An Evaluation of Sleepiness, Performance, and Workload Among Operators During a Real-Time Reactive Telerobotic Lunar Mission Simulation

Zachary Glaros, MA[1,2], Robert E. Carvalho, BS[3], Erin E. Flynn-Evans, PhD MPH[1], [1]Fatigue Countermeasures Laboratory, Human Systems Integration Division, NASA Ames Research Center, Mountain View, CA, USA, [2]San José State University, San José, CA, USA, [3]Intelligent Systems Division, Mountain View, CA, USA

Objective: We assessed operator performance during a real-time reactive telerobotic lunar mission simulation to understand how daytime versus nighttime operations might affect sleepiness, performance, and workload.

Background: Control center operations present factors that can influence sleepiness, neurobehavioral performance, and workload. Each spaceflight mission poses unique challenges that make it difficult to predict how long operators can safely and accurately conduct operations. We aimed to evaluate the performance impact of time-on-task and time-of-day using a simulated telerobotic lunar rover to better inform staffing and scheduling needs for the upcoming Volatiles Investigating Polar Exploration Rover (VIPER) mission.

Methods: We studied seven trained operators in a simulated mission control environment. Operators completed two five-hour simulations in a randomized order, beginning at noon and midnight. Performance was evaluated every 25 minutes using the Karolinska Sleepiness Scale, Psychomotor Vigilance Task, and NASA Task Load Index.

Results: Participants rated themselves as sleepier (5.06 ± 2.28) on the midnight compared to the noon simulation (3.12 ± 1.44; $p < .001$). Reaction time worsened over time during the midnight simulation but did not vary between simulations. Workload was rated higher during the noon (37.93 ± 20.09) compared to the midnight simulation (32.09 ± 21.74; $p = .007$).

Conclusion: Our findings suggest that work shifts during future operations should be limited in duration to minimize sleepiness. Our findings also suggest that working during the day, when distractions are present, increases perceived workload. Further research is needed to understand how working consecutive shifts and taking breaks within a shift influence performance.

Address correspondence to Zachary Glaros, Fatigue Countermeasures Laboratory, Human Systems Integration Division, NASA Ames Research Center, Moffett field, Mountain View, CA 94035, USA.
Email: zachary.l.glaros@nasa.gov

HUMAN FACTORS
Vol. 0, No. 0, ■■ ■, pp. 1-10
DOI:10.1177/00187208211056756
Article reuse guidelines: sagepub.com/journals-rmissions

Keywords: fatigue, workload, teleoperation, vigilance

INTRODUCTION

Telerobotic operations, in which an operator is physically controlling a device from a remote location, are increasing in use in many occupational environments (e.g., military operations, robotic surgery, and search and rescue), but most notably for extraplanetary exploration (Chen et al., 2007). This inherent separation from the device that is being controlled deprives the operator of sensory feedback and requires the operator to instead rely on visual displays, which can increase mental workload demand (Schipani, 2003). Often, this involves vigilant monitoring of displays for extended durations of time to preserve the robot and mission. In addition, spaceflight missions typically require 24-hour operations. Many studies have demonstrated that working at night is associated with performance impairment due to working under conditions of sleep loss or when the circadian rhythm is promoting sleep (Åkerstedt & Wright, 2009; Boivin & Boudreau, 2014). However, it is unclear how the demands of performing a sustained, continuous telerobotic operation might interact with sleep loss and circadian misalignment to influence an operator's ability to maintain performance. Furthermore, it is unclear whether an operator's time-on-task should be limited due to these factors. The upcoming Volatiles Investigating Polar Exploration Rover (VIPER) mission will require teams of human operators to control a lunar rover remotely from an Earth-based mission control center. Due to the relative lack of prior research on sustained real-time reactive mission control operations, we aimed to evaluate sleepiness, performance, and workload during a simulated operation to better inform scheduling and staffing requirements in preparation for the VIPER mission.

METHODS

Participants

This study was approved by the NASA Ames Human Research Institutional Review Board (HRI-359). A total of 16 individuals were trained to operate a lunar telerobotic simulation at the time of data collection and were thus invited to participate. Of the total operators, seven (n = 1 female) volunteered and provided informed consent.

Participants completed a background questionnaire that included general demographic information; the Morningness-Eveningness Questionnaire (MEQ) was used to determine the diurnal preference of the participant (Duffy et al., 2001; Horne & Östberg, 1976). The Pittsburgh Sleep Quality Index (PSQI) and Epworth Sleepiness Scale (ESS) were used to determine the likelihood of an individual having a sleep disorder (Buysse et al., 1989; Johns, 1991).

During the initial meeting, each participant was given an activity monitor (Actiwatch Spectrum, Respironics Inc®, Bend, OR, USA) to wear on their non-dominant wrist for one week prior to each simulation. The Actiwatch was used to evaluate participants' sleep timing, duration, and quality before and during the simulation. Participants were also asked to complete a daily sleep diary to document sleep, including naps, to supplement the actigraphy analysis.

Simulation

Each simulation required a pair of two role-types: that of a driver, and a "real-time" science operator (specific descriptions provided below). Due to the uneven distribution of drivers and real-time scientists, two researchers from the Fatigue Countermeasures Laboratory at NASA Ames Research Center were trained to perform the duties of the real-time scientist and completed all study procedures but were not included in analyses. The study involved each pair of operators completing two different simulations (described below). The driver was always paired with the same real-time scientist or confederate during the simulation runs.

The role of the telerobotic driver was to operate and manage the safety of the rover. During the simulation, the driver viewed a 3D projected world space of the moon. There were two levels of the world space: (1) the drivable terrain, and (2) the projected image of the surface of the moon. A high-resolution visual environment of the moon's surface was displayed (i.e., 1 km x 1 km), although the drivers only had a portion of that (i.e., 320 m x 320 m) available for driving. If the rover was driven past the physical terrain space, the simulation would error; therefore, drivers were instructed to remain within the mapped simulation. The driver issued a series of commands to the rover, such as precise directional input, waypoints, simulated captured images, and projected hazards (i.e., artificial obstructions in path). The simulation accounted for the communications latency and network delays for the real mission, projected to be 6–10 seconds per command. Therefore, the rover was operated with a waypoint driving approach as opposed to a joystick or wheel steering approach. The driver received imagery and data regarding the immediate area surrounding the rover and chose a waypoint within the camera's field of view to travel in a straight line. Each driver traversed a pre-selected path that was the same for each trial run so that the drivers would experience the same obstacles and challenges during the day compared to at night. To minimize learning effects, the traverse plan was reversed for everyone's simulation. In other words, the midnight traverse plan was the reverse of the noon traverse plan or vice versa depending on the order of the participant's simulation (two of the five drivers started with the noon traverse plan, while the other three started with the midnight traverse plan). Each traverse plan contained two Area of Interest Maps (AIM) which showed a colored square

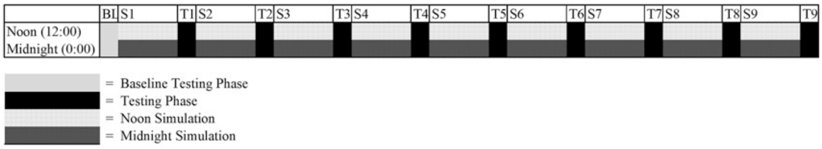

Figure 1. Experimental procedure outlining each simulation run; baseline testing (BL) was performed prior to each run and lasted roughly twice as long as each testing phase due to the extra time to complete the baseline NASA-TLX. Each simulation run (S1–S9) lasted approximately 25 minutes, followed by a testing phase (T1–T9) immediately after which lasted approximately 6 minutes. Testing sessions began every 25 and 55 minutes throughout the simulation run, totaling five hours. The drive was continuous, in that after each testing session, participants resumed where they left off in the simulation to complete the traverse plan.

around the projected world space, indicating areas that would be of interest to the real-time scientist.

The role of the real-time scientist was to observe the flow of data collected from hydrogen sensors on the rover. The real-time scientist consistently monitored a two-channel count rate, comparing water-equivalent hydrogen abundance and burial depth, which provided information about the concentration of water and how deep below it would be located. As the driver traversed the path and AIM, the real-time scientist provided continuous input on where specifically the rover should travel based on the flow of data.

Experimental Procedures

Participants completed two five-hour simulations in a randomized order, one beginning at noon and the other beginning at midnight. Four participants had their day simulation (i.e., 12:00–17:00) first and three participants had their midnight simulation (i.e., 0:00–05:00) first. Due to scheduling constraints, the time between the simulations varied from two to 10 days. Prior to the start of the simulation, the acting flight director provided a brief rundown on the procedures (e.g., purpose of simulation, traverse plan, and review of controls). During this time, the participants were asked by the study staff whether they consumed caffeine and, if so, what type and when they consumed it. Participants were then asked to complete a baseline cognitive test battery (tests described below) that was repeated approximately every 25 minutes throughout the experiment. In total, participants were scheduled to complete nine test batteries (not including the baseline assessment), until they completed five hours of continuous driving (Figure 1). During each cognitive battery, participants were moved to a nearby, isolated meeting space to minimize distractions. Participants were monitored by a study staff member for compliance during all tests.

Participants rated their sleepiness levels using the Karolinska Sleepiness Scale (KSS) (Åkerstedt & Gillberg, 1990), which is a nine-point Likert scale ranging from 1 = extremely alert to 9 = fighting sleep.

Participants completed a handheld, five-minute visual Psychomotor Vigilance Task (PVT.) to assess neurobehavioral performance (Arsintescu et al., 2019; Dinges et al., 1997; Dinges & Powell, 1985). Participants were instructed how to use the NASA-PVT prior to testing and completed a practice session on the day they received the actiwatch. The PVT has been shown to be free of learning effects during repeated measures (Basner et al., 2018).

Participants rated their workload on a validated touch-screen version of the NASA Task Load Index (NASA-TLX), which is a multidimensional scale designed to obtain estimates of workload from operators. Participants completed a baseline assessment of workload prior to both simulation sessions to create individual weighted ratings. This consisted of a 15-item comparison in which participants were asked to compare which workload characteristics were

Table 1. Demographic Information.

Variable	*M*(*SD*)	Range
Age	47.14 (12.32)	28–65
BMI	26.30 (5.52)	21.6–27
Typical weekday bedtime	23:17 (1:02)	22:00–01:00
Typical weekday waketime	06:34 (0:43)	06:00–08:00
Typical weekend bedtime	23:42 (0:59)	22:00–01:00
Typical weekend waketime	07:34 (0:56)	07:00–09:30
MEQ	53.36 (8.09)	41–64.5
ESS	6.57 (3.26)	2–10
PSQI	6.57 (4.39)	2–12

Note. MEQ = Morningness-Eveningness questionnaire; ESS = Epworth Sleepiness Scale; PSQI = Pittsburgh Sleep Quality Index. Sleep outcomes were estimated using sleep diaries.

more demanding. Each of the six workload domains (i.e., Mental, Physical, and Temporal demands, and Frustration, Effort, and Performance) were weighted against an individual's perceived ratings.

Statistical Analysis

We calculated the following metrics on the NASA-PVT: mean reaction time (RT), number of lapses (i.e., RT > 500 ms), optimum response times (i.e., the fastest 10% of response times for all trials), and cognitive slowing (i.e., the slowest 10% of response times for all trials). In addition to the NASA-PVT metrics, we calculated the weighted average for each participant's NASA-TLX ratings and compared the differences between the noon and midnight simulations through paired-sample t-tests. To assess the change in performance over time, each of the variables above were analyzed through mixed-effects modeling, with participant included as a random effect. Hedge's G was calculated as a measure of effect size for each of the tests performed. All calculations were performed using R Studio (1.3.1) on a Windows 10 Pro (20H2) operating system.

RESULTS

The participants showed evidence of sleep deficiency on average, according to the ESS and PSQI, with a wide range of scores (Table 1). They obtained more sleep in the 24 hours prior to the midnight simulation compared with the noon simulation, but less sleep in the 12 hours prior to the midnight simulation (Table 2). All but one participant consumed caffeine prior to the noon simulation (ranging from 5 to .5 h prior). For the midnight simulation, four of the seven participants consumed caffeine (ranging from 16 to 0 h prior), though only one consumed it at the start of the run while the rest consumed it many hours prior.

Participants rated themselves as significantly sleepier during the midnight simulation ($M = 5.06$, $SD = 2.28$), compared to the noon simulation ($M = 3.12$, $SD = 1.44$, $t\,(65) = -9.13$, $p < .001$, $g = .85$; Figure 2). Sleepiness significantly increased over time in the noon simulation ($F\,(9, 56) = 3.48$, $p = .0017$, $\eta_p{}^2 = .39$), but not the midnight simulation.

There were no significant differences across any of the NASA-PVT metrics between the noon and midnight simulations. Mean RT during the noon simulation ($M = 237.73$, $SD = 60.99$) was not significantly different from the midnight simulation ($M = 238.73$, $SD = 57.85$, $t\,(65) = .12$, $p = .91$; Figure 3). There were no significant differences in mean RT for the fastest 10% of responses comparing the noon simulations ($M = 170.83$, $SD = 12.23$) to the midnight simulations ($M = 175.11$, $SD = 18.41$, $t\,(65) = -1.73$, $p = .088$). There were no significant differences in mean RT for the slowest 10% of responses for the noon ($M = 363.0$, $SD = 79.91$) and midnight simulations ($M = 364.4$, $SD = 99.25$, $t\,(65) = -.19$, $p = .85$). Lastly, there were no significant

Table 2. Sleep Information.

Variable	*M*(*SD*)	Range
Sleep duration for prior week (h)	6.50 (1.75)	5.50–8.23
WASO for prior week's sleep	50.58 (40.52)	29.03–102.3
Sleep in prior 12 h (noon sim)	5.28 (.29)	4.97–5.73
Sleep in prior 24 h (noon sim)	6.02 (.83)	4.97–7.19
Sleep in prior 12 h (midnight sim)	1.19 (1.21)	.00–1.83
Sleep in prior 24 h (midnight sim)	7.44 (2.54)	4.12–11.00

Note. h = hours; WASO = Wake after sleep onset.

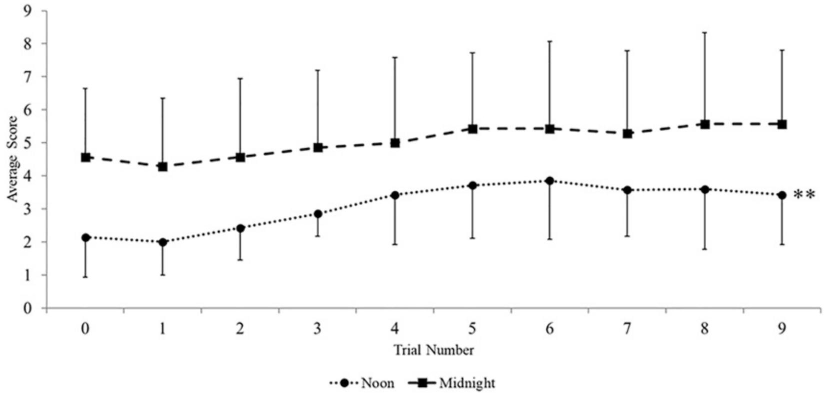

Figure 2. Mean KSS score (*y*-axis; larger number indicates greater sleepiness) plotted by trial number (*x*-axis); The noon scores are represented by the dotted line with circles ($^{**}p < .01$), with midnight represented by the hashed lines with squares. Error bars represent standard deviation. *Note*. KSS = Karolinska Sleepiness Scale.

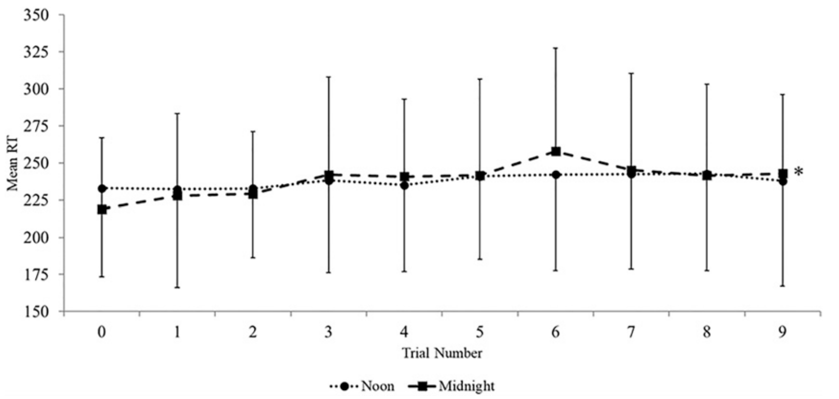

Figure 3. NASA-PVT performance (mean RT; *y*-axis) plotted by trial number (*x*-axis). Noon performance is represented by the dotted line with circles, with midnight represented by the hashed lines with squares ($^{*}p < .05$). Error bars represent standard deviation. *Note*. PVT = Psychomotor Vigilance Task

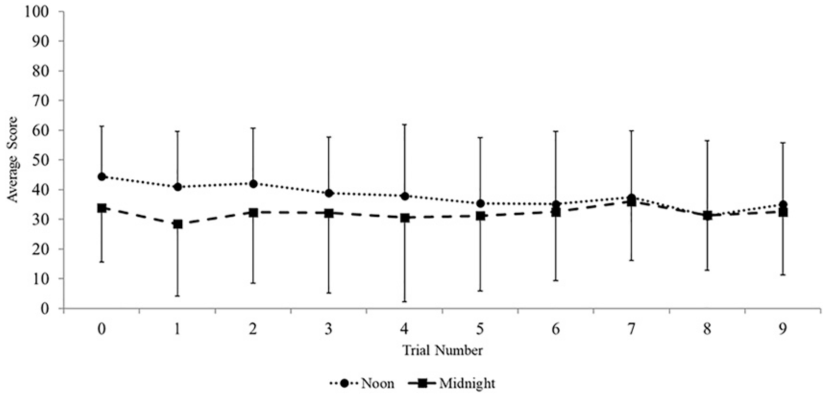

Figure 4. NASA-TLX performance (weighted workload; *y*-axis) plotted by trial number (*x*-axis). Noon performance is represented by the dotted line with circles, with midnight represented by the hashed lines with squares. Error bars represent standard deviation. *Note*. TLX = Task Load Index

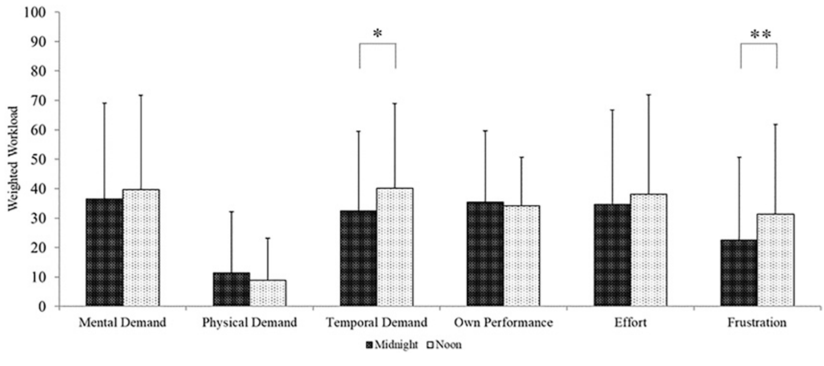

Figure 5. NASA-TLX performance (mean weighted workload; *y*-axis) plotted against individual workload domains (*x*-axis). Noon performance is represented by the lightly dotted bars, and midnight with the heavily dotted bars. Error bars represent standard deviation. (*$p < .05$, ** $p < .01$). *Note*. TLX = Task Load Index

differences across lapses in performance between the noon ($M = .35$, $SD = .54$) and midnight simulations ($M = .46$, $SD = .88$, $t\,(65) = .61$, $p = .55$).

Mean reaction time significantly increased over time during the midnight simulation ($F\,(9, 60) = 2.17$, $p = .036$, $\eta_p^2 = .27$), but not during the noon simulation. The rest of the NASA-PVT metrics (i.e., fastest 10% of responses, slowest 10% of responses, and lapses) did not yield statistically significant differences.

Participants rated workload as significantly higher during the noon simulation ($M = 37.93$, $SD = 20.09$), compared to the midnight simulation ($M = 32.09$, $SD = 21.74$, $t\,(65) = 2.81$, $p = .007$, $g = .26$; Figure 4).

Workload did not significantly differ over time during either simulation.

For individual workload dimensions, Temporal Demand was significantly higher in the noon simulation ($M = 40.15$, $SD = 28.74$) compared to the midnight simulation ($M = 32.50$, $SD = 27.04$), $t\,(65) = 2.62$, $p = .011$, $g = .26$; Figure 5). Frustration was also significantly higher in the noon simulation ($M = 31.36$, $SD = 30.44$) compared to the midnight simulation ($M =$

22.57, $SD = 28.09$, $t\ (65) = 3.34$, $p = .0014$, $g = .31$). The individual workload domains did not significantly differ over time during either simulation.

DISCUSSION

We evaluated the changes in sleepiness, performance, and workload over time during a lunar telerobotic simulation starting at noon and at midnight. We found that participants rated themselves sleepier during the midnight compared to the noon simulation. Objective performance was not different overall between the noon and midnight simulations, but reaction time slowed over time during the midnight run. Participants rated their overall workload as higher during the noon compared to the midnight simulation, but workload ratings remained stable over time for each. Our findings suggest that time-on-task and time-of-day both influence alertness, performance, and perceived workload during real-time reactive telerobotic operations.

We found that participants rated themselves as sleepier during the midnight simulations, compared to the noon simulations. This finding was expected, as individuals typically feel sleepier when they are required to stay up at night compared to during the day. This finding also aligns with previous research on subjective sleepiness among shift workers, where extended work shifts (Amann et al., 2014; Barger et al., 2014), as well as working night shifts (Garbarino et al., 2002; Härmä et al., 2002; Kazemi et al., 2016), have been shown to have similar effects on perceived sleepiness. These findings are also consistent with studies of individuals working 24-hour operations in control centers, including during pre-planned sequence operations (e.g., Mars rovers; Barger et al., 2014; Barger et al., 2012), during space transportation systems operations (Kelly et al., 1998), and during real-time reactive human operations (e.g., International Space Station Mission Control; Barger et al., 2021). Interestingly, in our study, sleepiness ratings increased over time in the noon simulation, suggesting that increasing time-on-task may unmask underlying sleepiness during daytime operations. This makes sense given that the participants averaged 6.5 hours of sleep per night during the week prior to the simulations. These findings highlight the importance of introducing mitigations to combat sleepiness during both day and night-time operations.

We did not find that participants exhibited poorer performance overall during the midnight, compared to the noon simulations. This is surprising as numerous studies have demonstrated that neurobehavioral performance is typically worse during the night compared to the daytime (Boivin & Boudreau, 2014; Short et al., 2015). This finding may result from the low number of participants that were recruited. However, consistent with prior studies in other control center operations, we did find that performance worsened over time during the midnight simulations (Barger et al., 2012, 2014, 2021). This is consistent across other operational contexts, where the PVT has been correlated with operational performance in aviation (Russo et al., 2005) and in driving (Jackson et al., 2013). Overall, these findings support the need to limit the time-on-task that an operator spends actively controlling the rover to mitigate the decline in performance that occurs with time-on-shift.

We found that participants rated their workload as significantly higher during the noon simulation compared to the midnight simulation. As the traverse path was identical (but reversed) between the noon and midnight simulations, these findings suggest that time-of-day differences affect perceived workload. Participants were not allowed to engage in other work tasks during either simulation, but it is possible that completing the simulation during the day when other work demands may have been competing for a participant's mental attention, could have amplified their perception of workload. It is also possible that shifting from one set of tasks (i.e., one's "day job") to another (i.e., the simulation) may have required increased mental effort at the beginning of the simulation. A similar phenomenon was observed among Naval crewmembers operating in a 5-day, 24-hour setting where workload was perceived as higher at the beginning of a work shift (Grech et al., 2009). Of note, the participants reported that their temporal demand and frustration were higher during the noon simulations compared to the midnight simulations, but these did not differ across time.

Although our study is one of the first to evaluate sleepiness, performance, and workload during a real-time, reactive telerobotic operation, it was not without limitation. We did not control the participants' sleep schedules, nor did we control their caffeine intake because individuals are able to sleep as they choose and freely access caffeine during real operations. Caffeine is a powerful stimulant, and it is possible that the NASA-PVT results, for example, were affected by caffeine use (O'Callaghan et al., 2018; Temple et al., 2018). Participants were also driven to and from the testing site during the night shifts as a precautionary measure to prevent fatigue-related accidents. It is possible that if participants were required to transport themselves, they would have consumed more caffeine during the evening. Future research should explore how individual strategies for preparing for work shifts impact on-the-job alertness and performance. In addition, due to the small number of participants and uneven sex distribution, our study findings may not be generalizable to all individuals completing future mission operations. We also utilized two members from the Fatigue Countermeasures Laboratory to act as confederates due to the uneven distribution of drivers to real-time scientist operators. Although their data was not utilized in the analyses, the performance of the confederates acting in this role may have affected the drivers in some way. As with all simulation studies, our study also lacked consequences for errors made during the simulation. Although the participants were instructed by both the researchers and flight director, there were no repercussions for making an error during the simulation. In addition, the lack of multiple formal pre-study practice sessions may have also affected the outcome of the results. Finally, we only studied a single, five-hour daytime shift compared to a single, five-hour nighttime shift. Real operations will require operators to work consecutive and potentially rotating shifts. Future research will need to be conducted to determine how these factors influence operator performance.

We evaluated sleepiness, performance, and workload among remote operators completing a lunar telerobotic simulation during the day and at night. In all, our findings support limiting the duration of time that a telerobotic operator works in order to preserve cognitive abilities and prevent fatigue. Our findings have practical implications for individuals working in similar real-time reactive telerobotic operation environments. Our findings also support the implementation of thoughtful schedule design that accounts for time-on-task in addition to traditional best practices shift scheduling (i.e., management of shift duration, timing, rotation, and time off). Overall, our study suggests that limiting the number of hours that an operator continuously works should mitigate against cognitive degradation. Future research should explore how rotating operators into active and passive roles of different durations during a work shift may influence alertness, performance, and workload.

KEY POINTS

- Real-time reactive telerobotic operations are understudied in terms of how the time-of-day affects performance
- We studied seven trained operators during two randomized, five-hour mission control simulations: one at noon, and one at midnight, assessing subjective sleepiness, reaction time, and perceived workload
- Sleepiness was rated higher and neurobehavioral performance degraded over time during the nighttime simulations, while perceived workload was higher during the daytime simulations
- Shift design should be considered during real-time reactive telerobotic operations to minimize sleepiness and performance decrement; future research should explore how consecutive shifts and taking breaks within a given shift might influence performance

ACKNOWLEDGMENTS

The authors wish to thank Nathan Feick and Jeremy Hirschberg for technical assistance with the study, Gregory Costedoat for review of this manuscript, and Kevin Gregory and Sean Pradhan for assistance with statistical analyses.

AUTHOR'S NOTE

Précis: We evaluated performance during a simulated teleoperation during the day and night and found that operators were sleepier at night but reported higher workload during the day.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the NASA Engineering Safety Center (NESC).

ORCID iD

Zachary Glaros https://orcid.org/0000-0001-7687-5856

REFERENCES

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*(1–2), 29–37. https://doi.org/10.3109/00207459008994241

Åkerstedt, T., & Wright, K. P. (2009). Sleep loss and fatigue in shift work and shift work disorder. *Sleep Medicine Clinics*, *4*(2), 257–271. https://doi.org/10.1016/j.jsmc.2009.03.001

Amann, U., Holmes, A., Caldwell, J., & Hilditch, C. (2014). Sleep and sleepiness of pilots operating long-range airplane emergency medical missions. *Aviation, Space, and Environmental Medicine*, *85*(9), 954–959. https://doi.org/10.3357/ASEM.3986.2014

Arsintescu, L., Kato, K. H., Cravalho, P. F., Feick, N. H., Stone, L. S., & Flynn-Evans, E. E. (2019). Validation of a touchscreen psychomotor vigilance task. *Accident Analysis & Prevention*, *126*, 173–176. https://doi.org/10.1016/j.aap.2017.11.041

Barger, L. K., Sullivan, J. P., Lockley, S. W., & Czeisler, C. A. (2021). Exposure to short wavelength-enriched white light and exercise improves alertness and performance in operational NASA flight controllers working overnight shifts. *Journal of Occupational & Environmental Medicine*, *63*(2), 111–118. https://doi.org/10.1097/JOM.0000000000002054

Barger, LK, Sullivan, JP, Vincent, AS, Fiedler, ER, McKenna, LM, Flynn-Evans, EE, Gilliland, K, Sipes, WE, Smith, PH, Brainard, GC, & Lockley, SW (2012). Learning to live on a Mars day: Fatigue countermeasures during the Phoenix Mars Lander mission. *Sleep*, *35*(10), 1423–1435. https://doi.org/10.5665/sleep.2128

Barger, L. K., Wright, K. P., Burke, T. M., Chinoy, E. D., Ronda, J. M., Lockley, S. W., & Czeisler, C. A. (2014). Sleep and cognitive function of crewmembers and mission controllers working 24-h shifts during a simulated 105-day spaceflight mission. *Acta Astronautica*, *93*, 230–242. https://doi.org/10.1016/j.actaastro.2013.07.002

Basner, M., Hermosillo, E., Nasrini, J., McGuire, S., Saxena, S., Moore, T. M., Gur, R. C., & Dinges, D. F. (2018). Repeated administration effects on psychomotor vigilance test performance. *Sleep*, *41*(1), 1-6. https://doi.org/10.1093/sleep/zsx187.

Boivin, D. B., & Boudreau, P. (2014). Impacts of shift work on sleep and circadian rhythms. *Pathologie Biologie*, *62*(5), 292–301. https://doi.org/10.1016/j.patbio.2014.08.001

Buysse, D. J., Reynolds, C. F., III, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Chen, J. Y. C., Haas, E. C., & Barnes, M. J. (2007). Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, *37*(6), 1231–1245. https://doi.org/10.1109/tsmcc.2007.905819

Dinges, DF, Pack, F, Williams, K, Gillen, KA, Powell, JW, Ott, GE, Aptowicz, C, & Pack, AI (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night. *Sleep*, *20*(4), 267–277

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*(6), 652–655. https://doi.org/10.3758/BF03200977

Duffy, J. F., Rimmer, D. W., & Czeisler, C. A. (2001). Association of intrinsic circadian period with morningness-eveningness, usua2l wake time, and circadian phase. *Behavioral Neuroscience*, *115*(4), 895–899. https://doi.org/10.1037//0735-7044.115.4.895

Garbarino, S., Nobili, L., Beelke, M., Balestra, V., Cordelli, A., & Ferrillo, F. (2002). Sleep disorders and daytime sleepiness in state police shiftworkers. *Archives of Environmental Health: An International Journal*, *57*(2), 167–173. https://doi.org/10.1080/00039890209602932

Grech, M. R., Neal, A., Yeo, G., Humphreys, M., & Smith, S. (2009). An examination of the relationship between workload and fatigue within and across consecutive days of work: Is the relationship static or dynamic? *Journal of Occupational Health Psychology*, *14*(3), 231–242. https://doi.org/10.1037/a0014952

Härmä, M, Sallinen, M, Ranta, R, Mutanen, P, & Müller, K (2002). The effect of an irregular shift system on sleepiness at work in train drivers and railway traffic controllers. *Journal of Sleep Research*, *11*(2), 141–151. https://doi.org/10.1046/j.1365-2869.2002.00294.x

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

Jackson, M. L., Croft, R. J., Kennedy, G. A., Owens, K., & Howard, M. E. (2013). Cognitive components of simulated driving performance: Sleep loss effects and predictors. *Accident Analysis & Prevention*, *50*, 438-444. https://doi.org/10.1016/j.aap.2012.05.020.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: The epworth sleepiness scale. *Sleep*, *14*(6), 540–545. https://doi.org/10.1093/sleep/14.6.540

Kazemi, R, Haidarimoghadam, R, Motamedzadeh, M, Golmohamadi, R, Soltanian, A, & Zoghipaydar, MR (2016). Effects of shift work on cognitive performance, sleep quality, and sleepiness among petrochemical control room operators. *Journal of Circadian Rhythms*, *14*(1), 1–8. https://doi.org/10.5334/jcr.134

Kelly, SM, Rosekind, MR, Dinges, DF, Miller, DL, Gillen, KA, Gregory, KB, Aguilar, RD, & Smith, RM (1998). Flight controller alertness and performance during spaceflight shiftwork operations. *Human Performance in Extreme Environments : The Journal of the Society for Human Performance in Extreme Environments*, *3*(1), 100–106.

O'Callaghan, F., Muurlink, O., & Reid, N. (2018). Effects of caffeine on sleep quality and daytime functioning. *Risk Management and Healthcare Policy*, *11*, 263. https://doi.org/10.2147/RMHP.S156404.

Russo, MB, Kendall, AP, Johnson, DE, Sing, HC, Thorne, DR, Escolas, SM, Santiago, S, Holland, DA, Hall, SW, & Redmond, DP (2005). Visual perception, psychomotor performance, and complex motor performance during an overnight air refueling simulated flight. *Aviation, Space, and Environmental Medicine*, *76*(7), C92–C103.

Schipani, S. P. (2003). *An evaluation of operator workload, during partially-autonomous vehicle operations*. Army Research Lab Aberdeen Proving Ground MD.

Short, M. A., Agostini, A., Lushington, K., & Dorrian, J. (2015). A systematic review of the sleep, sleepiness, and performance implications of limited wake shift work schedules, *Scandinavian Journal of Work, Environment & Health*, *41*(5), 425–440. https://doi.org/10.5271/sjweh.3509

Temple, J. L., Hostler, D., Martin-Gill, C., Moore, C. G., Weiss, P. M., Sequeira, D. J., Condle, J. P., Lang, E. S., Higgins, J. S., & Patterson, P. D. (2018). Systematic review and meta-analysis of the effects of caffeine in fatigued shift workers: Implications for emergency medical services personnel. *Prehospital Emergency Care*, *22*(supp1), 37-46.

Zachary Glaros obtained his MA in research and experimental psychology at San Jose State University in 2020 and is now a civil servant at NASA Ames Research Center through the Pathways Intern Program. Zachary is currently working in the Fatigue Countermeasures Laboratory. His work involves evaluating sleep and sleep loss in NASA missions.

Robert Carvalho is a Systems Engineer and Mission Operator at NASA Ames Research Center. Robert has a BS in Computer Engineering (1995). Robert has designed and executed operations for multiple missions, including IRIS and STP Sat-5. Robert is currently designing operations for the VIPER mission and looks forward to operating on the Moon.

Erin E. Flynn-Evans obtained her PhD from the University of Surrey in 2010 and MPH from Harvard School of Public Health in 2014. She is currently the Director and Principal Investigator of the Fatigue Countermeasures Laboratory at NASA Ames Research Center.