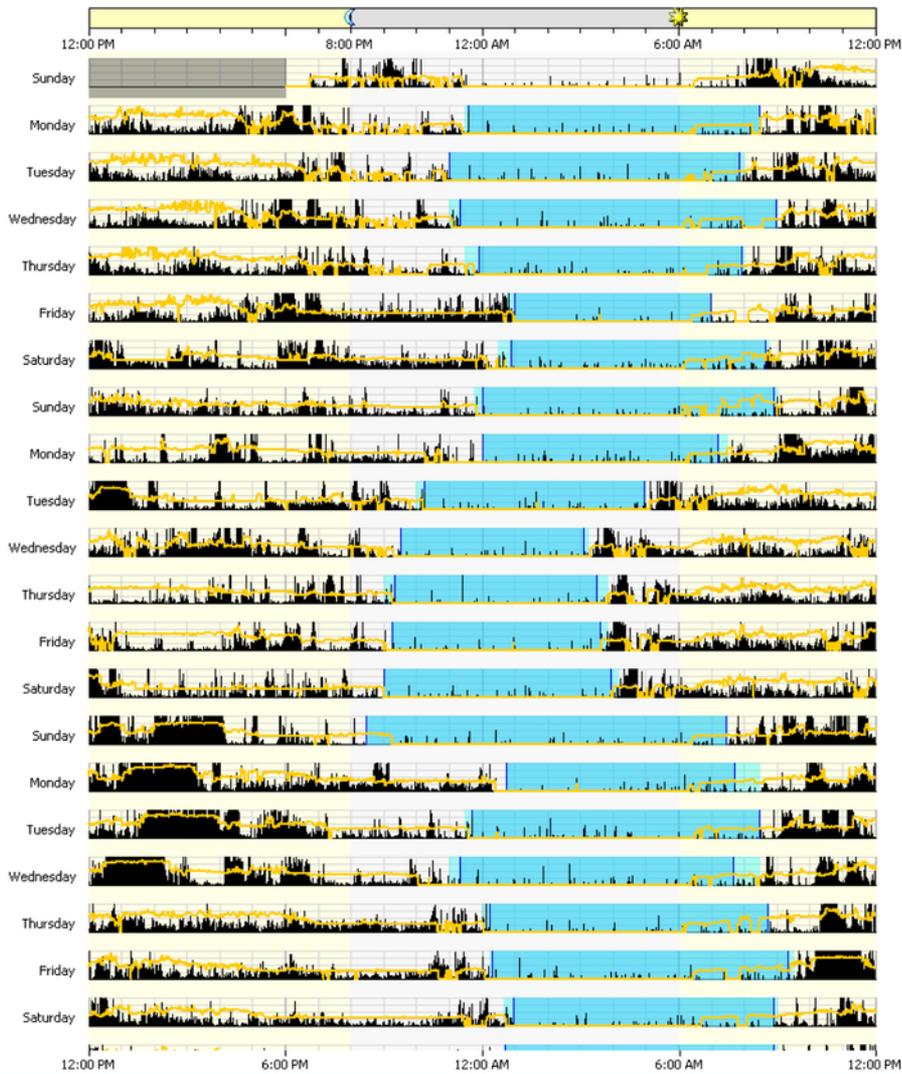
Figure 7: Actogram of sleep-wake cycles over 24 h for 14 days. The dark blue color represents the sleep periods; the light blue represents the rest periods. The black color represents movement. The yellow color represents the light. Please click here to view a larger version of this figure.
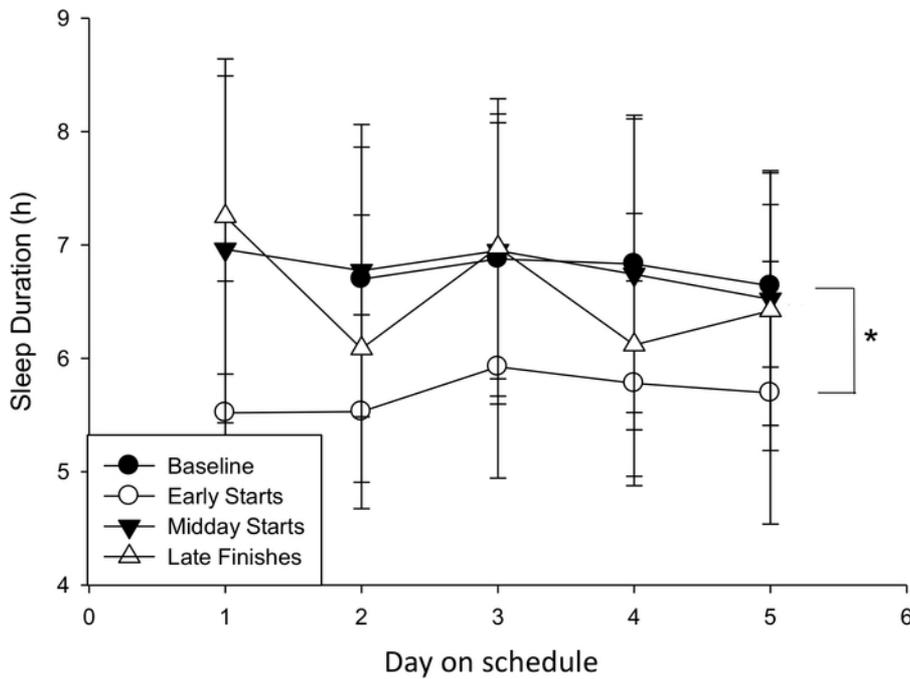
Figure 8: Mean actigraphy-derived sleep duration by day on each schedule type. Day 1 represents the night of sleep before the first work period of a given block. An asterisk designates a significant difference (*p < 0.05, **p < 0.01) in the means between the baseline condition and early starts block. Please click here to view a larger version of this figure.
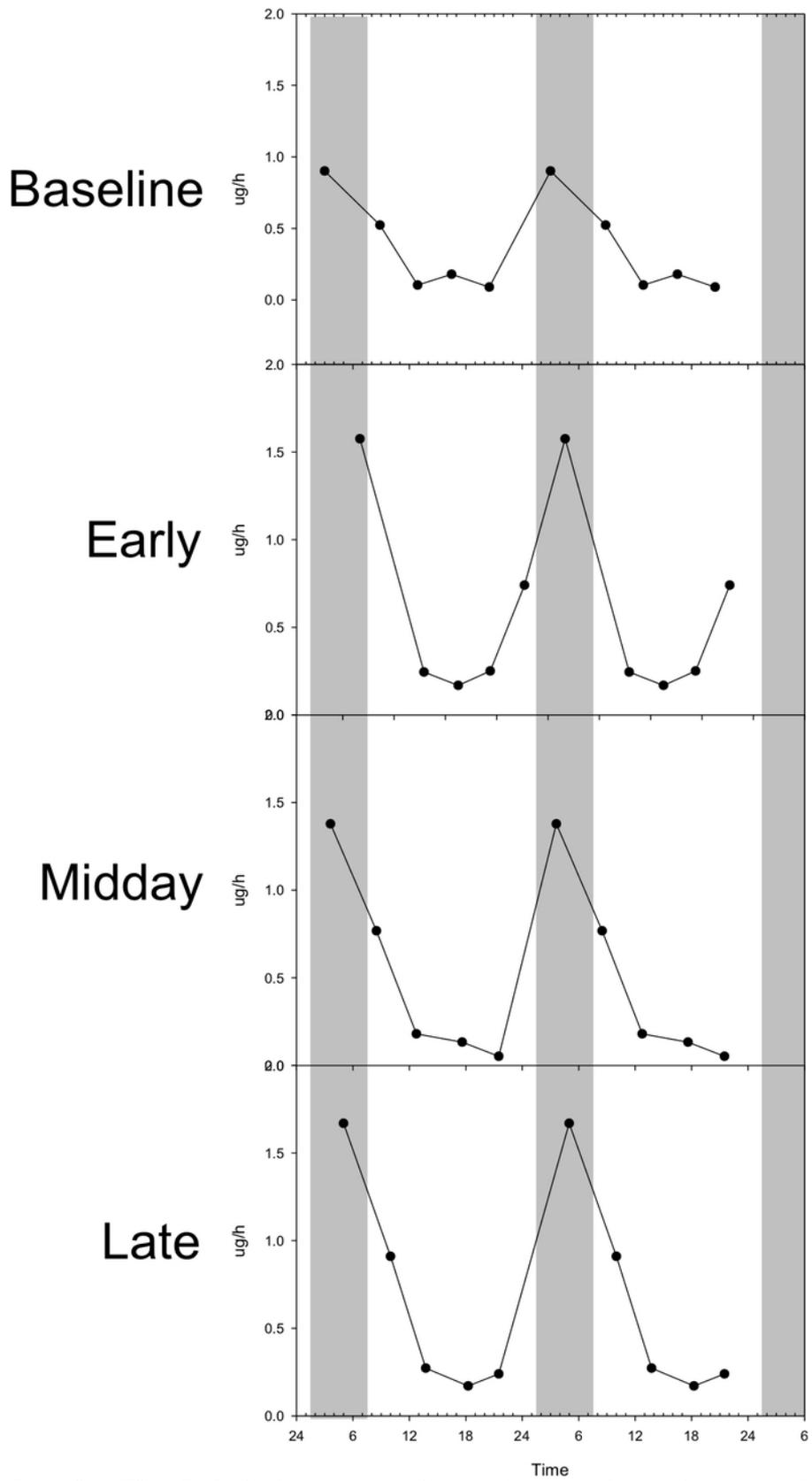
Figure 9: aMT6 profile for the five urine collection bins for each data collection episode for a single participant. Data are double-plotted.
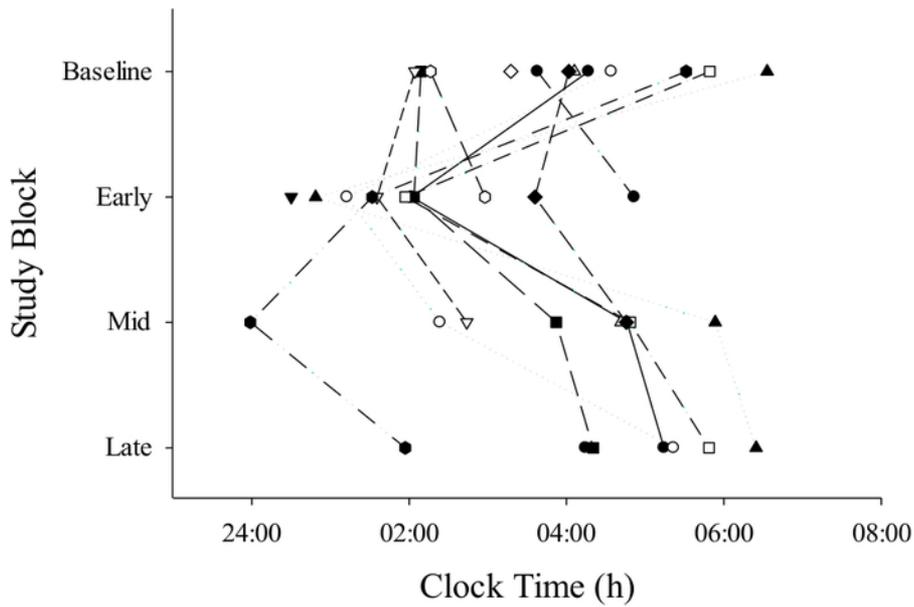Please click here to view a larger version of this figure.

Figure 10: 6-sulfatoxymelatonin (aMT6) acrophase (peak) by time (24 h clock) of circadian nadir and schedule type for each individual. Filled and open circles, triangles, squares represent individual participants. This figure is reproduced from Flynn-Evans et al.[7] Please click here to view a larger version of this figure.
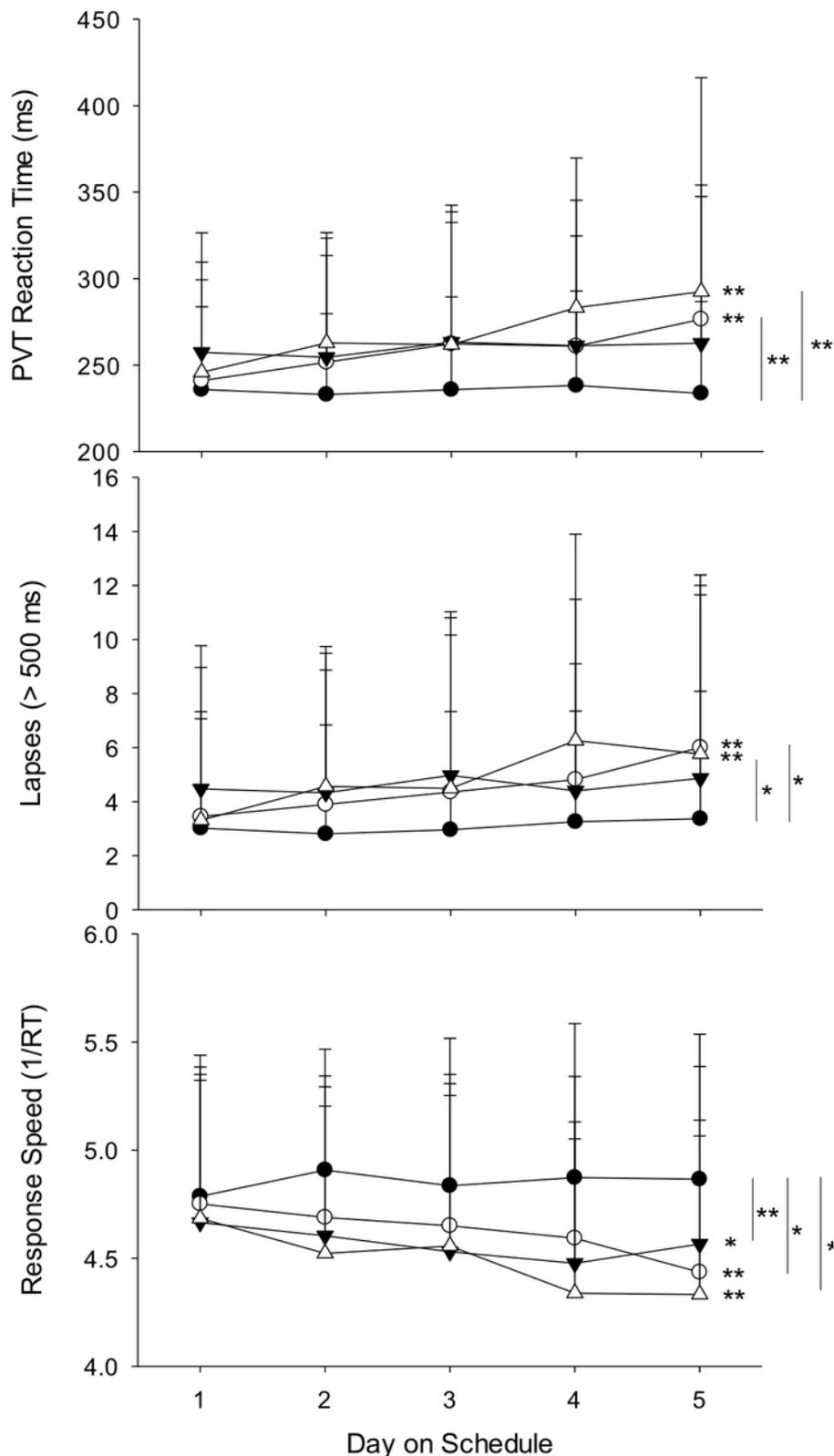
Figure 11: Psychomotor Vigilance Task (PVT) mean reaction time (RT), lapses (RT >500ms), and response speed (mean 1/RT) by day on each schedule type. Asterisks following each slope indicate changes in performance by day in that condition. Brackets indicate differences in the slope between baseline performance and the slope in performance in each of the other conditions (*p < 0.05, **p < 0.01). Baseline = filled circles, early = open circles, mid-day = filled triangles, late = open triangles. Please click here to view a larger version of this figure.

Work schedule	n	Bedtime (h, SD)	Wake time (h, SD)	Sleep duration (h, SD)	Sleep quality (SD)
Baseline (ref.)	39	23:10 (1:41)	7:20 (1:49)	8.2 (0.9)	2.4 (0.7)
Early	42	21:14 (1:01)**	4:29 (0:47)	7.4 (0.9)**	2.5 (0.6)
Midday	41	01:19 (0:43)**	9:11 (0:58)	7.9 (1.1)	2.3 (0.6)
Late	40	02:18 (1:07)**	9:57 (1:11)	7.8 (1.4)*	2.3 (0.7)
Rest days	42	23:47 (0:50)**	8:16 (0:58)	8.5 (0.9)*	2.4 (0.5)

Table 1: Sleep diary-derived sleep outcomes (bedtime, wake time, sleep duration and sleep quality) by schedule type. *p < 0.05, **p < 0.01; h = hour, SD = standard deviation. This table is reproduced from Flynn-Evans et al.[7]

Work schedule	Mean (SD)	p-value
Baseline	3.51 (0.80)	ref.
Early duty	4.03 (0.88)	< 0.001
Midday duty	3.85 (0.90)	< 0.001
Late duty	3.85 (0.89)	< 0.001

Table 2: Means and standard deviation for Samn-Perelli (SP) scores by duty block. A higher rating indicates greater fatigue.

Work schedule	n (participants)	Mean reaction time (ms, SD)	Response speed (s, SD)	Mean Lapses (> 500 ms)
Baseline (ref.)	38	236 (48)	4.84 (0.61)	3.1 (4.1)
Early	40	257 (70)**	4.63 (0.66)**	4.4 (5.4)**
Midday	39	261 (62)**	4.56 (0.66)**	4.7 (5.1)*
Late	38	266 (64)**	4.51 (0.63)**	4.7 (5.0)**
Rest days	40	249 (56)	4.69 (0.62)	4.0 (4.5)

Table 3: Psychomotor Vigilance Task (PVT) mean reaction time (RT), response speed (mean 1/RT), and lapses (RT > 500 ms) by schedule type. *p < 0.05, **p < 0.01; this table is reproduced from Flynn-Evans et al.[7]

Discussion

The methods described in this manuscript provide insight into sleep patterns, circadian phases, fatigue ratings, and performances of pilots during daytime flights including early starts, high workload mid-day flights, and late finishes. The combination of these methods demonstrated that these factors are all affected by modest changes in work start time and workload. By evaluating a systematic study schedule and integrating these measures into an easy-to-use touch-screen application, a large amount of data was collected in a challenging environment. Using this combination of methods allowed for a clearer interpretation of changes in alertness and performance during non-traditional daytime work shifts.

This design and implementation of methods measuring objective sleep, circadian, fatigue, and performance data were critical in allowing the determination of how work start-time influences pilots during daytime flights in the absence of jet lag. The protocol was designed to allow for systematic comparisons between conditions, while also minimizing the inconvenience to participants and maximizing data collection at operationally relevant timepoints. These are critical steps to collecting meaningful data in operational environments. The measures have been validated in both laboratory and field studies, which is important for interpreting results. Although the study was designed to enable participants to complete study procedures independently, the pre-study briefing session was crucial to ensure that volunteers understood the study procedures and importance of maintaining consistency when completing study tests and questions, particularly for the PVT.

The finding that sleep duration and timing changes according to work start time is consistent with prior studies in smaller samples of individuals that used PSG to evaluate sleep timing[59,60]. Although early starts and late finishes may be expected to encroach on sleep timing, the large sample of data collected in an operational environment provides insight into the unexpected ways that the participants lose sleep. For example, the wake maintenance zone, which represents the strongest drive to be awake, occurs just before a habitual bedtime. In laboratory studies, participants have been shown to have difficulty sleeping during the wake maintenance zone[61,62,63]. It was expected that participants may try to go to bed a few hours earlier than normal in order to prepare for early starts. It was also expected that as a result of trying to initiate sleep during the wake maintenance zone, participants may exhibit a long sleep latency during the sleep preceding early starts; however, this was not the case. These data highlight important differences between the laboratory and field, and they demonstrate the need for collecting sleep data in operational environments.

Although circadian phase information was obtained in a subset of individuals, the circadian phase changes observed in each schedule type mirrored the changes observed in sleep timing. The addition of the circadian phase to this protocol enhanced the ability to understand why fatigue ratings and performance changed by work start-time. Alertness and performance follow a circadian rhythm, with the lowest alertness and poorest performance typically coinciding with timing of the aMT6s acrophase. Although it was found that the circadian rhythms of most participants shifted in the expected direction relative to the imposed work schedule, it was also found that this shifting was variable between

individuals. This suggests that some individuals may have more difficulty adapting to early or late schedules, causing modest circadian misalignment. The combination of these methods enhanced interpretation of these conclusions.

The sleep data collected also allowed for a better understanding of why fatigue ratings and performance changed relative to the different work schedules. For example, it was found that during early starts and late finishes, Samn-Perelli ratings and PVT performance was poorer by day on each of those schedules. This makes sense, because pilots obtained less sleep during early starts and late finishes relative to baseline, which meant that they were accruing sleep debt with each day on those schedules. In contrast, PVT performance was also poorer by day during the high workload mid-day start schedules. During the mid-day schedule, the amount of sleep the pilots obtained was not different from sleep duration during the baseline data collection. As a result, this finding suggests that the poorer performance observed during the mid-day work schedules was not likely to be driven by acute sleep restriction. It would have been very difficult to interpret the fatigue ratings and performance data without the sleep data, making the combination of these methods important.

Although these methods were designed and implemented successfully, this approach can involve some challenges. For example, it is possible that participants may forget when or how to complete some procedures. It is helpful to communicate with volunteers regularly to confirm that they are completing tasks according to the protocol, especially during the first phase of urine collection. In addition, the risk of data loss increases as the length of the study increases, because individuals may lose or damage their study devices. If a study is scheduled for several weeks, as was the case for this study, then it may be desirable to download data at the study midpoint to reduce potential data loss and review compliance with the protocol. Insufficient or missing data may reduce interpretability of the results, so care must be taken to ensure that individuals are collecting data appropriately.

There are many possible applications for these methods in other operational settings. These methods may be used to characterize sleep, circadian phase, fatigue, and performance in occupations with unusual scheduling practices or environmental considerations, such as during spaceflight or military operations. In addition, there are many promising interventions and countermeasures evaluated in laboratory environments, such as the use of blue-enriched light to accelerate circadian phase shifting, strategic on-the-job napping, hypnotics to maximize sleep opportunities, and stimulants such as caffeine to improve alertness. Although such approaches may be shown to be effective under controlled laboratory conditions, the deployment of such tools and technology in operational environments must be evaluated to confirm their efficacy in reducing fatigue in the real world. The combination of actigraphy, sleep diaries, circadian phase information, fatigue ratings, and PVT collection, combined with an easy-to-use software application to facilitate administration of tasks, provides adequate data for evaluating the effectiveness of interventions. The combination of these methods has significant translational potential for other complex operational environments, where it may be difficult to deploy more invasive data collection efforts.

Disclosures

Acknowledgments

References

1. Fischer, D., Lombardi, D. A., Folkard, S., Willetts, J., Christiani, D. C. Updating the "Risk Index": A systematic review and meta-analysis of occupational injuries and work schedule characteristics. *Chronobiology International.* **34** (10), 1423-1438 (2017).
2. Williamson, A. *et al.* The link between fatigue and safety. *Accident Analysis and Prevention.* **43** (2), 498-515 (2011).
3. Dawson, D., Reid, K. Fatigue, alcohol and performance impairment. *Nature.* **388** (6639), 235 (1997).
4. Van Dongen, H. P., Maislin, G., Mullington, J. M., Dinges, D. F. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep.* **26** (2), 117-126 (2003).
5. Santhi, N., Horowitz, T. S., Duffy, J. F., Czeisler, C. A. Acute sleep deprivation and circadian misalignment associated with transition onto the first night of work impairs visual selective attention. *PLoS One.* **2** (11), e1233 (2007).
6. Gander, P. H. Evolving Regulatory Approaches for Managing Fatigue Risk in Transport Operations. *Reviews of Human Factors and Ergonomics.* 253-271 (2015).
7. Flynn-Evans, E. E. *et al.* Sleep and neurobehavioral performance vary by work start time during non-traditional day shifts. *Sleep Health.* **4** (5), 476-484 (2018).
8. Gander, P. *et al.* in *Principles and practice of sleep medicine* Vol. 6th eds M. *Kryger, T. Roth, & W. C. Dement Elsevier.* (2016).
9. Mantua, J., Gravel, N., Spencer, R. M. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. *Sensors (Basel).* **16** (5), 646 (2016).
10. de Zambotti, M., Claudatos, S., Inkelis, S., Colrain, I. M., Baker, F. C. Evaluation of a consumer fitness-tracking device to assess sleep in adults. *Chronobiology Internations.* **32** (7), 1024-1028 (2015).
11. Flynn-Evans, E. E., Barger, L. K., Kubey, A. A., Sullivan, J. P., Czeisler, C. A. Circadian misalignment affects sleep and medication use before and during spaceflight. *Npj Microgravity.* **2,** 15019 (2016).
12. Gander, P., Millar, M., Webster, C., Merry, A. Sleep loss and performance of anaesthesia trainees and specialists. *Chronobiology International.* **25** (6), 1077-1091 (2008).

August 2019 | 150 | e59851 | Page 19 of 21

13. Gander, P., van den Berg, M., Signal, L. Sleep and sleepiness of fishermen on rotating schedules. *Chronobiology International.* **25** (2-3), 389-398 (2008).

14. Roach, G. D., Darwent, D., Sletten, T. L., Dawson, D. Long-haul pilots use in-flight napping as a countermeasure to fatigue. *Applied Ergonomics.* **42** (2), 214-218 (2011).

15. Signal, T. L., Gale, J., Gander, P. H. Sleep measurement in flight crew: comparing actigraphic and subjective estimates to polysomnography. *Aviation Space and Environmental Medicine.* **76** (11), 1058-1063 (2005).

16. Czeisler, C. A., Gooley, J. J. Sleep and circadian rhythms in humans. *Cold Spring Harbor and Symposia on Quantitative Biology.* **72,** 579-597 (2007).

17. Bojkowski, C. J., Arendt, J., Shih, M. C., Markey, S. P. Melatonin secretion in humans assessed by measuring its metabolite, 6-sulfatoxymelatonin. *Clinical Chemistry.* **33** (8), 1343-1348 (1987).

18. Lockley, S.W., Skene, D.J., Arendt, J., Tabandeh, H., Bird, A.C., Defrance, R. Relationship between melatonin rhythms and visual loss in the blind. *The Journal of Clinical Endocrinology & Metabolism.* **82** (11), 3763-70 (1997).

19. Samn, S. W., Perelli, L. P. Estimating aircrew fatigue: a technique with application to airlift operations. *School of Aerospace Medicine Brooks AFB, TX.* (1982).

20. Åkerstedt, T., Gillberg, M. Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience.* **52** (1-2), 29-37 (1990).

21. Jay, S. M., Dawson, D., Ferguson, S. A., Lamond, N. Driver fatigue during extended rail operations. *Applied Ergonomics.* **39** (5), 623-629 (2008).

22. Dorrian, J., Baulk, S. D., Dawson, D. Work hours, workload, sleep and fatigue in Australian Rail Industry employees. *Applied Ergonomics.* **42** (2), 202-209 (2011).

23. Ferguson, S. A., Baker, A. A., Lamond, N., Kennaway, D. J., Dawson, D. Sleep in a live-in mining operation: the influence of start times and restricted non-work activities. *Applied Ergonomics.* **42** (1), 71-75 (2010).

24. Gander, P. H. *et al.* Crew fatigue safety performance indicators for fatigue risk management systems. *Aviation Space and Environmental Medicine.* **85** (2), 139-147 (2014).

25. Kaida, K. *et al.* Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology.* **117** (7), 1574-1581 (2006).

26. Short, M. A. *et al.* The effect of split sleep schedules (6h-on/6h-off) on neurobehavioural performance, sleep and sleepiness. *Applied Ergonomics.* **54,** 72-82 (2016).

27. Reyner, L., Horne, J. A. Falling asleep whilst driving: are drivers aware of prior sleepiness? *International Journal of Legal Medicine.* **111** (3), 120-123 (1998).

28. Samn, S., Perelli, L. Estimating Aircrew Fatigue: A Technique with Application to Airlift Operations. *Brooks Air Force Base, San Antonio, TX.* (1982).

29. Dinges, D. F., Powell, J. W. Microcomputer analysis of performance on a portable, simple visual RT task sustained operations. *Behavior Research Methods, Instruments & Computers.* **17** (6), 3 (1985).

30. Basner, M., Dinges, D. F. Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep.* **34** (5), 581-591 (2011).

31. Grant, D. A., Honn, K. A., Layton, M. E., Riedy, S. M., Van Dongen, H. P. A. 3-minute smartphone-based and tablet-based psychomotor vigilance tests for the assessment of reduced alertness due to sleep deprivation. *Behavioral Research Methods.* **49,** 9 (2017).

32. Veksler, B. Z., Gunzelmann, G. Functional equivalence of sleep loss and time on task effects in sustained attention. *Cognitive Science.* **42** (2), 32 (2018).

33. Belenky, G. *et al.* Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research.* **12** (1), 1-12 (2003).

34. Lamond, N. *et al.* The impact of a week of simulated night work on sleep, circadian phase, and performance. *Occupational Environmental Medicine.* **60** (11), e13 (2003).

35. Thorne, D. R. *et al.* The Walter Reed palm-held psychomotor vigilance test. *Behav Res Methods.* **37** (1), 111-118 (2005).

36. Dinges, D. F., Kribbs, N. B. in *Sleep, sleepiness and performance.* 97-128 *(*eds S. Folkard & T. H. Monk, John Wiley & Sons) (1991).

37. Jota, R., Ng, A., Dietz, P., Wigdor, D. How fast is fast enough?: a study of the effects of latency in direct-touch pointing tasks. *CHI Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Pages.* 2291-2300. (2013).

38. Arsintescu, L., Mulligan, J. B., Flynn-Evans, E. E. Evaluation of a Psychomotor Vigilance Task for Touch Screen Devices. *Human Factors.* **59** (4), 661-670 (2017).

39. Kay, M. *et al.* in *The 7th Conference on Pervasive Computing Technologies for Healthcare* (ed IEEE) 248-251 (Proceedings of Pervasive Health, New York, NY) (2013).

40. Bourgeois-Bougrine, S., Carbon, P., Gounelle, C., Mollard, R., Coblentz, A. Perceived fatigue for short- and long-haul flights: a survey of 739 airline pilots. *Aviation Space and Environmental Medicine.* **74** (10), 1072-1077 (2003).

41. Petrilli, R. M., Roach, G. D., Dawson, D., Lamond, N. The sleep, subjective fatigue, and sustained attention of commercial airline pilots during an international pattern. *Chronobiology International.* **23** (6), 1357-1362 (2006).

42. Rai, B., Foing, B. H., Kaur, J. Working hours, sleep, salivary cortisol, fatigue and neuro-behavior during Mars analog mission: five crews study. *Neuroscence Letters.* **516** (2), 177-181 (2012).

43. Barker, L. M., Nussbaum, M. A. Fatigue, performance and the work environment: a survey of registered nurses. *Journal of Advanced Nursing.* **67** (6), 1370-1382 (2011).

44. Samel, A., Wegmann, H. M., Vejvoda, M. Aircrew fatigue in long-haul operations. *Accidend Analysis & Prevention.* **29** (4), 439-452 (1997).

45. Reis, C., Mestre, C., Canhao, H. Prevalence of fatigue in a group of airline pilots. *Aviation Space and Environmental Medicine.* **84** (8), 828-833 (2013).

46. Marcus, J. H., Rosekind, M. R. Fatigue in transportation: NTSB investigations and safety recommendations. *Inury Prevention.* **23** (4), 232-238 (2017).

47. Goode, J. H. Are pilots at risk of accidents due to fatigue? *Journal of Safety Research.* **34** (3), 309-313 (2003).

48. O'Hagan, A. D., Issartel, J., Fletcher, R., Warrington, G. Duty hours and incidents in flight among commercial airline pilots. *International Journal of Occupational Safety and Ergonomics.* **22** (2), 165-172 (2016).

49. National Transportation Safety Board. Uncontrolled collision with terrain. *American International Airways Flight 808. NTSB/AAR-94/04.* National Transportation Safety Board, Washington, DC (1994).

50. Federal Aviation Administration. FAA Report. Regulatory Impact Analysis. *Flightcrew Member Duty and Rest Requirements PART 117.* https://www.faa.gov/regulations_policies/rulemaking/recently_published/media/2120-AJ58RegEval.pdf (2011).

51. Buysse, D. J., Reynolds, C. F. III, Monk, T. H., Berman, S. R., Kupfer, D. J. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research.* **28** (2), 193-213 (1989).

52. Krupp, L. B., LaRocca, N. G., Muir-Nash, J., Steinberg, A. D. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology.* **46** (10), 1121-1123 (1989).

53. Johns, M. W. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep.* **14** (6), 540-545 (1991).

54. Vercoulen, J. H. *et al.* Dimensional assessment of chronic fatigue syndrome. *Journal of Psychosomatic Research.* **38** (5), 383-392 (1994).

55. Horne, J. A., Ostberg, O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology.* **4** (2), 97-110 (1976).

56. Zaidi, F.H., Hull, J.T., Peirson, S.N., Wulff, K., Aeschbach, D., Gooley, J.J., Brainard, G.C., Gregory-Evans, K., Rizzo III, J.F., Czeisler, C.A., Foster, R.G. Short-wavelength light sensitivity of circadian, pupillary, and visual awareness in humans lacking an outer retina. *Current Biology.* **17** (24) 2122-8 (2007).

57. Honn, K. A., Riedy, S. M., Grant, D. A. Validation of a portable, touch-screen psychomotor vigilance test. *Aerospace Medicine and Human Performance.* **86** (5), 428-434 (2015).

58. Aldhous, M. E., Arendt, J. Radioimmunoassay for 6-sulphatoxymelatonin in urine using an iodinated tracer. *Annals Clinical Biochemistry.* **25** (Pt 3), 298-303 (1988).

59. Kecklund, G., Akerstedt, T. Effects of timing of shifts on sleepiness and sleep duration. *Journal of Sleep Research.* **4** (S2), 47-50 (1995).

60. Folkard, S., Barton, J. Does the 'forbidden zone' for sleep onset influence morning shift sleep duration? *Ergonomics.* **36** (1-3), 85-91 (1993).

61. Lavie, P. Ultrashort sleep-waking schedule. III. "Gates" and "forbidden zones" for sleep. *Electroencephalography and Clinical Neurophysiology.* **63** (5), 414-425 (1986).

62. Shekleton, J. A. *et al.* Improved neurobehavioral performance during the wake maintenance zone. *Journal of Clinical Sleep Medicine.* **9** (4), 353-362 (2013).

63. Strogatz, S. H., Kronauer, R. E., Czeisler, C. A. Circadian pacemaker interferes with sleep onset at specific times each day: role in insomnia. *American Journal of Physiology.* **253** (1 Pt 2), R172-178 (1987).