Received: 2 August 2021 | Revised: 15 October 2021 | Accepted: 5 November 2021

DOI: 10.1111/jsr.13521

Journal of Sleep Research

RESEARCH ARTICLE

Early starts and late finishes both reduce alertness and performance among short-haul airline pilots

Lucia Arsintescu[1] | **Sean Pradhan**[1,2] | **Ravi G. Chachad**[1] | **Kevin B. Gregory**[3] | **Jeffrey B. Mulligan**[4] | **Erin E. Flynn-Evans**[3]

[1]Fatigue Countermeasures Laboratory, San Jose State University, Moffett Field, California, USA

[2]School of Business, Menlo College, Atherton, California, USA

[3]Fatigue Countermeasures Laboratory, NASA Ames Research Center, Moffett Field, California, USA

[4]Ocular Imaging Laboratory, NASA Ames Research Center, Moffett Field, California, USA

Correspondence
Lucia Arsintescu, NASA Ames Research Center, Fatigue Countermeasures Laboratory, MS: 262-4, CA 94035, USA.
Email: lucia.arsintescu-1@nasa.gov

Summary

Flight crews are frequently required to work irregular schedules and, as a result, can experience sleep deficiency and fatigue. This study was conducted to determine whether perceived fatigue levels and objective performance varied by time of day, time awake, and prior night's sleep duration. Ninety-five pilots (86 male, 9 female) aged 33 years (±8) volunteered for the study. Participants completed a daily sleep diary, Samn-Perelli fatigue scale, and psychomotor vigilance task that were completed before and after each flight duty period and at the top-of-descent for each flight. Pilots experienced higher self-reported fatigue (*EMM* = 3.92, SE = 0.09, $p < 0.001$) and worse performance (Response speed: *EMM* = 4.27, SE = 0.08, $p = 0.004$) for late-finishing duties compared with early-starting duties (Samn-Perelli: *EMM* = 3.74, SE = 0.08; Response speed: *EMM* = 4.37, SE = 0.08), but had shorter sleep before early-starting duties (early: *EMM* = 6.94, SE = 0.10; late: *EMM* = 8.47, SE = 0.14, $p < 0.001$). However, pre-duty Samn-Perelli and response speed were worse ($z = 4.18$, $p < 0.001$; $z = 3.05$, $p = 0.03$; respectively) for early starts compared with late finishes (*EMM* = 2.74, SE = 0.19), while post-duty Samn-Perelli was worse for late finishes (*EMM* = 4.74, SE = 0.19) compared with early starts (*EMM* = 4.05, SE = 0.12). The results confirm that duty time has a strong influence on self-reported fatigue and performance. Thus, all flights that encroach on a biological night are targets for fatigue risk management oversight.

KEYWORD
aviation, crew, duty limits, fatigue, short-haul,

1 | INTRODUCTION

Fatigue remains a challenge in short-haul aviation due to (1) irregular schedules that prevent adequate duration and timing of sleep (Borgeois-Bougrine et al., 2003); (2) early and late duty times that encroach on the nightly sleep opportunity (Åkerstedt et al., 2021; Flynn-Evans et al., 2018; Roach et al., 2012; Vejvoda et al., 2014); and (3) workload factors, such as duty length and the number of sectors (Flynn-Evans et al., 2018; Goffeng et al., 2019; Honn et al., 2016; Powell et al., 2007). The International Civil Aviation Organization (ICAO) defines fatigue as "a physiological state of reduced mental or physical performance capability resulting from sleep loss or extended wakefulness, circadian phase, or workload (mental and/or physical activity) that can impair a crew member's alertness and ability to safely operate an aircraft or perform safety-related duties" (International Air Transport Association (IATA;2015). To minimise these risks, industry regulators prescribe limits for duty hours that airlines should follow when generating pilot work schedules. The European Aviation Safety Agency (EASA) and the Federal Aviation Administration (FAA) impose many limits on pilot duty time, including restricting the duration of work by the time of day that a duty period starts (Commission Regulation (EU) No 83/2014; FAA, 14 CFR Parts 117, 119, and 121, 2009). Both regulations allow for longer work episodes when work shifts begin earlier in the day relative to shifts that begin later. It is unclear, however, whether such duty time restrictions align with actual fatigue and performance.

In 2014, Vejvoda and colleagues conducted a study to examine the effects of time awake on late-finishing flights (defined by EASA as flights "finishing in the period between 00:00 and 01:59 in the time zone to which the crew is acclimatized", EASA, EU No83/, 2014). Pilots rated their fatigue levels at the end of each flight and duty period. They found that for late-finishing flight duty periods (FDPs), pilots were awake longer and reported moderate to severe fatigue levels at duty end compared with at the end of FDPs that started early in the morning (between 05:00 and 06:59). However, this study only included self-report measures of fatigue and may not have captured the influence of time of day on pilot performance for the early starts because their measures were collected at the end of the flight, which would occur during a time of circadian alertness for flights starting in the early morning (i.e., the measures assessing early starts were collected in the late morning or early afternoon; Bermudez et al., 2016). As early start flights often begin before typical wake times (e.g., 05:00–07:00), it would be expected that the influence of the circadian system might cause pilots to feel sleepy and perform poorly at the beginning of their shift . In addition, early start times reduce the amount of sleep that pilots obtain the night before a flight (Åkerstedt et al., 2021; Flynn-Evans et al., 2018; Roach et al., 2012). For instance, Roach et al. (2012) found that for every hour of duty start before 9:00 am, pilots lose approximately 15 min of sleep.

Chronic sleep loss causes a decline in psychomotor vigilance, working memory, and cognitive throughput performance (Van Dongen et al., 2003). Even relatively moderate sleep restriction—if sustained night after night—can seriously impair waking neurobehavioural functions in healthy adults. Thus, the objective of this study was to examine the effects of time of day on fatigue and pilot performance among short-haul airline pilots. Furthermore, we expanded on the investigation of the effects of time awake on pilot fatigue as described in Vejvoda et al. (2014) by using a large data set and investigating performance as measured by the psychomotor vigilance task (PVT) in addition to self-report ratings.

2 | METHODS

Ninety-five pilots (86 male, 9 female) from a short-haul airline volunteered to participate in the study during normal airline operations over 36 consecutive days (including days off). One pilot was removed from analyses due to incomplete data.

The study was reviewed and approved by the NASA Ames Research Center (ARC) Institutional Review Board (IRB)(HRI-312; HRI-319). Participation in the study was voluntary. Pilots were contacted via their company email address and through flyers. Volunteers were invited for a training session where they provided informed consent and were trained on the study procedures. Each participant was provided with an iPod (5th generation, iOS 6.8.53, Apple, Cupertino, CA) for completing the tests during the study period. Using a custom-built mobile application that was administered on an iPod, they filled in a demographic questionnaire, a Morningness-Eveningness Questionnaire (MEQ; Horne & Östberg, 1976), and the Epworth Sleepiness Scale (ESS; Johns, 1991). In addition, they completed a practice data collection session under the supervision of a study researcher to ensure that they understood the study procedures and tests. All pilots flew a pre-designed schedule of four duty blocks separated by 3 days off and an attempt was made to have all pilots fly the same type of schedule. Each duty block contained five duty days and each duty day had either two or four flight sectors. The same type of duty (early, midday [morning-afternoon], late) was scheduled for all five days in a duty block, but operational constraints yielded some modest variations from this schedule. All pilots were scheduled for blocks of all types of flights. Each duty block was separated by 3–4 days off. All pilots flew during the day and their duty start time varied from early morning (e.g., ~05:00) to late afternoon, while end duty time varied from late morning to late night (e.g., ending ~00:00). Each pilot returned to their home base at the end of the FDP on each day of the study. The duty schedules were obtained from the airline at the end of the study to confirm the actual time of the flights. In order to maintain consistency with the analyses conducted by Vejvoda et al. (2014), we classified early duties as starting between 05:00 and 6:59, mid-morning duties as starting between 07:00 and 10:59, afternoon duties as starting between 13:00 and 16:59, and evening duties as starting between 17:00 and 20:59. Furthermore, as with Vejvoda et al. (2014), we compared early starts beginning from 05:00 and 06:59 to late finishes ending between 00:00 and 01:59.

Participants completed a sleep diary twice per day (upon bedtime and wakeup) throughout the study where they entered information about bedtime, wakeup time, and sleep quality and rated their fatigue level using a Samn-Perelli (SP) fatigue scale (Samn & Perelli, 1982). The SP fatigue scale is a 7-point Likert type scale ranging from "1 = fully alert, wide awake" to "7 = completely exhausted, unable to function effectively". On duty days, participants also completed a SP fatigue scale and a 5-minute PVT (Arsintescu et al., 2017) on the top-of-descent (TOD) of each flight, and at the end of the duty period. In addition, 52 pilots completed the SP and PVT before duty (in the briefing room).

2.1 | Data analysis

All statistical analyses were performed using RStudio (Version 1.3.1056) and IBM SPSS Statistics (Version 25). In cases where tests were taken within 30 min of each other, we excluded the second SP or PVT. For all sleep analyses, we evaluated the prior night's sleep in relation to duty day. In order to estimate sleep on the night prior to duty we excluded the naps from main sleep analyses. Naps were taken 0.08% of times across all duty days and were usually taken after early starts. Touch events from the iPod included a latency of 68.53 ms relative to traditional PVT boxes (Arsintescu et al., 2017). This was subtracted from each PVT raw trial before PVT analyses to accurately assess lapses (i.e., reaction times > 500 ms) and response times. The following metrics were assessed in our PVT analyses: (1) *response speed*–the reciprocal response time (mean (1/RT) × 1000), and (2) *lapses*–the reaction times exceeding 500 ms. A PVT response was considered valid if the reaction time (RT) was >100 ms. Responses with an RT ≤ 100 ms were considered false starts and were removed from analyses.

Prior to data analyses, we removed outliers from our dataset using a cutoff criterion of three standard deviations above or below the mean for each of our measures (i.e., response speed, lapses, and SP fatigue scores). The relationships between subjective fatigue reported via the SP fatigue scale and outcomes of the PVT (i.e., response speed and lapses) were analysed using repeated-measures correlations. To investigate the impact of the duty start time on pilots' subjective fatigue and PVT performance, a series of mixed-effects models (using data collected during the entire duty period) were evaluated. The rmcorr (Bakdash & Marusich, 2017), lme4 (Bates et al., 2019), lmerTest (Kuznetsova, Brockhoff, & Christensen, 2019), and multcomp packages (Hothorn et al., 2019) for R were utilised for these tests. A linear model was assumed for response speed, while a negative binomial distribution was specified for lapses due to overdispersion. To examine the heightened risk of subjective fatigue, scores on the SP were dichotomised using a cut-off criterion ≥5 (IATA, ICAO, & IFALPA, 2015; Samn & Perelli, 1982). Further analysis on this binary indicator of subjective fatigue was performed to compare early-starts and late-finishes using a mixed-effects logistic regression controlling for prior sleep. Additional mixed-effects models were also conducted on SP fatigue, PVT response speed, and lapses to evaluate performance change as a function of time of day. All *p*-values for group comparisons were adjusted using the Bonferroni correction.

3 | RESULTS

3.1 | Participants

The pilots were 33 (±8) years old (range 21–54) and reported 7.76 (±0.75) h of sleep need per night to feel fully alert. Sixty-five percent of the pilots had <4000 h of total flight experience, 23% had between 4000–9500 h, and 12% had more than 10,000 h of total flight experience. Demographic characteristics for the 94 pilots who provided complete data sets are shown in Table 1.

TABLE 1 Demographic characteristics of participating pilots (n = 94)

	M (*SD*)	Range
Age	32.9 (8.03)	21–54
Weight (kg)	77.23 (11.61)	45.5–100.5
Height (m)	1.79 (0.07)	1.60–1.93
BMI (kg/m²)	23.99 (2.93)	16.20–32.70
Sleep need (h)	7.76 (0.75)	5–10
MEQ score	51.5 (5.91)	36–64
ESS score	5.37 (3.72)	0–19

Abbreviation: BMI, body mass index; ESS, Epworth sleepiness scale; M, mean; MEQ, Morningness-eveningness questionnaire; SD, standard deviation.

TABLE 2 Sleep characteristics prior to duty by duty start time

Start of duty	Waketime (hh:mm)	Bedtime (hh:mm)	Sleep duration (h)
05:00–06:59 (early)	04:19 (01:14)	21:15 (02:01)	6.90 (1.30)
07:00–10:59 (mid-morning)	06:30 (01:22)	22:29 (01:27)	7.54 (1.57)
13:00–16:59 (afternoon)	09:16 (01:19)	01:05 (01:21)	8.39 (1.47)
17:00–20:59 (evening)	09:27 (01:37)	02:09 (01:31)	8.26 (1.73)

Sleep characteristics are reported as mean (standard deviation).
Abbreviation: hh:mm, hours and minutes.

Detailed information about sleep, including bedtimes, wake times, and sleep duration prior to duty by duty start times, is provided in Table 2. Participants reported obtaining less sleep than their average sleep need when duty started early (Table 2).

3.2 | Flight duty periods

Data were collected during 1476 FDPs ($M = 7.40 \pm 1.91$ h) that comprised 2738 flights among study participants. FDP information by duty start time is provided in Table 3.

3.3 | Effects of duty start time on SP fatigue and PVT performance

The repeated-measures correlations demonstrated that the SP fatigue scores were negatively and weakly correlated with response speed ($r(4470) = -0.26$, $p < 0.001$, 95% CI: [−0.29, −0.23]), and positively and weakly correlated with lapses ($r(4470) = 0.14$, $p < 0.001$, 95% CI: [0.11, 0.17]). The results from the linear mixed-effects models revealed main effects of duty start time on SP scores ($F[3, 2987.50] = 28.38$, $p < 0.001$, $R^2_{\text{Marginal}} = 0.02$, $R^2_{\text{Conditional}} = 0.29$), response speed, ($F$[3,

TABLE 3 Flight duty period information by duty start time

Start of duty	Flights	Sectors	Flight duration (h)	Flight duty period duration (h)
05:00–06:59	848	2.19 (0.58)	2.18 (0.70)	6.60 (1.67)
07:00–10:59	296	2.17 (0.56)	2.51 (0.66)	7.30 (1.63)
13:00–16:59	834	2.41 (0.82)	2.48 (1.01)	8.02 (1.99)
17:00–20:59	117	1.99 (0.12)	2.35 (0.50)	6.42 (1.22)

Flights are provided as frequencies. Sectors and durations are reported as mean (standard deviation).

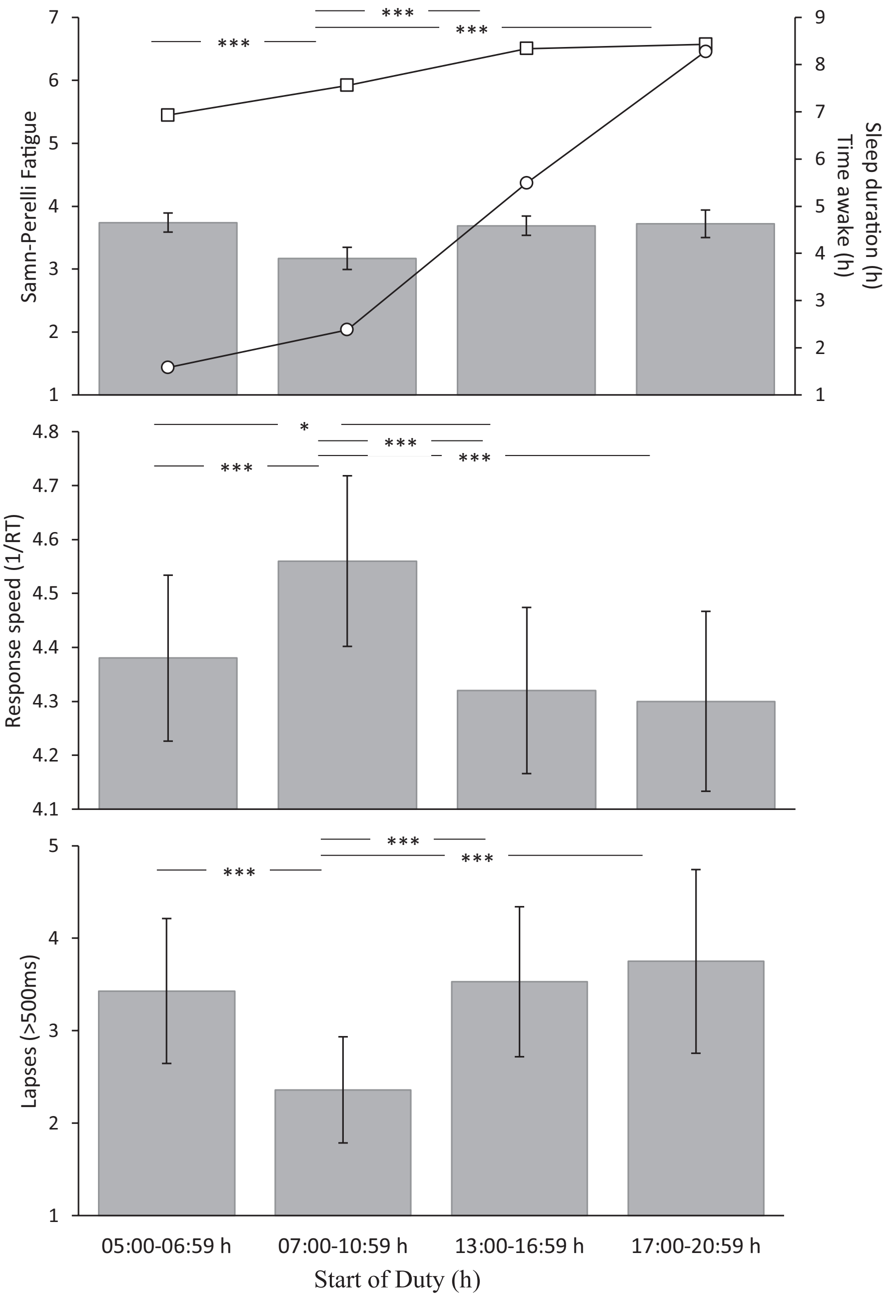

FIGURE 1 Estimated marginal mean Samn-Perelli fatigue across duty (grey bars + 95% CI; top panel), response speed (grey bars + 95% CI; middle panel), and lapses (grey bars + 95% CI; bottom panel) plotted as a function of duty start time (05:00–06:59 h [early morning]; 07:00–10:59 h [mid-morning]; 13:00–16:59 h [afternoon]; 17:00–20:59 h [evening]). Estimated marginal means are reported to adjust for other model terms. Top panel secondary vertical axis (right): Sleep period time (sleep duration) in the previous night = open squares; Time awake at duty start = open circles; ms, milliseconds; h, hours; *$p < 0.05$, ***$p < 0.001$

2952.10] = 26.61, $p < 0.001$, $R^2_{\text{Marginal}} = 0.01$, $R^2_{\text{Conditional}} = 0.74$), and lapses, ($\chi^2[3] = 17.93$, $p < 0.001$, $R^2_{\text{Marginal}} = 0.01$, $R^2_{\text{Conditional}} = 0.57$). Pilots reported significantly higher fatigue on the SP (Figure 1, top panel) when their FDP started in the early-morning, afternoon, and evening compared with mid-morning ($p < 0.001$). These results were nearly identical for response speed and lapses (Figure 1, middle panel and bottom panel). The response speed decreased when the FDP started in the early-morning, afternoon, and evening compared with mid-morning ($p < 0.001$). Pilots were also significantly slower during FDPs that began in the afternoon as opposed to the early morning ($p = 0.03$). Pilots had significantly more lapses (Figure 1, bottom panel) in the early morning, afternoon, and evening compared with mid-morning starts ($p < 0.001$). In addition, total time awake at duty start was related to decreased response speed ($F[1, 2495.50] = 8.08$, $p = 0.005$, $R^2_{\text{Marginal}} = 0.001$, $R^2_{\text{Conditional}} = 0.73$), while the prior main sleep duration was not associated with statistically clear changes in response speed ($p > 0.05$). Ultimately, prior sleep duration and total time awake at duty start did not meaningfully affect subjective fatigue or lapses across the duty periods ($p > 0.05$).

3.4 | Early start vs. late finish duties effects on SP fatigue and PVT performance

In order to replicate Vejvoda et al. (2014), we conducted a series of comparisons between early-starting (05:00–06:59 clock h) and late-finishing (00:00–01:59 clock h) FDPs to examine differences on SP subjective fatigue. To extend the findings of Vejvoda et al., we conducted the same analyses to assess changes in PVT performance (Figure 2). The results from these analyses indicated that pilots reported that their fatigue was significantly higher ($z = 2.90$, $p = 0.004$) for late-finishing duties ($EMM = 3.92$, SE = 0.09) compared with early starts ($EMM = 3.74$,

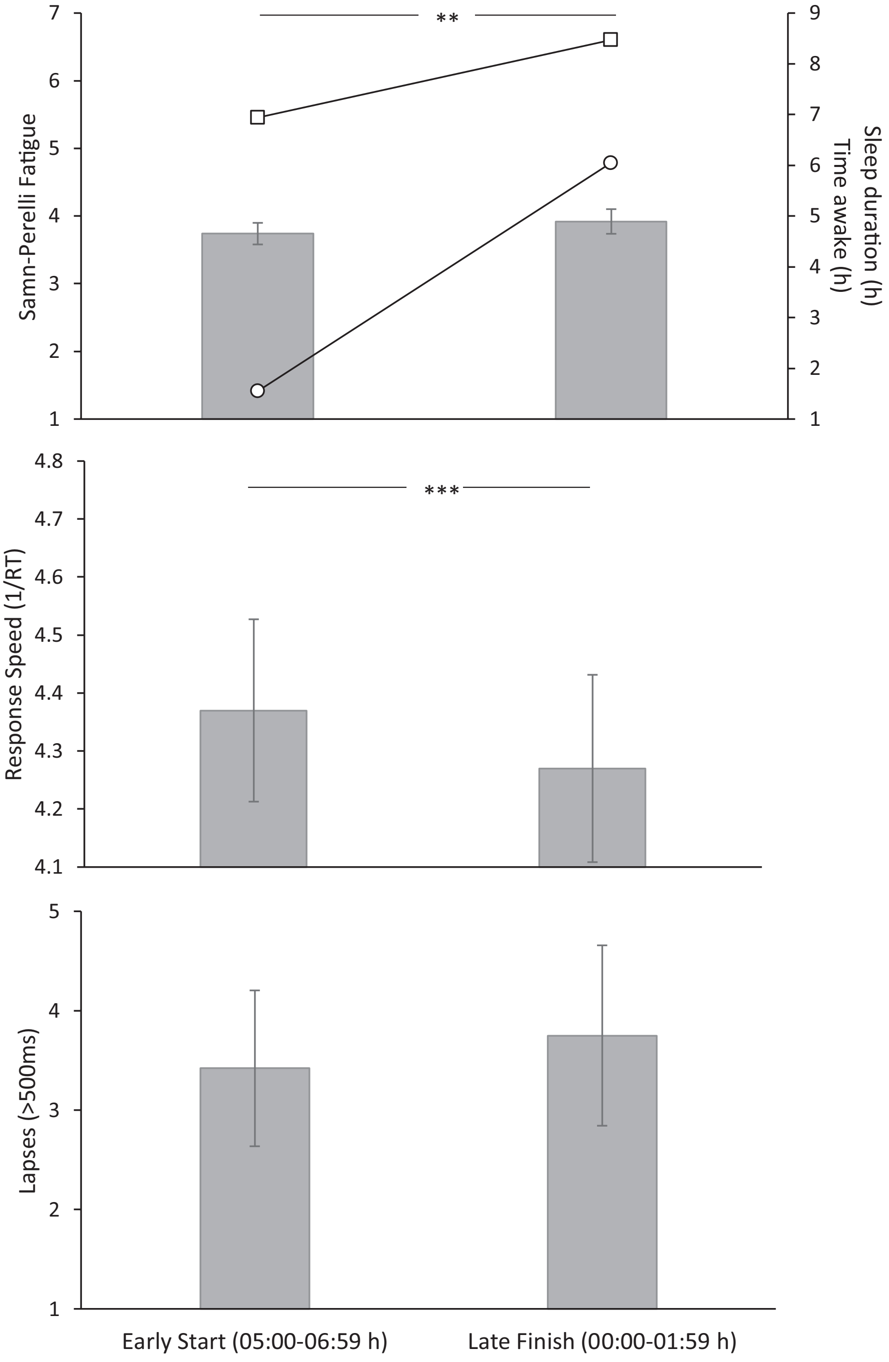

FIGURE 2 Estimated marginal mean Samn-Perelli fatigue across duty (grey bars + 95% CI; top panel), Response speed (grey bars + 95% CI; middle panel), and lapses (grey bars + 95% CI; bottom panel) for early-start and late-finishing FDPs. Top panel secondary vertical axis (right): Sleep period time (sleep duration) in the previous night = open squares; Time awake at duty start = open circles; ms, milliseconds; h, hours; **$p < 0.01$, ***$p < 0.001$

SE = 0.08; Figure 2, top panel). The response speed was significantly worse ($z = -3.64$, $p < 0.001$) for late-finishing duties (*EMM* = 4.27, SE = 0.08) compared with early-starting FDPs (*EMM* = 4.37, SE = 0.08; Figure 2, middle panel). There were no significant differences in the number of lapses between the two duty types ($p > 0.05$; Figure 2, bottom panel). For late finishes (*EMM* = 8.47, SE = 0.14), prior sleep periods were significantly longer ($z = 11.21$, $p < 0.001$) compared with early-starting duties (*EMM* = 6.94, SE = 0.10). In addition, the FDP duration for late finishes (*EMM* = 9.78, SE = 0.14) was significantly longer ($z = 17.35$, $p < 0.001$) than early starts (*EMM* = 7.12, SE = 0.08). However, pilots completing late-finishing duties (*EMM* = 6.05, SE = 0.11) were awake for significantly longer periods of time prior to their FDP start time ($z = 39.85$, $p < 0.001$) compared with those with early starts (*EMM* = 1.55, SE = 0.07).

The mixed-effects logistic regression showed a significant effect of duty type on high subjective fatigue, $\chi^2(1) = 16.50$, $p < 0.001$, $R^2_{\text{Marginal}} = 0.02$, $R^2_{\text{Conditional}} = 0.20$. Specifically, pilots with late-finishing duties ($\hat{p} = 0.42$, $SE = 0.07$) were more likely ($z = 2.32$, $p = 0.02$; $OR = 02.11$, $SE = 00.15$) to report high subjective fatigue than those with early-starting duties ($\hat{p} = 0.26$, $SE = 0.03$).

Given that pre-duty data were only collected among 52 participants, specificity analyses using identical mixed-effects models with Bonferroni-corrected post-hoc tests on SP scores, response speed, and lapses were performed. We sought to compare differences among pre-duty, in-flight, and post-duty responses for early-starting and late-finishing duties. Results from these analyses indicated that pre-duty scores on the SP were higher ($z = 4.18$, $p < 0.001$; Figure 3, top panel) for early starts (*EMM* = 3.45, SE = 0.12) than for late finishes (*EMM* = 2.74, SE = 0.19). However, post-duty SP scores were higher ($z = 4.02$, $p < 0.001$) for late finishes (*EMM* = 4.74, SE = 0.19) compared with early starts (*EMM* = 4.05, SE = 0.12). Pre-duty response speed was lower for pilots ($z = 3.05$, $p = 0.03$; Figure 3, middle panel) during early-starting duties (*EMM* = 4.19, SE = 0.11) than late-finishing ones (*EMM* = 4.42, SE = 0.13). There were no meaningful differences among pre-duty, in-flight, and post-duty lapses for early starts and late finishes ($p > 0.05$; Figure 3, bottom panel).

3.5 | Effects of time of day on SP fatigue and PVT performance

Additional mixed-effects models showed that SP fatigue scores followed a significant quadratic pattern by time of day, with pilots experiencing less fatigue during daytime hours and more fatigue during night hours (i.e., early morning or late night; $F[2, 4177] = 188.11$, $p < 0.001$, $R^2_{\text{Marginal}} = 0.06$, $R^2_{\text{Conditional}} = 0.25$; Figure 4, top panel). PVT response speed also followed a significant quadratic trend, with better performance occurring during the day and poorer performance occurring during early morning and night hours ($F[2, 4187] = 15.67$, $p < 0.001$, $R^2_{\text{Marginal}} = 0.03$, $R^2_{\text{Conditional}} = 0.56$; Figure 4, middle panel). Lapses followed a similar pattern and were elevated during the early morning, decreased during the day, and increased again during the night ($F[2, 4187] = 6.55$, $p = 0.001$, $R^2_{\text{Marginal}} = 0.02$, $R^2_{\text{Conditional}} = 0.40$; Figure 4, bottom panel).

4 | DISCUSSION

We found that fatigue ratings and performance varied by time of day among short-haul pilots flying during the day. We also detected similar outcomes when examining the SP fatigue ratings and PVT performance (response speed and lapses) by the start time of duty. Pilots were more fatigued and showed slower PVT response speed and an increased number of lapses when the duty started in the early-morning, afternoon, and evening compared with mid-morning start times. Although sleep duration was significantly shorter before early starts, prior night's sleep duration did not affect any of the outcome variables. Collectively, our findings suggest that encroachment on the biological night is associated with worse fatigue and performance. It is likely that sleep deficiency, combined with time awake and time of day interact, which would explain why the worst fatigue and performance occurred at the end of late finishes (Cohen et al., 2010).

Our findings confirm that fatigue follows a time of day pattern among short-haul pilots (Powell et al., 2007), and extends prior research (Vejvoda et al., 2014) by providing objective measures of performance from the PVT. A large study by Powell et al. showed that SP fatigue ratings followed a similar time of day pattern, with higher fatigue ratings in the early morning, the lowest fatigue ratings in the late morning, and the highest fatigue ratings for duties ending around midnight (Powell et al., 2007). We found that fatigue ratings followed a similar pattern and demonstrate that objective performance also varies in the same manner. Notably, the performance impairment that we observed was not extreme in the morning or after late finishes (i.e., <4 lapses), likely because the pilots had days off that allowed for recovery sleep and access to countermeasures such as caffeine. However, the relatively worse performance that we observed likely reflects increasing "state instability," suggesting that early starts and late finishes reduce pilots' capacity to sustain attention (Doran, Van Dongen, & Dinges, 2001). Consistent with previous research (Powell et al., 2007; Sallinen et al., 2021; Vejvoda et al., 2014), we found that fatigue and performance were worse at the end of the late finishes compared with at the end of the early starts. Importantly though, while the Vejvoda study only included data from the end of duty, we collected pre-duty data from a subset of our participants. In these analyses, we found that pre-duty fatigue and performance were worse during early starts compared with late finishes. Our findings are also consistent with Sallinen et al., who found that pilots experienced reduced alertness on both early and late FDPs with an increase level of fatigue across flights on late FDPs (Sallinen et al., 2017). Collectively, these findings suggest that both early starts and late finishes should be considered targets for fatigue risk management, but that these duties each require unique fatigue management strategies.

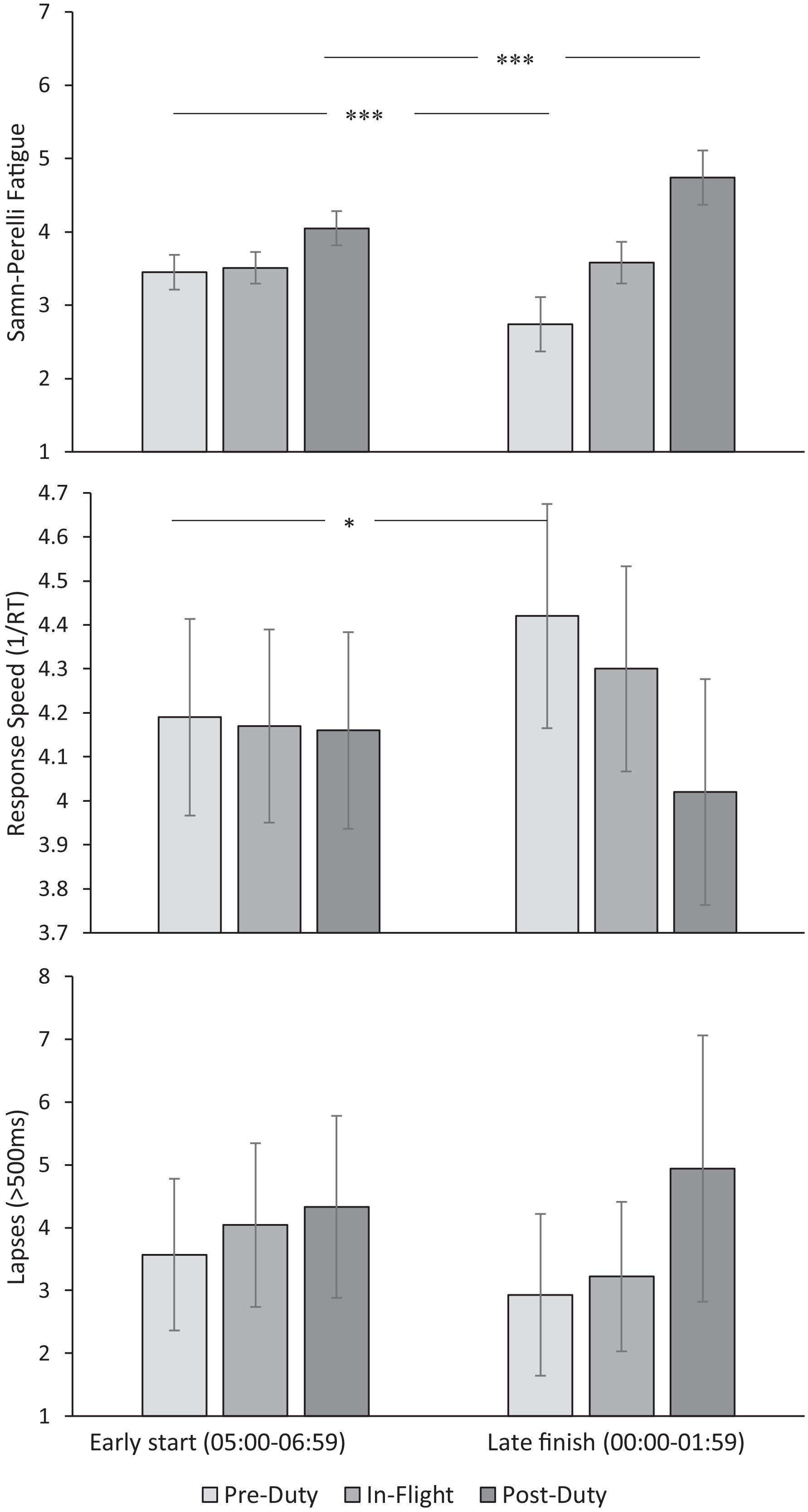

FIGURE 3 Estimated marginal mean Samn-Perelli fatigue by pre-duty, in-flight, and post-duty (95% CI; top panel), Response speed (95% CI; middle panel), and lapses (95% CI; bottom panel) for early-start and late-finishing FDPs. ms, milliseconds; $*p < 0.05$, $***p < 0.001$

Others have also demonstrated that fatigue ratings are higher during early starts compared with later morning starts (Åkerstedt et al., 2021; Bourgeois-Bougrine et al., 2003; Flynn-Evans et al., 2018; Roach et al., 2012; Sallinen et al., 2017). This likely relates to later starts affording a longer sleep opportunity and to the circadian drive for alertness promoting waking during the biological day (Roach et al., 2012). That is, when the pilots completed preflight assessments of fatigue and sleepiness prior to early starts, they were

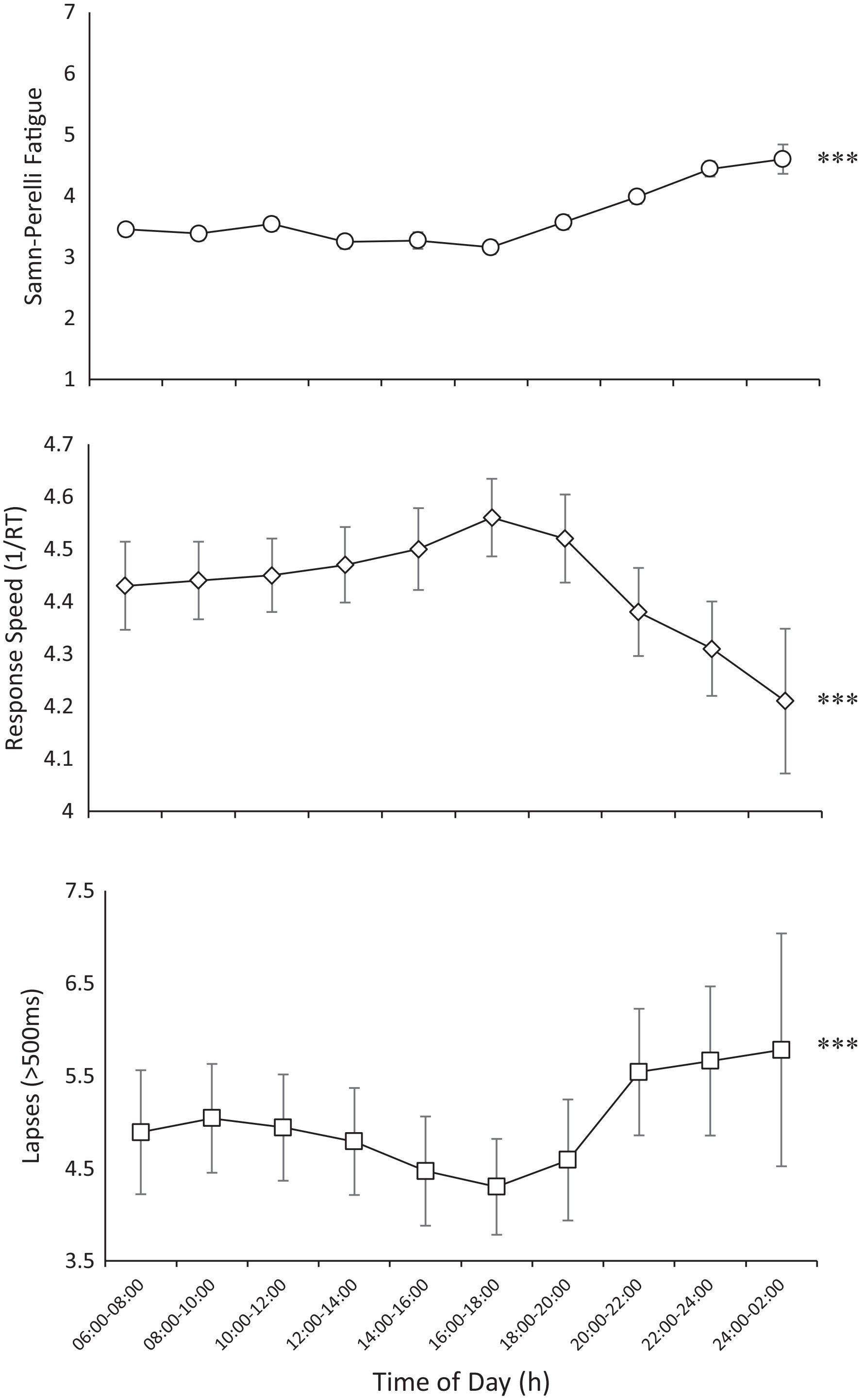

FIGURE 4 Mean Samn-Perelli by time of day (open circles + 95% CI; top panel), Response speed (open diamonds + 95% CI; middle panel), and lapses (open squares + 95% CI; bottom panel). ms, milliseconds; h, hours; ***$p < 0.001$

awake at a time when they would typically be asleep. In contrast, later morning starts would have involved providing fatigue assessments at times that were likely during habitual wake periods. Åkerstedt et al. (2021) found that sleep was the strongest predictor of fatigue over a 24-period followed by duty type (including both early and late duties) and duty time. Although we did not find a statistically significant association between prior night's sleep duration and subsequent fatigue or performance, the amount of sleep that the pilots reported prior to early starts was 6.90 h, which is considered insufficient relative to consensus recommendations (Hirshkowitz et al., 2015; Watson et al., 2015). Given that sleep information in our study was assessed via a sleep diary, it is possible that the participants misestimated the sleep they obtained. It is also possible that the pilots' use of caffeine and other countermeasures dampened our ability to observe an effect of prior nights' sleep duration. Prior studies have demonstrated that sleep has an immediate restorative effect on performance, but sleep-deprived individuals experience faster deterioration of performance relative to those who are rested (Cohen et al., 2010). This may explain why participants in our study experienced worsening performance over time awake on early starts despite the onset of the circadian alerting signal. These findings suggest that strategies to minimise sleep loss and to increase pilot sleep duration prior to early starts should be evaluated.

Duty time limitations vary by time of day in both Europe (EU No83/, 2014) and the US (Federal Aviation Administration & Department of Transportation, 2009), with longer work hours allowed when work starts earlier in the day relative to later in the day. Our findings provide support for restricting duty periods by time of day. We found that fatigue is lowest and performance is best during flights that start between 07:00 and 10:59. Flights that started earlier or later than that range of time were associated with elevated levels of fatigue and poorer performance. Such duty periods have the potential to encroach on a pilot's biological drive for sleep, likely accounting for the elevated fatigue and reduced performance that we observed among early starts and late finishes. Notably, the average FDP in our study was much shorter than the allowable maximum duty periods allowed in both Europe and the US. Our finding that fatigue and performance were poorest during late finishes suggests that both time of day and time awake effects interact.

In our study, the average duty duration was around 7 h. This means that a pilot who began duty at 15:29 would have finished the duty at around 22:30 in our study, but the maximal allowable limit when starting duty at 15:29 is 12 hours in both Europe and the US. Given the influence of the circadian rhythm, coupled with presumably elevated sleep pressure due to the additional time awake, it is likely that such a scenario would be associated with even worse performance at the end of the duty relative to what we observed. Others have demonstrated that the length of duty impacts fatigue ratings at the TOD, with longer duties resulting in higher fatigue ratings (Powell et al., 2007). Further research is needed to better understand how the duration of a duty period affects alertness and performance at different times of day.

Although we conducted a large study to evaluate fatigue and performance during daytime short-haul operations, our study is not without limitations. In our study, the duty duration for the late-finishing duties was longer and the number of sectors was higher compared with early starts, and this may have had an additional impact on the increased fatigue levels and worse performance during these duties as has been shown in other studies (Goffeng et al., 2019; Honn et al., 2016; Powell et al., 2007; Sallinen et al., 2021). In addition, the FDPs that we studied were much shorter than the maximal allowable limits, and we did not evaluate flights starting or ending between approximately 01:00 and 04:00. For example, Sallinen et al. (2021) found that fatigue was strongly predicted by night FDPs (between 02:00 and 05:59). Further research is needed to understand how duty periods close to the allowable limit might interact with the time-of-day effects that we observed. Finally, participants self-reported sleep and, as a result, they may have mis-estimated their sleep duration. We conducted additional analyses to investigate the impact of prior sleep on subjective fatigue and PVT performance while controlling for time awake at duty start and found that prior sleep was not related to any of the outcome measures (see Supplemental material).

Overall, we found that flights that encroach on the biological night (the time when the circadian rhythm is not promoting wakefulness (e.g., often between 21:00 and 07:00 with large individual variations in entrained individuals [Arendt, 2010]) are associated with higher self-reported fatigue and poorer performance compared with flights that start in the middle of the day. Early starts were associated with higher fatigue and reduced performance at the beginning of duty compared with duties that started later in the day. However, fatigue and performance were worse at the end of late-finishing duties compared with at the end of early starts. Taken together, our findings suggest that any flight that encroaches on the biological night should be a target for fatigue risk management. Future studies should evaluate countermeasure strategies such as strategic use of caffeine, light, napping, and sleep hygiene to identify tools and approaches that may minimise the deficits associated with working during times that one would typically be asleep.

ACKNOWLEDGMENTS

We would like to thank the pilots for volunteering to participate in this study and to the staff of the airline for help in data collection. We would also like to thank Mr. Gregory Costedoat for reviewing this manuscript. This research was supported by the NASA Systemwide Safety Program.

CONFLICT OF INTEREST

No conflict of interest has been declared by the authors.

AUTHOR CONTRIBUTIONS

LA: conception, methods, data collection, data analysis, wrote the manuscript; SP: data analysis, wrote the manuscript; RC: data analysis; KG: data analysis, manuscript review; JM: development of iPod application, conception, methods, EF: conception, methods, data collection, data analysis, wrote the manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are subject to third-party privacy restrictions. Data may be available upon request on a case-by-case basis with permission from third party.

ORCID

Lucia Arsintescu https://orcid.org/0000-0003-0919-1563
Sean Pradhan https://orcid.org/0000-0003-0957-9816

REFERENCES

Åkerstedt, T., Klemets, T., Karlsson, D., Häbel, H., Widman, L., & Sallinen, M. (2021). Acute and cumulative effects of scheduling on aircrew fatigue in ultra-short-haul operations. *Journal of Sleep Research*, *30*(5), e13305. https://doi.org/10.1111/jsr.13305

Arendt, J. (2010). Shift work: Coping with the biological clock. *Occupational Medicine*, *60*(1), 10–20. https://doi.org/10.1093/occmed/kqp162

Arsintescu, L., Mulligan, J.B., & Flynn-Evans, E.E. (2017). Evaluation of a psychomotor vigilance task for touch screen devices. *Human Factors*, *59*(4), 661–670. https://doi.org/10.1177/0018720816688394

Bakdash, J.Z., & Marusich, L.R. (2017). Repeated measures correlation. *Frontiers in Psychology*, *8*, 456. https://doi.org/10.3389/fpsyg.2017.00456

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Fox, J. (2019). lme4: Linear mixed-effects models using 'eigen' and S4. R package version 1.1-21.

Bermudez, E.B., Klerman, E.B., Czeisler, C.A., Cohen, D.A., Wyatt, J.K., & Phillips, A.J. (2016). Prediction of vigilant attention and cognitive performance using self-reported alertness, circadian phase, hours since awakening, and accumulated sleep loss. *PLoS One*, *11*(3), e0151770. https://doi.org/10.1371/journal.pone.0151770

Bourgeois-Bougrine, S., Cabon, P., Mollard, R., Coblentz, A., & Speyer, J.J. (2003). Fatigue on aircrew from short-haul flights in civil aviation: The effects of work schedules. *Human Factors and Aerospace Safety*, *3*, 177–188.

Cohen, D.A., Wang, W., Wyatt, J.K., Kronauer, R.E., Dijk, D.J., Czeisler, C.A., & Klerman, E.B. (2010). Uncovering residual effects of chronic sleep loss on human performance. *Science Translational Medicine*, *2*(14), 14ra3. https://doi.org/10.1126/scitranslmed.3000458

Commission Regulation (EU) No 83/2014. Official Journal of the European Union, Brussels, L28/170-L28-29.

Doran, S.M., VanDongen, H.P.A., & Dinges, D.F. (2001). Sustained attention performance during sleep deprivation: Evidence of state instability. *Archives Italiennes De Biologie*, *139*, 253–267.

Federal Aviation Administration, Department of Transportation. (2009). *Flightcrew member duty and rest requirements*. 14 CFR Parts 117, 119, and 121. Docket No.: FAA-2009-1093; Amdt. Nos. 117-1, 119-16, 121-357 RIN 2120–AJ58.

Flynn-Evans, E.E., Arsintescu, L., Gregory, K., Mulligan, J., Nowinski, J., & Feary, M. (2018). Sleep and neurobehavioral performance vary by work start time during non-traditional day shifts. *Sleep Health*, *4*(5), 476–484. https://doi.org/10.1016/j.sleh.2018.08.002

Goffeng, E.M., Wagstaff, A., Nordby, K.C., Meland, A., Goffeng, L.O., Skare, Ø., … Lie, J.S. (2019). Risk of fatigue among airline crew during 4 consecutive days of flight duty. *Aerospace Medicine and Human Performance*, *90*(5), 466–474. https://doi.org/10.3357/AMHP.5236.2019

Hirshkowitz, M., Whiton, K., Albert, S.M., Alessi, C., Bruni, O., DonCarlos, L., … Adams Hillard, P.J. (2015). National Sleep Foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health*, *1*(1), 40–43. https://doi.org/10.1016/j.sleh.2014.12.010

Honn, K.A., Satterfield, B.C., McCauley, P., Caldwell, J.L., & Van Dongen, H.P. (2016). Fatiguing effect of multiple take-offs and landings in regional airline operations. *Accident Analysis & Prevention*, *86*, 199–208. https://doi.org/10.1016/j.aap.2015.10.005

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

Hothorn, T., Bretz, F., Westfall, P., Heiberger, R.M., Schuetzenmeister, A., & Scheibe, S. (2019). multcomp: Simultaneous inference in general parametric models. R package version 1.4-10.

International Air Transport Association (IATA), International Civil Aviation Organization (ICAO), and International Federation of Air Line Pilots' Associations (IFALPA). (2015). *Fatigue management guide for the airline operators* (2nd edn). Retrieved from https://www.icao.int/safety/fatiguemanagement/FRMS%20Tools/FMG%20for%20Airline%20Operators%202nd%20Ed%20(Final)%20EN.pdf

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

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2019). lmerTest: Tests in linear mixed effects models. R package version 3.1-0.

Powell, D., Spencer, M.B., Holland, D., Broadbent, E., & Petrie, K.J. (2007). Pilot fatigue in short-haul operations: Effects of number of sectors, duty length, and time of day. *Aviation, Space, and Environmental Medicine*, *78*(7), 698–701.

Roach, G.D., Sargent, C., Darwent, D., & Dawson, D. (2012). Duty periods with early start times restrict the amount of sleep obtained by short-haul airline pilots. *Accident Analysis & Prevention*, *45*, 22–26. https://doi.org/10.1016/j.aap.2011.09.020

Sallinen, M., Sihvola, M., Puttonen, S., Ketola, K., Tuori, A., Härmä, M., … Åkerstedt, T. (2017). Sleep, alertness and alertness management among commercial airline pilots on short-haul and long-haul flights. *Accident Analysis and Prevention*, *98*, 320–329. https://doi.org/10.1016/j.aap.2016.10.029

Sallinen, M., van Dijk, H., Aeschbach, D., Maij, A., & Åkerstedt, T. (2021). A large-scale European Union study of aircrew fatigue during long night and disruptive duties. *Aerospace Medicine and Human Performance*, *91*(8), 628–635. https://doi.org/10.3357/AMHP.5561.2020

Samn, S.W., & Perelli, L.P. (1982). *Estimating aircrew fatigue: A technique with application to airlift operations* (Technical Report No. SAM-TR-82-21). USAF School of Aerospace Medicine.

Van Dongen, H.P.A., Maislin, G., Mullington, J.M., & Dinges, D.F. (2003). 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. https://doi.org/10.1093/sleep/26.2.117

Vejvoda, M., Elmenhorst, E.-M., Pennig, S., Plath, G., Maass, H., Tritschler, K., … Aeschbach, D. (2014). Significance of time awake for predicting pilots' fatigue on short-haul flights: Implications for flight duty time regulations. *Journal of Sleep Research*, *23*(5), 564–567. https://doi.org/10.1111/jsr.12186

Watson, N.F., Badr, M.S., Belenky, G., Bliwise, D.L., Buxton, O.M., Buysse, D., … Tasali, E. (2015). Recommended amount of sleep for a healthy adult: A joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. *Journal of Clinical Sleep Medicine*, *11*(6), 591–592. https://doi.org/10.5665/sleep.4716

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Arsintescu, L., Pradhan, S., Chachad, R. G., Gregory, K. B., Mulligan, J. B., & Flynn-Evans, E. E. (2022). Early starts and late finishes both reduce alertness and performance among short-haul airline pilots. *Journal of Sleep Research*, 31, e13521. https://doi.org/10.1111/jsr.13521