

Journal of Visualized Experiments

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

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE59851R1
Full Title:	Methods for Collecting Sleep, Circadian, Fatigue, and Performance Data in Complex Operational Environments
Keywords:	Fatigue, sleep, circadian phase, alertness, psychomotor vigilance task, PVT, NASA-PVT, melatonin, 6-sulfatoxymelatonin, actigraphy, sleep diaries
Corresponding Author:	Erin Flynn-Evans NASA Ames Research Center Moffett Field, CA UNITED STATES
Corresponding Author's Institution:	NASA Ames Research Center
Corresponding Author E-Mail:	erin.e.flynn-evans@nasa.gov
Order of Authors:	Lucia Arsintescu Kenji H Kato Cassie J Hilditch Kevin B Gregory Erin Flynn-Evans
Additional Information:	
Question	Response
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US\$2,400)
Please indicate the city, state/province, and country where this article will be filmed . Please do not use abbreviations.	Moffett Field CA USA

TITLE:

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

AUTHORS & AFFILIATIONS:

Lucia Arsintescu¹, Kenji H. Kato², Cassie J. Hilditch¹, Kevin B. Gregory¹, Erin E. Flynn-Evans³

¹San Jose State University Research Foundation, San Jose, CA, USA

²ASRS Research and Technology Solutions, Beltsville, MD, USA

³Fatigue Countermeasures Laboratory, NASA Ames Research Center, Mountain View, CA, USA

Corresponding Author:

Erin E. Flynn-Evans (erin.e.flynn-evans@nasa.gov)

Tel: (650)-279-3459

Email Addresses of Co-authors:

Lucia Arsintescu (lucia.arsintescu-1@nasa.gov)

Kenji H. Kato (kenji.h.kato@nasa.gov)

Cassie J. Hilditch (cassie.j.hilditch@nasa.gov)

Kevin B. Gregory (kevin.b.gregory@nasa.gov)

KEYWORDS:

fatigue, sleep, circadian phase, alertness, psychomotor vigilance task, NASA-PVT

SUMMARY:

Sleep loss and circadian misalignment contribute to numerous operational accidents and incidents. The effectiveness of countermeasures and work scheduling designs aimed at mitigating fatigue can be challenging to evaluate in operational environments. This manuscript summarizes an approach for collecting sleep, circadian, fatigue, and performance data in complex operational environments.

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.

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 performance³⁻⁵. These studies form the basis for fatigue risk management recommendations and work scheduling practices in operational environments⁶.

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 settings⁷. 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 operation⁸. 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 occupations¹¹⁻¹⁴ and have been validated against PSG, showing 0.85 concordance for sleep duration¹⁵. 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 hypothalamus¹⁶. 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 work¹¹. 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 environments¹⁷.

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)¹⁸ and 9-point Karolinska Sleepiness Scale (KSS)¹⁹. The SP is also commonly used in field studies of shift workers across a wide range of occupations²⁰⁻²³. The KSS has been validated against objective measures of sleepiness such as electroencephalography (EEG) and slow rolling eye movements^{19,24}, as well as performance²⁴. This scale is commonly used in studies both in the laboratory and the field^{23,25}. 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^{24,26}, while Samn-Perelli ratings relate directly to flying duties²⁷. For the study

described in this manuscript, it was chosen to assess the Samn-Perelli since it was originally developed as a subjective fatigue measure in a study population consisting of pilots.²⁷

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 possible²⁸. The PVT has been validated under conditions of acute and chronic sleep loss and circadian misalignment^{4,5,29}. The duration of the task can be varied based on the design of the study^{30,31}; although, the traditional 10 min duration is preferred in laboratory studies^{32,33}. while a 5 min duration PVT is typically more feasible in field studies where operational demands can interfere with the administration of the test³⁴.

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 testing³⁵. 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^{36,37}. 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^{37,38}.

Each of these elements of data collection is important and these tools have been used in other operational studies in the past³⁹⁻⁴². 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^{39,43,44}. The high intensity of airline operations is conducive to degradation of crew performance and increases the risk of fatigue-related incidents⁴⁵⁻⁴⁹. 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 midday: 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

1.1. Determining the study schedule

1.1.1. Include a baseline data collection period in order to assess outcomes in the absence of intervention.

1.2. Identifying when to collect data during the experiment

1.2.1. 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.

1.2.2. 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.

1.3. 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.

1.3.1. Have participants engage with study personnel directly and assure them that employment decisions will not be made based on individual data.

1.3.2. 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.

1.3.3. 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.

1.4. Invite participants who volunteered in the study to attend a 30–60 min training/briefing session.

1.4.1. 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)⁵⁰, Fatigue Severity Scale (FSS)⁵¹, Epworth Sleepiness Scale (ESS)⁵², Checklist of Individual Strength (CIS)⁵³, Morningness/Eveningness Questionnaire (MEQ)⁵⁴] to estimate the prevalence of baseline sleep issues in the population. These questionnaires may also be used as covariates in data analysis.

1.4.2. 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 each 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.

1.4.3. 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.

1.4.4. 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

2.1. 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 diagramed in **Figure 1**.

2.1.1. Data collection protocol on non-work days:

2.1.1.1. Instruct participants to complete the sleep diary each day within 30 min of waking up and going to bed.

2.1.1.2. Ask participants to wear their activity monitor continuously.

2.1.1.3. 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).

2.1.2. Data collection protocol on work days:

2.1.2.1. Instruct participants to complete the sleep diary each day within 30 min of waking up and going to bed.

2.2. 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

3.1. 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.

3.2. 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.

3.3. 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.

3.4. 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

4.1. 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.

4.2. 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.

4.3. 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.

4.3.1. 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.

4.3.2. Bedtime sleep diary: just prior to going to bed, instruct participants to document the duration and timing of any naps and anticipated bedtimes.

4.4. PVT: instruct participants to take the PVT at pre-determined times throughout the study using precise methodology as described below (in section 6).

4.5. 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

5.1. Use urine collection to measure aMT6s production to estimate circadian phase.

5.1.1. 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 jug, 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).

5.1.2. 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 four-hour blocks during the day and an eight-hour block overnight. In total, each 24 h collection includes five samples.

5.1.3. 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.

5.2. Use the following urine collection instructions for participants:

“Do not collect your first morning void on the first day of collection. When you wake up on the first day urinate in the toilet as usual. Begin collecting your urine after your first void.

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). This is fine. All of the urine produced during the collection block will mix together. 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 jug. 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 jug 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. 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 jug). 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 jug 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.

It is essential that ALL the urine produced is collected. Missing or losing any urine during any collection block can affect our ability to detect the sleep hormones in your urine. Urine can be lost because you accidentally urinate into the toilet during a collection period. If this happens, please note the time of the missed urine and contact your study coordinator to receive instructions on how to proceed.”

6. PVT administration methods

NOYR: 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^{34,55} or checkboard pattern³⁸. The NASA-PVT was designed to mimic the laboratory PVT-192 device in which the target is in the form of a milliseconds counter.

6.1. 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.

6.2. 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 time³⁷.

6.3. 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 pilots⁷. 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 examining 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 phase⁵⁶. **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 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 AND TABLE LEGENDS:

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.

Figure 2: The activity monitor/accelerometer device worn on the wrist of the non-dominant hand.

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.

Figure 4: Urine kit. The kit contains (A) a urinal hat or urinal jug, (B) pipettes, (C) urine collection tube, (D) white sticker labels, (E) a bio-hazard bag, (F) ice pack, and (G) shipping materials.

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.

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.

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.

Figure 9: aMT6 profile for the five urine collection bins for each data collection episode for a single participant. Data are double-plotted.

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

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.

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

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

Table 3: Psychomotor Vigilance Task (PVT) mean reaction time (RT), response speed (mean 1/RT), and lapses (RT >500ms) 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^{57,58}. 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 zone⁵⁹⁻⁶¹. 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.

ACKNOWLEDGMENTS:

We thank the study participants and airline personnel for their support in data collection. We also thank the members of the Fatigue Countermeasures Laboratory at NASA Ames Research

Center for their assistance with this project. This research was supported by the NASA Systemwide Safety Program.

DISCLOSURES:

EEFE is a consultant for Baby Sleep Science and she has received travel funding from Washington State University, the University of Chicago, the Puget Sound Pilots, the National Safety Council, and the American Academy of Sleep Medicine and Sleep Research Society. The other authors have no disclosures to report.

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, doi:10.1080/07420528.2017.1367305 (2017).
- 2 Williamson, A. *et al.* The link between fatigue and safety. *Accident Analysis and Prevention*. **43** (2), 498-515, doi:10.1016/j.aap.2009.11.011 (2011).
- 3 Dawson, D., Reid, K. Fatigue, alcohol and performance impairment. *Nature*. **388** (6639), 235, doi:10.1038/40775, (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, doi:10.1371/journal.pone.0001233 (2007).
- 6 Gander, P. H. Evolving Regulatory Approaches for Managing Fatigue Risk in Transport Operations. *Reviews of Human Factors and Ergonomics*. 253-271, doi:10.1177/1557234X15576510 (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, doi:10.1016/j.sleh.2018.08.002 (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. **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 International*. **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, doi:10.1038/npjmgrav.2015.19 (2016).

746 12 Gander, P., Millar, M., Webster, C., Merry, A. Sleep loss and performance of anaesthesia
747 trainees and specialists. *Chronobiology International*. **25** (6), 1077-1091 (2008).

748 13 Gander, P., van den Berg, M., Signal, L. Sleep and sleepiness of fishermen on rotating
749 schedules. *Chronobiology International*. **25** (2-3), 389-398 (2008).

750 14 Roach, G. D., Darwent, D., Sletten, T. L., Dawson, D. Long-haul pilots use in-flight napping
751 as a countermeasure to fatigue. *Applied Ergonomics*. **42** (2), 214-218,
752 doi:10.1016/j.apergo.2010.06.016 (2011).

753 15 Signal, T. L., Gale, J., Gander, P. H. Sleep measurement in flight crew: comparing
754 actigraphic and subjective estimates to polysomnography. *Aviation Space and
755 Environmental Medicine*. **76** (11), 1058-1063 (2005).

756 16 Czeisler, C. A., Gooley, J. J. Sleep and circadian rhythms in humans. *Cold Spring Harbor
757 and Symposia on Quantitative Biology*. **72**, 579-597, doi:10.1101/sqb.2007.72.064 (2007).

758 17 Bojkowski, C. J., Arendt, J., Shih, M. C., Markey, S. P. Melatonin secretion in humans
759 assessed by measuring its metabolite, 6-sulfatoxymelatonin. *Clinical Chemistry*. **33** (8),
760 1343-1348 (1987).

761 18 Samn, S. W., Perelli, L. P. Estimating aircrew fatigue: a technique with application to airlift
762 operations. *School of Aerospace Medicine Brooks AFB, TX*. (1982).

763 19 Åkerstedt, T., Gillberg, M. Subjective and objective sleepiness in the active individual.
764 *International Journal of Neuroscience*. **52** (1-2), 29-37 (1990).

765 20 Jay, S. M., Dawson, D., Ferguson, S. A., Lamond, N. Driver fatigue during extended rail
766 operations. *Applied Ergonomics*. **39** (5), 623-629 (2008).

767 21 Dorrian, J., Baulk, S. D., Dawson, D. Work hours, workload, sleep and fatigue in Australian
768 Rail Industry employees. *Applied Ergonomics*. **42** (2), 202-209,
769 doi:10.1016/j.apergo.2010.06.009 (2011).

770 22 Ferguson, S. A., Baker, A. A., Lamond, N., Kennaway, D. J., Dawson, D. Sleep in a live-in
771 mining operation: the influence of start times and restricted non-work activities. *Applied
772 Ergonomics*. **42** (1), 71-75 (2010).

773 23 Gander, P. H. *et al.* Crew fatigue safety performance indicators for fatigue risk
774 management systems. *Aviation Space and Environmental Medicine*. **85** (2), 139-147
775 (2014).

776 24 Kaida, K. *et al.* Validation of the Karolinska sleepiness scale against performance and EEG
777 variables. *Clinical Neurophysiology*. **117** (7), 1574-1581 (2006).

778 25 Short, M. A. *et al.* The effect of split sleep schedules (6h-on/6h-off) on neurobehavioural
779 performance, sleep and sleepiness. *Applied Ergonomics*. **54**, 72-82,
780 doi:10.1016/j.apergo.2015.12.004 (2016).

781 26 Reyner, L., Horne, J. A. Falling asleep whilst driving: are drivers aware of prior sleepiness?
782 *International Journal of Legal Medicine*. **111** (3), 120-123 (1998).

783 27 Samn, S., Perelli, L. Estimating Aircrew Fatigue: A Technique with Application to Airlift
784 Operations. *Brooks Air Force Base, San Antonio, TX*. (1982).

785 28 Dinges, D. F., Powell, J. W. Microcomputer analysis of performance on a portable, simple
786 visual RT task sustained operations. *Behavior Research Methods, Instruments &
787 Computers*. **17** (6), 3 (1985).

788 29 Basner, M., Dinges, D. F. Maximizing sensitivity of the psychomotor vigilance test (PVT) to
789 sleep loss. *Sleep*. **34** (5), 581-591 (2011).

790 30 Grant, D. A., Honn, K. A., Layton, M. E., Riedy, S. M., Van Dongen, H. P. A. 3-minute
791 smartphone-based and tablet-based psychomotor vigilance tests for the assessment of
792 reduced alertness due to sleep deprivation. *Behavioral Research Methods*. **49**, 9 (2017).

793 31 Veksler, B. Z., Gunzelmann, G. Functional equivalence of sleep loss and time on task
794 effects in sustained attention. *Cognitive Science*. **42** (2), 32 (2018).

795 32 Belenky, G. *et al.* Patterns of performance degradation and restoration during sleep
796 restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep*
797 *Research*. **12** (1), 1-12 (2003).

798 33 Lamond, N. *et al.* The impact of a week of simulated night work on sleep, circadian phase,
799 and performance. *Occupational Environmental Medicine*. **60** (11), e13 (2003).

800 34 Thorne, D. R. *et al.* The Walter Reed palm-held psychomotor vigilance test. *Behav Res*
801 *Methods*. **37** (1), 111-118 (2005).

802 35 Dinges, D. F., Kribbs, N. B. in *Sleep, sleepiness and performance*. 97-128 (eds S. Folkard &
803 T. H. Monk, John Wiley & Sons). (1991).

804 36 Jota, R., Ng, A., Dietz, P., Wigdor, D. in *CHI Proceedings of the SIGCHI Conference on*
805 *Human Factors in Computing Systems Pages*. 2291-2300.

806 37 Arsintescu, L., Mulligan, J. B., Flynn-Evans, E. E. Evaluation of a Psychomotor Vigilance
807 Task for Touch Screen Devices. *Human Factors*. **59** (4), 661-670,
808 doi:10.1177/0018720816688394 (2017).

809 38 Kay, M. *et al.* in *The 7th Conference on Pervasive Computing Technologies for Healthcare*
810 (ed IEEE) 248-251 (Proceedings of Pervasive Health, New York, NY). (2013).

811 39 Bourgeois-Bougrine, S., Carbon, P., Gounelle, C., Mollard, R., Coblentz, A. Perceived
812 fatigue for short- and long-haul flights: a survey of 739 airline pilots. *Aviation Space and*
813 *Environmental Medicine*. **74** (10), 1072-1077 (2003).

814 40 Petrilli, R. M., Roach, G. D., Dawson, D., Lamond, N. The sleep, subjective fatigue, and
815 sustained attention of commercial airline pilots during an international pattern.
816 *Chronobiology International*. **23** (6), 1357-1362, doi:10.1080/07420520601085925
817 (2006).

818 41 Rai, B., Foing, B. H., Kaur, J. Working hours, sleep, salivary cortisol, fatigue and neuro-
819 behavior during Mars analog mission: five crews study. *Neuroscience Letters*. **516** (2), 177-
820 181, doi:10.1016/j.neulet.2012.03.067 (2012).

821 42 Barker, L. M., Nussbaum, M. A. Fatigue, performance and the work environment: a survey
822 of registered nurses. *Journal of Advanced Nursing*. **67** (6), 1370-1382, doi:10.1111/j.1365-
823 2648.2010.05597.x (2011).

824 43 Samel, A., Wegmann, H. M., Vejvoda, M. Aircrew fatigue in long-haul operations. *Accident*
825 *Analysis & Prevention*. **29** (4), 439-452 (1997).

826 44 Reis, C., Mestre, C., Canhao, H. Prevalence of fatigue in a group of airline pilots. *Aviation*
827 *Space and Environmental Medicine*. **84** (8), 828-833 (2013).

828 45 Marcus, J. H., Rosekind, M. R. Fatigue in transportation: NTSB investigations and safety
829 recommendations. *Injury Prevention*. **23** (4), 232-238, doi:10.1136/injuryprev-2015-
830 041791 (2017).

831 46 Goode, J. H. Are pilots at risk of accidents due to fatigue? *Journal of Safety Research*. **34**
832 (3), 309-313 (2003).

833 47 O'Hagan, A. D., Issartel, J., Fletcher, R., Warrington, G. Duty hours and incidents in flight
834 among commercial airline pilots. *International Journal of Occupational Safety and*
835 *Ergonomics*. **22** (2), 165-172 (2016).

836 48 National Transportation Safety Board. National Transportation Safety Board, Washington
837 DC. (1994).

838 49 Federal Aviation Administration. Department of Transportation, Washington DC (2011).

839 50 Buysse, D. J., Reynolds, C. F. III, Monk, T. H., Berman, S. R., Kupfer, D. J. The Pittsburgh
840 Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry*
841 *Research*. **28** (2), 193-213 (1989).

842 51 Krupp, L. B., LaRocca, N. G., Muir-Nash, J., Steinberg, A. D. The fatigue severity scale.
843 Application to patients with multiple sclerosis and systemic lupus erythematosus.
844 *Archives of Neurology*. **46** (10), 1121-1123 (1989).

845 52 Johns, M. W. A new method for measuring daytime sleepiness: the Epworth sleepiness
846 scale. *Sleep*. **14** (6), 540-545 (1991).

847 53 Vercoulen, J. H. *et al.* Dimensional assessment of chronic fatigue syndrome. *Journal of*
848 *Psychosomatic Research*. **38** (5), 383-392 (1994).

849 54 Horne, J. A., Ostberg, O. A self-assessment questionnaire to determine morningness-
850 eveningness in human circadian rhythms. *International Journal of Chronobiology*. **4** (2),
851 97-110 (1976).

852 55 Honn, K. A., Riedy, S. M., Grant, D. A. Validation of a portable, touch-screen psychomotor
853 vigilance test. *Aerospace Medicine and Human Performance*. **86** (5), 428-434,
854 doi:10.3357/AMHP.4165.2015 (2015).

855 56 Aldhous, M. E., Arendt, J. Radioimmunoassay for 6-sulphatoxymelatonin in urine using an
856 iodinated tracer. *Annals Clinical Biochemistry*. **25** (Pt 3), 298-303 (1988).

857 57 Kecklund, G., Akerstedt, T. Effects of timing of shifts on sleepiness and sleep duration.
858 *Journal of Sleep Research*. **4** (S2), 47-50 (1995).

859 58 Folkard, S., Barton, J. Does the 'forbidden zone' for sleep onset influence morning shift
860 sleep duration? *Ergonomics*. **36** (1-3), 85-91 (1993).

861 59 Lavie, P. Ultrashort sleep-waking schedule. III. "Gates" and "forbidden zones" for sleep.
862 *Electroencephalography and Clinical Neurophysiology*. **63** (5), 414-425 (1986).

863 60 Shekleton, J. A. *et al.* Improved neurobehavioral performance during the wake
864 maintenance zone. *Journal of Clinical Sleep Medicine*. **9** (4), 353-362,
865 doi:10.5664/jcsm.2588 (2013).

866 61 Strogatz, S. H., Kronauer, R. E., Czeisler, C. A. Circadian pacemaker interferes with sleep
867 onset at specific times each day: role in insomnia. *American Journal of Physiology*. **253** (1
868 Pt 2), R172-178 (1987).

Figure 1

[Click here to access/download;Figure;Figure 1.tiff](#) 



Figure 2

[Click here to access/download;Figure;Figure 2.tiff](#) 



Figure 3

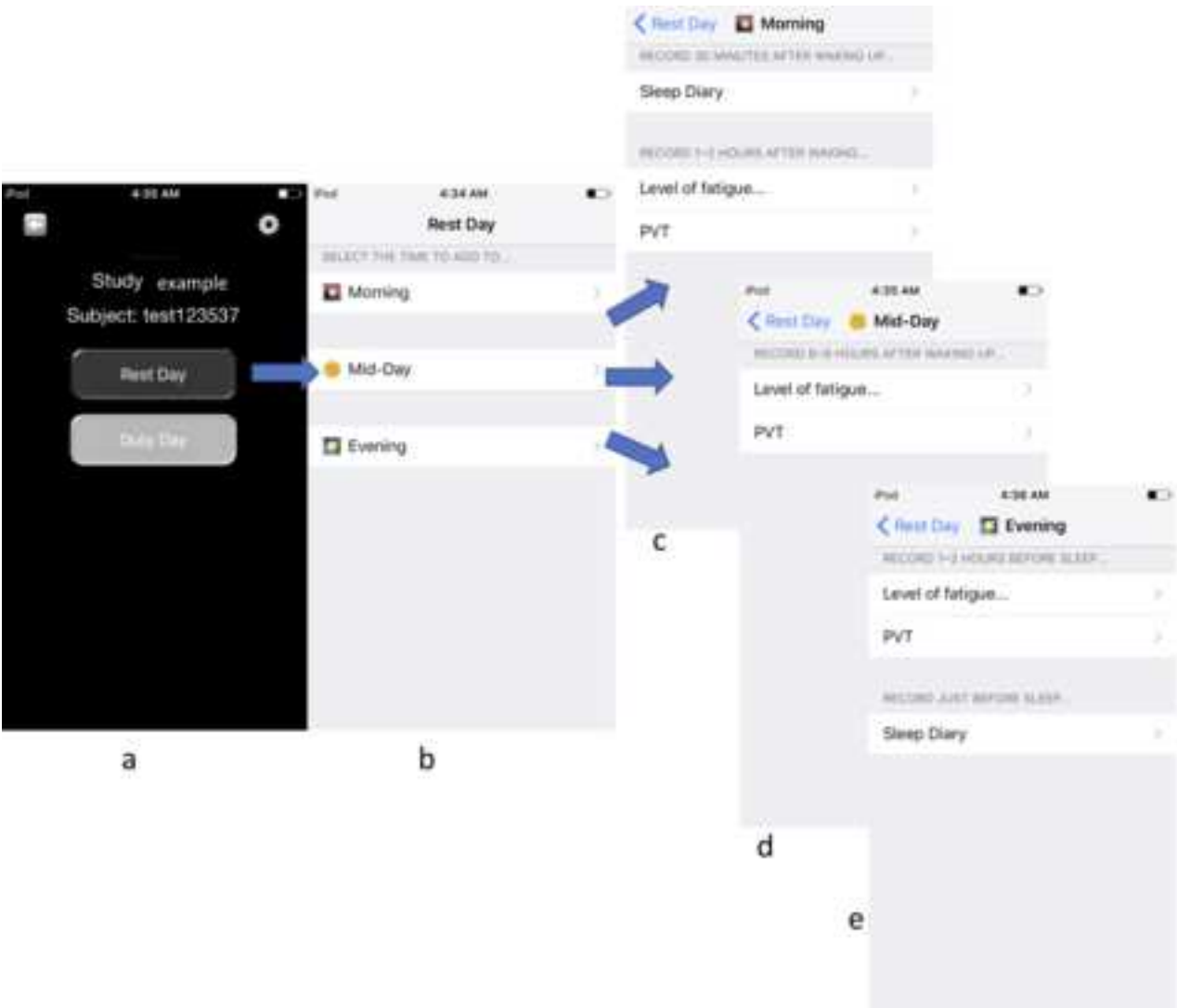
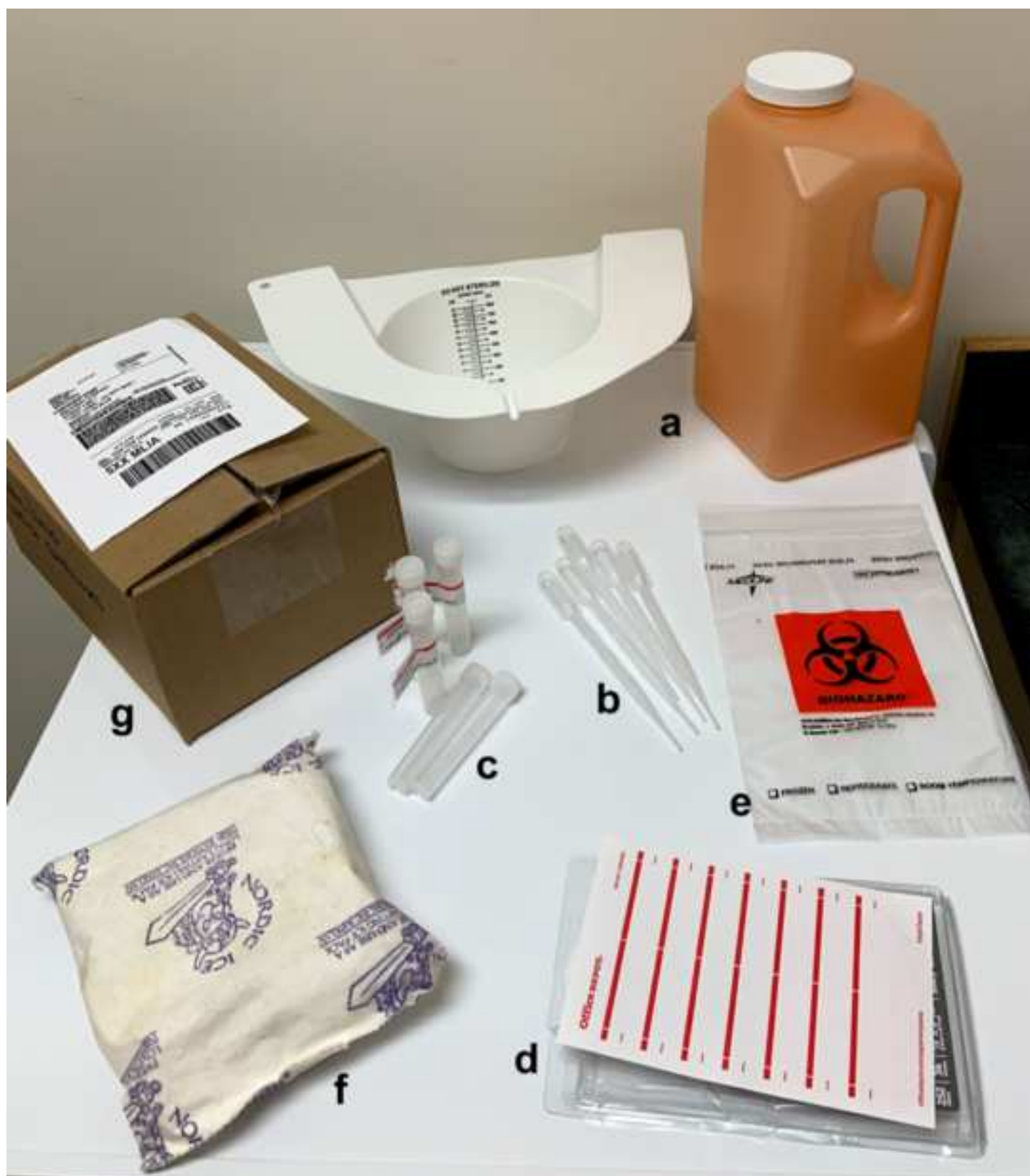


Figure 4

[Click here to access/download;Figure;Figure 4.tiff](#)



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

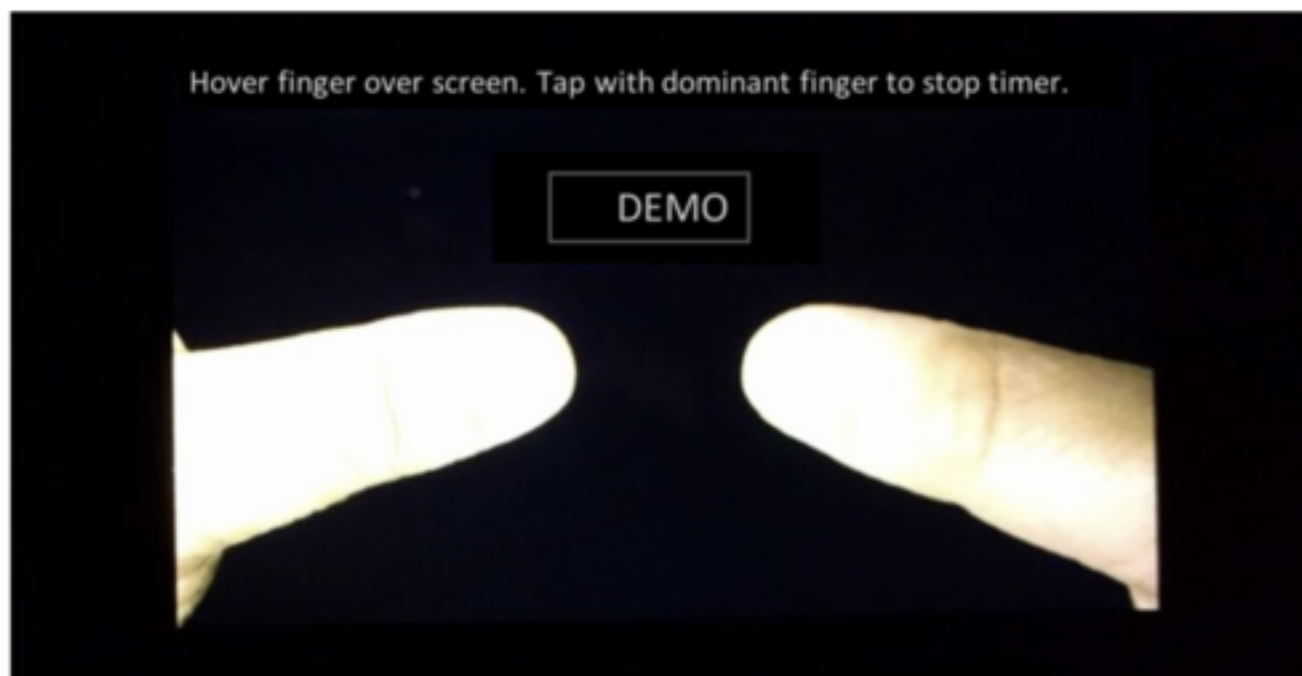
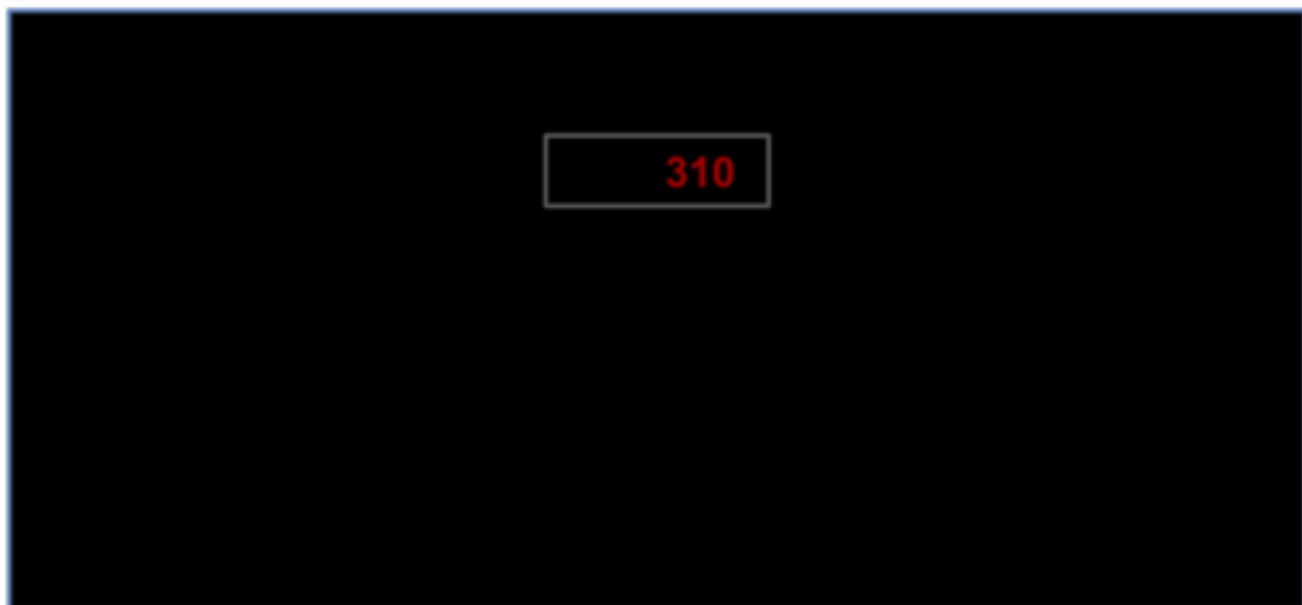
**a****b**

Figure 7

[Click here to access/download;Figure;Figure 7.tiff](#)

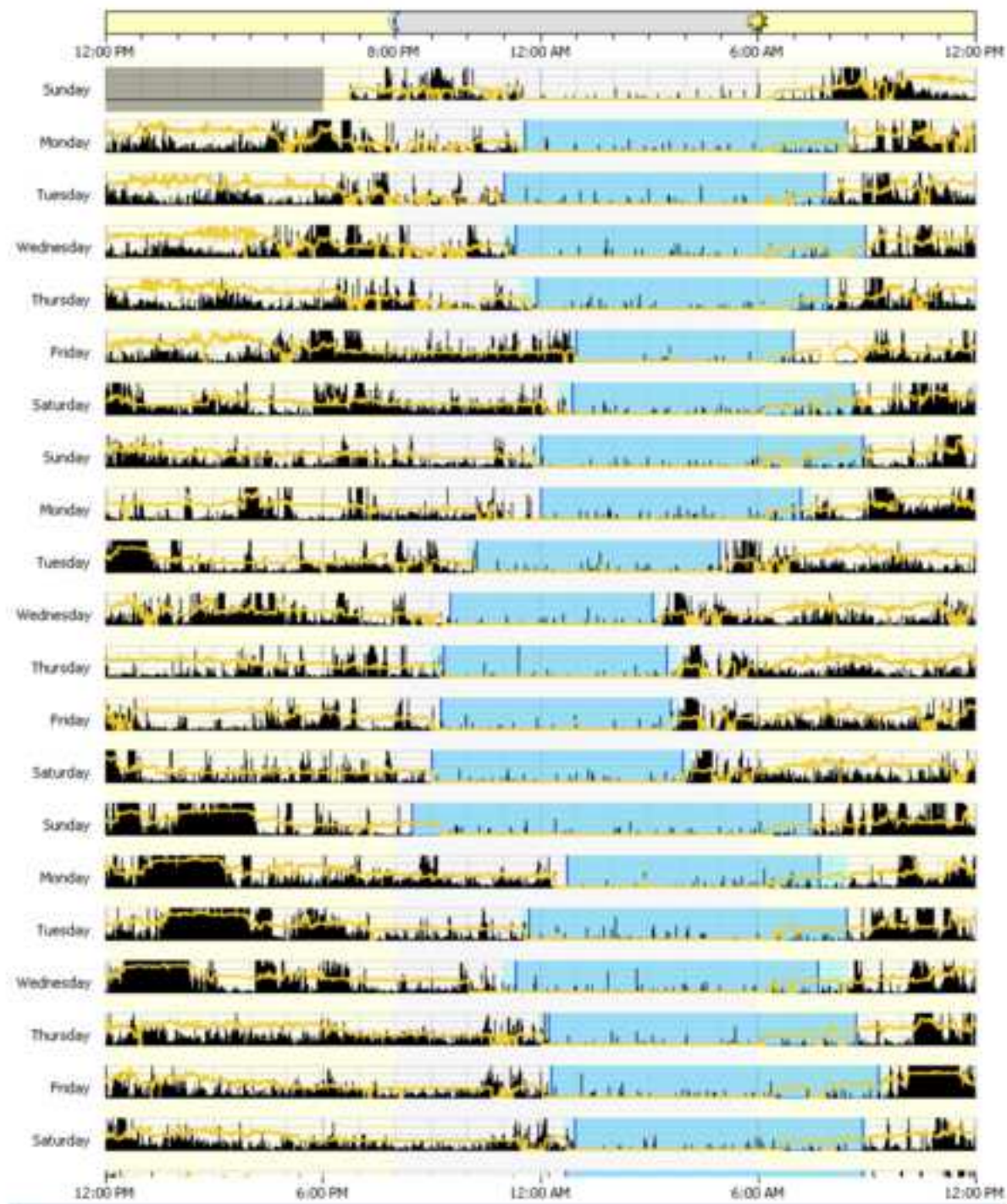


Figure 8

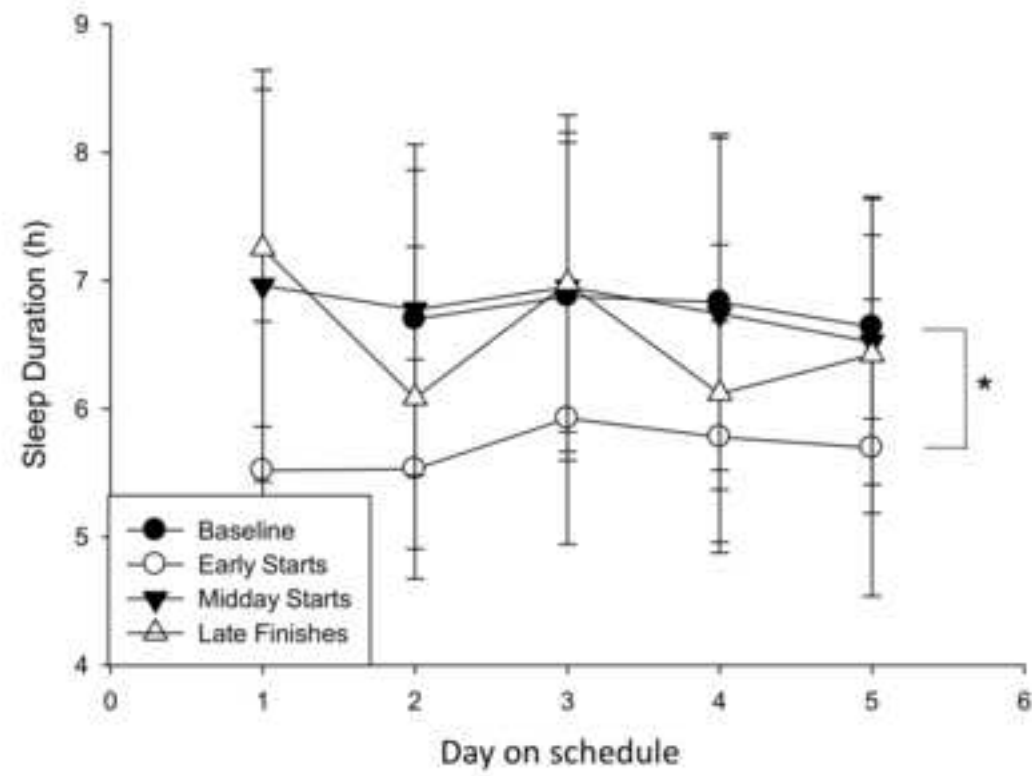
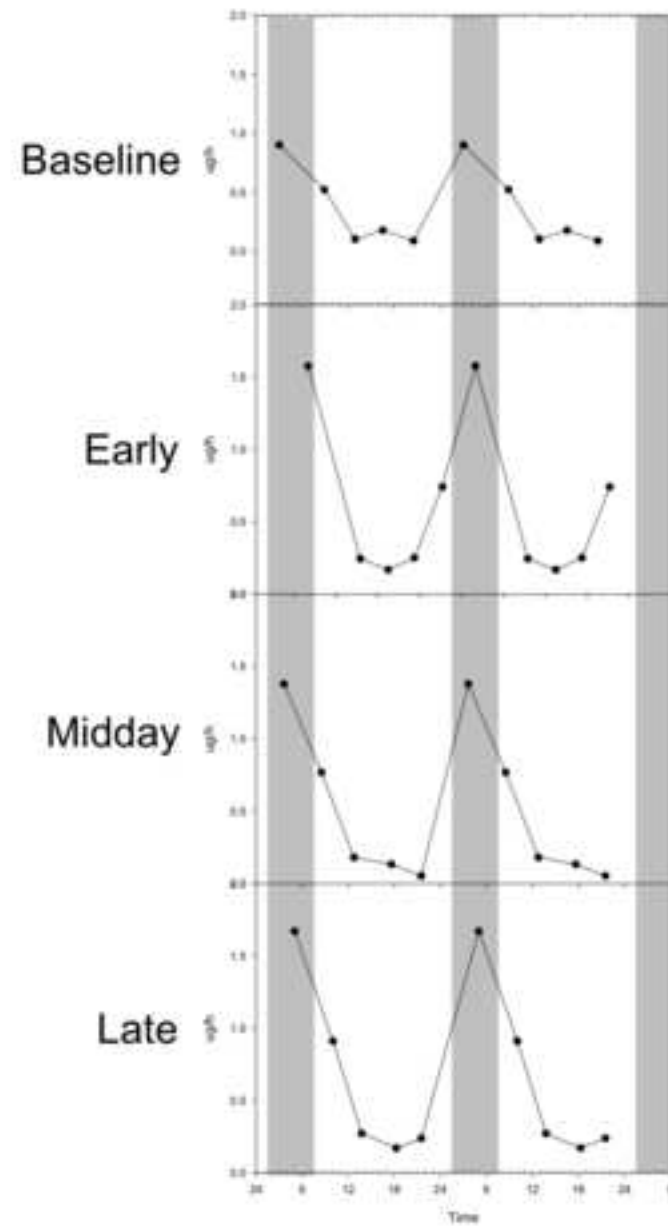
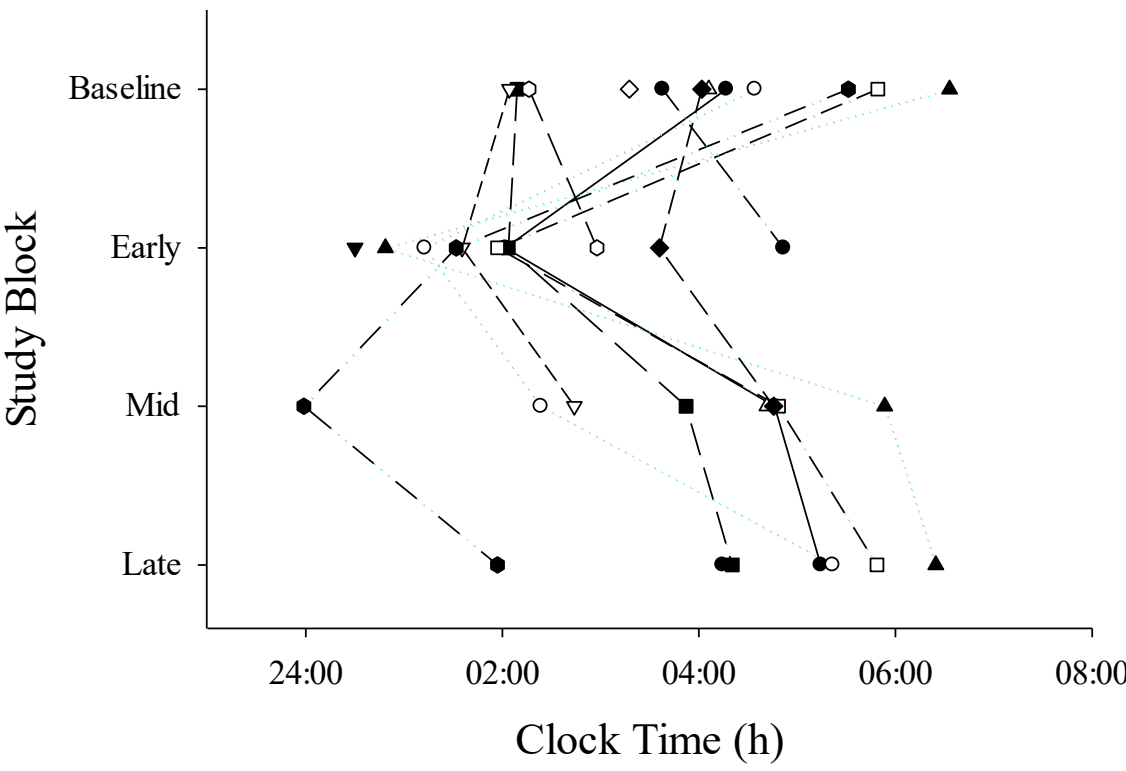
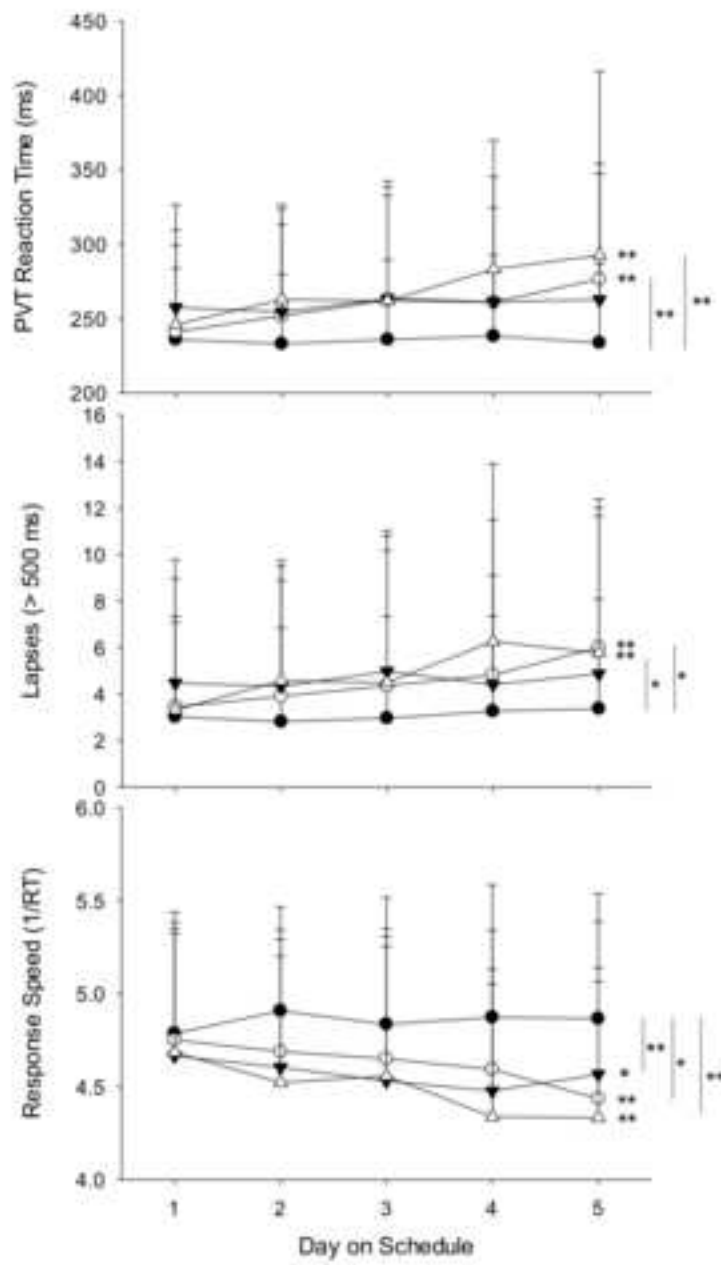


Figure 9







Work schedule	n	Bedtime (h, SD)
Baseline (ref.)	39	23:10 (1:41)
Early	42	21:14 (1:01)**
Midday	41	01:19 (0:43)**
Late	40	02:18 (1:07)**
Rest days	42	23:47 (0:50)**

Wake time (h, SD)	Sleep duration (h, SD)	Sleep quality (SD)
7:20 (1:49)	8.2 (0.9)	2.4 (0.7)
4:29 (0:47)	7.4 (0.9)**	2.5 (0.6)
9:11 (0:58)	7.9 (1.1)	2.3 (0.6)
9:57 (1:11)	7.8 (1.4)*	2.3 (0.7)
8:16 (0:58)	8.5 (0.9)*	2.4 (0.5)

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

Work schedule	n (participant	Mean reaction time (ms, SD)
Baseline (ref.)	38	236 (48)
Early	40	257 (70)**
Midday	39	261 (62)**
Late	38	266 (64)**
Rest days	40	249 (56)

Response speed (s, SD)	Mean Lapses (> 500 ms)
4.84 (0.61)	3.1 (4.1)
4.63 (0.66)**	4.4 (5.4)**
4.56 (0.66)**	4.7 (5.1)*
4.51 (0.63)**	4.7 (5.0)**
4.69 (0.62)	4.0 (4.5)

Name of Material/Equipment	Company	Catalog Number	Comments/Description
Actiwatch Spectrum Pro	Philips Respironics, Bend OR, USA	1099351	The number listed in the Catalog Number section
iPod Touch 5Th gen	Apple Inc., Cupertino CA, USA	A1509	The number listed in the Catalog Number section
Medline DYND30261 Zip-Style Biohazard Specimen Bags, Plastic, Latex Free, 9" Length, 6" Width, Clear	Medline Industries, Inc., Northfield IL	DYND30261	The number listed in the catalog Number section
Medline DYND80024 24 hours Urine Collection Bottle, 3000 mL	Medline Industries, Inc., Northfield IL	DYND80024	The number listed in the catalog Number section
Moveland 3ml Disposable Plastic Transfer Pipettes	Moveland		
Nordic Ice NOR1038 No-Sweat Reusable Long-Lasting Gel Pack, 16 oz. (Pack of 3)	Nordic Cold Chain Solutions	0858687005050	
Office Depot Brand Print-Or-Write Color Permanent Inkjet/Laser File Folder Labels, OD98817, 5/8" x 3 1/2", Dark Blue	Office Depot, Inc.Boca Raton FL, USA	660-426	
Philips Actiware 6.0.9	Respironics, Inc., Murrysville PA, USA	1104776	This software is used to analyze sleep recorded t
Push cap, neutral for 7 mL tubes	Sarstedt, Numbrecht, Germany	65.793	
SAS software 9.4	SAS Institute, Cary, NC	https://www.sas.com/en_us	This software is used to analyze the data. Any st
Shipping material	FedEx, USPS, UPS		Any company can be used.
Specimen Collector Urine/Stool White 26 oz.	McKesson Corporation, San Francisco CA	16-9522	The number listed in the catalog Number section
Tube 7 mL, 50x16mm, PS	Sarstedt, Numbrecht, Germany	58.485	

n is the Reference number for Actiwatch Spectrum Pro.

n is the Model number. Newer generations of iPods can be used for data collection.

n is the Part number

n is the Part number

through Actiwatch Spectrum Pro

statistical software (e.g., SPSS, R) can be used.

n is the Part number

612542.6 For questions, please contact us at submissions@jove.com or +1.617.945.9051.

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: **"Agreement"** means this Article and Video License Agreement; **"Article"** means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; **"Author"** means the author who is a signatory to this Agreement; **"Collective Work"** means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; **"CRC License"** means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: <http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode>; **"Derivative Work"** means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; **"institution"** means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; **"JOVE"** means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; **"Materials"** means the Article and / or the Video; **"Parties"** means the Author and JOVE; **"Video"** means any video(s) made by the Author, alone or in conjunction with any other parties, or by JOVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion

of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JOVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JOVE agreeing to publish the Article, the Author hereby grants to JOVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JOVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the

ARTICLE AND VIDEO LICENSE AGREEMENT

☒ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

Item 2: Please select one of the following items:

Item 1: The Author elects to have the Materials be made available (as described at <http://www.jove.com/publish>) via:

☒ Standard Access

☐ Open Access

Author(s):

Title of Article:

Methods for collecting data in complex operational environments, fatigue, and performance data

ARTICLE AND VIDEO LICENSE AGREEMENT

1 Alewife Center #200
Cambridge, MA 02140
Tel. 617.945.9051
www.jove.com

love
JOURNAL OF
VISUALIZED EXPERIMENTS

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JOVE in Section 3 above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the institution's website or the Author's personal website, in each case provided that a link to the Article on the JOVE website is provided and notice of JOVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This Section 5 applies if the "Standard Access" box has been checked in Item 1 above or if no box has been checked in Item 1 above. In consideration of JOVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to Section 7 below, JOVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JOVE.

6. **Grant of Rights in Video – Open Access.** This Section 6 applies only if the "Open Access" box has been checked in Item 1 above. In consideration of JOVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JOVE, subject to Section 7 below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this Section 6 is applicable, JOVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in Item 2 above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum

rights permitted under such statute. In such case, all such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JOVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JOVE to take steps in the Author(s) name and on their behalf if JOVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JOVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the safe, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JOVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and all other regulations, including, without limitation, animal treatment, laboratory, hospital, ethical, human and laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JOVE Discretion.** If the Author requests the assistance of JOVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JOVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JOVE may, in its sole

ARTICLE AND VIDEO LICENSE AGREEMENT

1 Alewife Center #200
Cambridge, MA 02140
Tel. 617 945.9051
www.jove.com



discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. Jove reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to Jove. Jove and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. Jove has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. Indemnification. The Author agrees to indemnify Jove and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless Jove from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damages to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to Jove, making of videos by Jove, or publication in Jove or elsewhere by Jove. The Author shall be responsible for, and shall hold Jove harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to

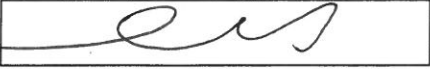
the making of a video by Jove its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include Jove's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of Jove, its employees, agents or independent contractors.

13. Fees. To cover the cost incurred for publication, Jove must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author or any submitted Materials after final peer review approval will result in a US\$1,200 fee to cover pre-production expenses incurred by Jove. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. Transfer, Governing Law. This Agreement may be assigned by Jove and shall inure to the benefit of any of Jove's successors and assigns. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Erin Flynn-Evans
Department:	Fatigue Countermeasures Laboratory
Institution:	NASA Ames Research Center
Title:	Director
Signature:	
Date:	2/14/2019

Please submit a signed and dated copy of this license by one of the following three methods:

1. Upload an electronic version on the Jove submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to Jove / Attn: Jove Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

612542.6 For questions, please contact us at submissions@jove.com or +1.617.945.9051.

March 14, 2019

Phillip Steindel, PhD

Review Editor

JoVE

Dear Dr. Steindel:

Please find enclosed our revised manuscript entitled “Methods for Collecting Sleep, Circadian, Fatigue, and Performance Data in Complex Operational Environments” for consideration in the Journal of Visual Experiments. We have revised the manuscript as recommended by editors and reviewers and have responded to their comments point-by-point below. We have italicized the reviewer comments to distinguish them from our responses. We hope that you will find our revised rationale clear.

Best regards,

Erin E. Flynn-Evans, PhD

Editorial comments:

General:

- 1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues.*

[We checked for spelling or grammar issues.](#)

- 2. Please ensure that the manuscript is formatted according to JoVE guidelines—letter (8.5” x 11”) page size, 1-inch margins, 12 pt Calibri font throughout, all text aligned to the left margin, single spacing within paragraphs, and spaces between all paragraphs and protocol steps/substeps.*

[The manuscript is formatted according to JoVE guidelines.](#)

- 3. JoVE cannot publish manuscripts containing commercial language. This includes trademark symbols (™), registered symbols (®), and company names before an*

instrument or reagent. Please limit the use of commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents.

For example: Actiwatch, Apple, iPod,

We removed the commercial language. We replaced the word Actiwatch with activity monitor and the word iPod with touchscreen device. We also removed the word Apple.

Protocol:

- 1. Please ensure that all text in the protocol section is written in the imperative tense as if telling someone how to do the technique (e.g., “Do this,” “Ensure that,” etc.). The actions should be described in the imperative tense in complete sentences wherever possible. Avoid usage of phrases such as “could be,” “should be,” and “would be” throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a “Note.” However, notes should be concise and used sparingly.*

We made the changes requested by the editor. However, there were situations where we described specific methods used in our study and therefore we included those descriptions as Notes but did not use the imperative tense.

- 2. Please include an ethics statement before the numbered protocol steps, indicating that the protocol follows the guidelines of your institution’s human research ethics committee.*

Ethics statement is included now at the beginning of the protocol (page 4, lines 200-202).

- 3. Protocol: Please adjust the numbering of the Protocol to follow the JoVE Instructions for Authors. For example, 1 should be followed by 1.1 and then 1.1.1 and 1.1.2 if necessary.*

We adjusted the numbering of the protocol to follow JoVE instructions.

- 4. Please split up longer Protocol steps so that individual steps contain only 2–3 actions and a maximum of 4 sentences.*

We tried to split up the steps to contain 2-3 actions. However, for the instructions of the urine collection this was not possible. Those instructions are given to the participants at home to refer to them whenever needed. If preferred, we could include that documentation as supplementary material.

5. *There is a 10 page limit for the Protocol, but there is a 2.75 page limit for filmable content. Please highlight 2.75 pages or less of the Protocol (including headers and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol. Remember that non-highlighted Protocol steps will remain in the manuscript, and therefore will still be available to the reader.*

We highlighted the filmable content in yellow.

6. *For each protocol step, please ensure you answer the “how” question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action. If revisions cause a step to have more than 2-3 actions and 4 sentences per step, please split into separate steps or substeps.*

We have modified the protocol steps to provide more detail on how each step is performed.

Specific Protocol steps:

1. 4: *This is unclear-are these collection methods all done using the same app? Can you include an example of the questionnaire?*

The collection methods use the same app. The PVT, questionnaire and sleep diary are all part of the app. We provided a reference to one of the questionnaires (Samn-Perelli). We revised this section to indicate that these data are collected using a single app.

Results:

1. *What statistical tests were used to obtain the p-values?*

We added a paragraph about the statistical test used for analyses (page 11, lines 518-524).

Discussion:

1. *Discussion: As we are a methods journal, please revise the Discussion to explicitly cover the following in detail in 3–6 paragraphs with citations:*
 - a) *Critical steps within the protocol*
 - b) *Any modifications and troubleshooting of the technique*
 - c) *Any limitations of the technique*

We have modified the discussion to include two paragraphs describing the critical steps in the protocol and the troubleshooting strategies and limitations (page 14, lines 629-639; page 15, lines 680-695).

References:

1. Please do not abbreviate journal titles.

We changed the journal titles to their full name.

Table of Materials:

1. Please ensure the Table of Materials has information on all materials and equipment used, especially those mentioned in the Protocol.

We made a new revision of the Table of Materials and included all materials used in the protocol.

Reviewers' comments:

Reviewer #1:

We thank the reviewer for taking time to read our manuscript and for providing valuable suggestions.

Manuscript Summary:

This manuscript details methods for collecting circadian, fatigue, sleep, and performance data in difficult-to-study field environments. This paper is unique and highly useful. Conventional peer-reviewed manuscripts are not often afforded the space to capture precise methodologies at this level of detail, and this reviewer considers this to be quite valuable to the field. Conducting field studies necessarily creates data with a lot more noise, and having standardized methods of collection will help reduce variability and increase generalizability across studies. Additionally, the representative data is from an unusually well-designed field study, in that the conditions are distinct, separated by time, and within-subjects. Further, the authors nicely dissociate the individual constructs of sleepiness, fatigue, and performance, which are often reduced to a single measure in many studies (eg, KSS only).

Major Concerns:

None

Minor Concerns:

1. I might consider adding the Actiwatch software and license to your materials list, since it is an expenditure that is necessary to download/interpret data from those devices.

As the reviewer suggested we added the Actiware software to the Materials list.

2. Consider removing the red "Active" markings from the actogram, or else describe them.

The red color represents the time when the participant is awake. However, because they may create confusion in reading the actogram we provided a figure without the red markings. (Figure 7).

3. I believe this will largely be resolved via the video, but the description of the app in 4.4.c was a little confusing.

We reworded this section and we hope it is more explicit now (page 8, lines 386-393, section 4 – 5.2., 5.3.). We also included a new figure (Figure 3) showing the flow of the app on a day off.

4. A little more detail regarding how the participants are notified about the diary within 30 min of waking could be good - how does the investigator know they are awake?

We apologize for this confusion. The app does not actively prompt the participant to complete the sleep diary within 30 minutes of waking, because, as the reviewer notes, it is impossible to predict when someone will wake up. Instead, the app flow is designed such that after an individual completes a task, the next task in the sequence is presented on the screen. The participants are told at the beginning of the study that they have to fill in the sleep diary within 30 minutes of being awake or going to bed. In addition, the sleep diary title displays in parentheses the words “within 30 minutes of waking up” or “within 30 minutes of going to bed”. Each test that is taken on the iPod is time-stamped and therefore the investigator can verify the times when the tests were taken. We have deleted the word “prompt” and have clarified this in the protocol description and have also included a new figure (Figure 3) to demonstrate how the app works.

Reviewer #2:

We thank the reviewer for taking time to read our manuscript and for providing valuable suggestions.

Manuscript Summary:

Description of the problem with using lab-validated fatigue measures in a complex operational environment. Suggested protocol for capturing subjective, objective, and physiological data across four representative flight schedules and analysis of findings. Discussion indicates some limitations with the protocol (primarily compliance) along with benefits for collecting multiple measures.

Major Concerns:

None. Intro, protocol, and discussion are very well explained.

Minor Concerns:

1) There is equal discussion of KSS as well as SP in the introduction, but only SP was used in the study with no further mention of KSS. Recommend adding a line explaining why SP was used over (or in addition to) KSS.

This is a good point. We use both scales in our studies depending on the outcomes that we are trying to measure. As the reviewer may know, the Samn-Perelli Fatigue Checklist was originally designed as a simple measure of subjective fatigue in pilots while Karolinska Sleepiness Scale was designed to measure sleepiness in the moment. Although studies showed that they are highly correlated, and it is hard to distinguish between sleepiness and fatigue we chose SP because it was developed mainly for pilots and used in numerous aviation studies. We added this explanation to the manuscript (page 3, lines 142-143).

2) Figures: Figure 5 is not an informative image of the iPod PVT screen, recommend replacing with better image.

Figure 5 (in the current revision Figure 6) represents the starting page of any PVT to remind participants about the landscape orientation of the device and the use of the thumb in answering to the PVT stimuli. To provide more clarification we added a second picture of the PVT task itself showing the display of the stimulus on the screen.

Figure 7 y-axis needs correcting/clarification, does it refer to actual clock time or hours awake?

The y-axis in Figure 7 refers to the actual clock time. Upon re-reading the manuscript, we felt that this figure was not reflective of all of the measures obtained in the sleep diary. We replaced this figure with a table (Table 1) describing all of the self-reported sleep outcomes.

Figure 10 is missing a legend and would benefit from an alternate presentation form, perhaps break into four graphs by schedule type.

This figure is reproduced from the manuscript that summarizes the scientific findings presented in this manuscript (Flynn-Evans et al. Sleep Health 2018). As a result, we feel it would be best to keep the figure in its current form, so that a reader may compare the scientific results directly against the methods presented in this paper. We apologize for the omission of the figure legend. We have added it and hope that it clarifies the data contained in the figure (page 13, lines 602-604)

Figure 11 is missing a legend and the significant differences symbology is a bit challenging to interpret - would this work better in table form?

We added the legend (page 13, lines 607-612) for Figure 11. In addition, we added Table 3 to support the visual depiction of the PVT findings in Figure 11.