Subjective Measurements of In-Flight Sleep, Circadian Variation, and Their Relationship with Fatigue

Margo J. van den Berg; Lora J. Wu; Philippa H. Gander

BACKGROUND: This study examined whether subjective measurements of in-flight sleep could be a reliable alternative to actigraphic measurements for monitoring pilot fatigue in a large-scale survey.

METHODS: Pilots (3-pilot crews) completed a 1-page survey on outbound and inbound long-haul flights crossing 1–7 time zones ($N = 586$ surveys) between 53 city pairs with 1-d layovers. Across each flight, pilots documented flight start and end times, break times, and in-flight sleep duration and quality if they attempted sleep. They also rated their fatigue (Samn-Perelli Crew Status Check) and sleepiness (Karolinska Sleepiness Scale) at top of descent (TOD). Mixed model ANCOVA was used to identify independent factors associated with sleep duration, quality, and TOD measures. Domicile time was used as a surrogate measure of circadian phase.

RESULTS: Sleep duration increased by 10.2 min for every 1-h increase in flight duration. Sleep duration and quality varied by break start time, with significantly more sleep obtained during breaks starting between (domicile) 22:00–01:59 and 02:00–05:59 compared to earlier breaks. Pilots were more fatigued and sleepy at TOD on flights arriving between 02:00–05:59 and 06:00–09:59 domicile time compared to other flights. With every 1-h increase in sleep duration, sleepiness ratings at TOD decreased by 0.6 points and fatigue ratings decreased by 0.4 points.

DISCUSSION: The present findings are consistent with previous actigraphic studies, suggesting that self-reported sleep duration is a reliable alternative to actigraphic sleep in this type of study, with use of validated measures, sufficiently large sample sizes, and where fatigue risk is expected to be low.

KEYWORDS: self-reported sleep duration, subjective fatigue and sleepiness, fatigue risk management systems.

van den Berg MJ, Wu LJ, Gander PH. *Subjective measurements of in-flight sleep, circadian variation, and their relationship with fatigue*. Aerosp Med Hum Perform. 2016; 87(10):869–875.

In commercial aviation, fatigue has been defined 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 crewmember's alertness and ability to safely operate an aircraft or perform safety-related duties."[12] As humans function optimally with unrestricted sleep during the biological night, fatigue is inevitable in 24/7 operations and, as a hazard to safety, must be managed accordingly. Data-driven fatigue risk management systems (FRMSs) are gaining traction as a more flexible approach for managing pilot fatigue than prescriptive flight and duty time limits. Using a safety management systems approach, FRMSs require a closed process loop consisting of four steps: 1) monitoring pilot fatigue levels; 2) identifying when/where fatigue could represent a hazard; 3) assessing the associated safety risk; and 4) if necessary, implementing additional mitigation strategies to lower the risk, the effectiveness of which will be evident in ongoing data collection in step 1.[12]

Data for the FRMS process loop can come from routine organizational data, for example, planned vs. actual schedules worked, or no-blame pilot fatigue reports. However, in some situations, pilot monitoring data are needed. Different measures can be used for this purpose and their choice needs to be appropriate to the anticipated levels of fatigue and safety risk.[5,8,11] Pilot monitoring data are relatively resource-intensive and time-consuming to collect by comparison with routine

---

From the Sleep/Wake Research Centre, Massey University, Wellington, NZ.

This manuscript was received for review in February 2016. It was accepted for publication in June 2016.

Address correspondence to: Margo J. van den Berg, Sleep/Wake Research Centre, Massey University, Private Box 756, Wellington, New Zealand; m.j.vandenberg@massey.ac.nz.

DOI: 10.3357/AMHP.4587.2016

operational data.[5,8] The measures chosen and acceptable levels for them (safety performance indicators or SPIs) must: be agreed on by regulators, operators, pilots, and scientists as being meaningful and reliable; not jeopardize pilots' ability to perform their operational duties; and have been widely used in aviation, so that data can be compared between different types of operations.[5,11]

Since performance tests and subjective ratings of fatigue and sleepiness capture only certain aspects of waking function, measuring pilots' sleep will arguably provide the most useful information on their fatigue status, since inadequate sleep affects many aspects of waking function.[8] Both the quantity and quality of sleep will determine subsequent alertness levels.[21] Polysomnography (PSG) is recognized as the gold standard for objectively measuring sleep and has been used in some flight crew studies.[4,22,26] However, measuring sleep with PSG is very costly, time consuming, and intrusive, so actigraphy has been widely used as a cheaper and less intrusive alternative.[7,18,27] There are nevertheless significant equipment costs associated with actigraphy, as well as the time required for collecting and analyzing the data. Such cost and time requirements can be important barriers to monitoring pilot sleep.

One study has compared 56 in-flight sleep episodes monitored by actigraphy with pilots' subjective estimates during a single break in the bunk on long-haul flights (mean actigraphic sleep duration on the outbound flight = 2.77 h, on the inbound flight = 4.42 h). For sleep duration, there was a reasonable correlation between the two measures ($R^2 = 0.63$), but with greater variability in the subjective estimates. On the other hand, only weak correlations were found between subjective sleep quality and actigraphic measures of sleep quality (mean activity score, fragmentation index). In addition, correlations for both sleep duration and quality decreased as sleep became shorter and more disrupted.[19]

A more recent study compared bunk sleep recorded with PSG, actigraphy, and self-report from 21 pilots during a 7-h in-flight rest break on ultra-long haul aircraft delivery flights. While average self-reported sleep duration was similar to the average sleep duration recorded with PSG, this study also found only a weak relationship between subjective sleep quality and PSG sleep efficiency.[25] Pilots overestimated their sleep on average by 11 min for sleep durations averaging 175 min, but there was large variability between individuals in the reliability of their estimates. Taken together, the findings of both studies indicate that self-reported sleep duration is unreliable for estimating the sleep durations of individual pilots, but that it may be a suitable, inexpensive, and easily-obtained measure for estimating average sleep duration in a large group of pilots.

The data analyzed in this paper come from a project that was designed to evaluate the potential safety risk associated with new requirements on the distribution of in-flight rest breaks that were introduced by the U.S. Federal Aviation Administration (FAA) in 2012.[2] It requires that the pilot flying during landing has at least 2 consecutive hours in the second half of the flight duty period available for in-flight rest and that the pilot monitoring at landing has a minimum break of 90 min during the flight duty period. The intent of these provisions is to protect the alertness of the landing pilot by providing an adequate opportunity for in-flight sleep and limiting time awake at top of descent, as well as ensuring an acceptable level of alertness of the pilot monitoring by providing a minimum opportunity for in-flight sleep. These requirements were designed for longer flights with four-pilot crews, but had the unintended consequence of forcing the landing pilot on shorter three-pilot flights to always take the last in-flight rest break. This contradicted the established practice of landing pilots on three-pilot flights most commonly selecting the second rest break, which often falls between meal services and, therefore, has the least disruption from activities in the passenger cabin. In addition, depending on the timing of the flight, the last rest break does not always provide the best sleep opportunity relative to the circadian pacemaker cycle. The project was considered low risk because it monitored pilots in three-person crews following their established practice of choosing in-flight rest breaks. There was no evidence that this causes elevated fatigue or safety risk, whereas the new regulatory requirement was untested.

To address this issue, the FAA and the airline scientific advisory team agreed on a study to compare sleep duration and quality of pilots in three-person crews taking the second vs. the third rest break on flights with landing times in three 4-h time bins: 02:00–05:59 and 06:00–09:59, when sleepiness and fatigue were expected to be high, and 22:00–01:59, when it was expected that in-flight sleep would be more difficult because these flights traversed the evening wake maintenance zone. The type of monitoring undertaken in this type of study should be commensurate with the expected level of safety risk.[11] An actigraphic study (estimated *N* required = 210 pilots) was considered inappropriate relative to the expected level of fatigue and safety risk. Therefore, a prospective survey study was undertaken using subjective measurements of in-flight sleep duration and quality. The aim of the present analysis was to determine whether subjective measurements of in-flight sleep collected in this large-scale survey reliably reflected fatigue during flight operations in relationship to: 1) expected circadian variation in these measures; and 2) findings from comparable analyses of data from a large multi-airline database in which sleep was estimated with actigraphy.[6,7]

METHODS

Subjects

The study was registered as a low-risk study with the Massey University Human Ethics Committee: Southern A, which is a registered Institutional Review Board (IRB # 00006014, FWA # 00011627). This means that the project was evaluated by peer review and judged to be low risk and a full ethics application was not required. Participation was voluntary and data confidentiality was strictly maintained. At the end of their participation, pilots were reimbursed per day of participation according to the conditions of an industrial agreement.

Materials

Pilots were asked to complete a one-page survey for an outbound flight and a one-page survey for the subsequent inbound flight. For each flight they were asked to record: times of blocks off (when the aircraft is pushed back from the departure gate) and blocks on (when the aircraft arrives at the destination gate); which rest break they took and when it began and ended; whether they tried to sleep; and if they slept, to estimate how long they slept. If they slept, pilots were asked to rate their sleep quality on a scale from 1 = extremely good to 7 = extremely poor.[25] They were also asked to rate their sleepiness at top of descent (TOD; the beginning of a high workload phase of flight where the procedures for landing are initiated) on the Karolinska Sleepiness Scale (1 = extremely alert, 3 = alert, 5 = neither sleepy nor alert, 7 = sleepy, but no difficulty remaining awake, or 9 = extremely sleepy, fighting sleep) and their fatigue at TOD on the Samn-Perelli Crew Status Check (1 = fully alert, wide awake, 2 = very lively, responsive, but not at peak, 3 = okay, somewhat fresh, 4 = a little tired, less than fresh, 5 = moderately tired, let down, 6 = extremely tired, very difficult to concentrate, or 7 = completely exhausted, unable to function effectively).

The Karolinska Sleepiness Scale (KSS) has been validated[13] and is used to measure subjective sleepiness in both laboratory[1] and field studies.[9,10] In controlled laboratory studies, values of 7 and above on the KSS have been associated with the occurrence of microsleeps (very short periods of uncontrolled sleep).[1]

The Samn-Perelli Crew Status Check (SP) was developed specifically for use with flight crew.[22,23] It has been used in studies focused on sleep loss, fatigue, and performance of flight crew,[16,20,22] as well as in laboratory studies.[3] There is less empirical evidence for a cutoff on the SP. However, airlines have been using values of 5 and above to indicate excessive pilot fatigue.[17]

Procedure

The airline provided information on the class of rest facility available for each study flight. All pilots eligible to fly the Atlanta-based trips targeted for the study were initially notified about the study via the company intranet and could contact the study team for further information. The targeted long-haul trips were operated by three-person crews consisting of a pilot flying, pilot monitoring, and a relief pilot. Each pilot was provided with one scheduled in-flight rest break on every flight. All pilots were acclimated to their base time at the beginning of their flight duty period and, to minimize circadian adaptation to the destination time zone, targeted trips were selected that had short, 1-d layovers between the outbound and inbound flights.

Study packages were made available in the briefing cubicle before each study trip, to be returned by mail after the return flight. The study package included: a numbered copy of the survey which did not include personal identifying information; an information sheet describing the study and what would be involved if a pilot chose to participate; a letter of support from the airline; a return envelope; and a payroll slip to claim reimbursement for their participation. The pay code used was a "miscellaneous additional flight time" so that study participation could not be tracked through the payroll system.

Data from the surveys were entered into a database and all records were cross-checked against the flight information provided by the airline. All departure and arrival times and break start times were converted to domicile time (the time zone of the pilot's crew base). This was assumed to be a reasonable surrogate measure for pilots' circadian phase on outbound and inbound flights, since minimal circadian adaptation was expected during the 1-d layovers. Domicile break start times and arrival times were categorized in the following 4-h time bins: 02:00–05:59; 06:00–09:59; 10:00–13:59; 14:00–17:59, 18:00–21:59, and 22:00–01:59.

Statistical Analysis

Descriptive statistics were calculated using IBM SPSS version 21 (IBM Corp, 2012). Chi-squared tests were undertaken in SAS 9.3 (SAS Institute Inc., 2010). For significant effects, Tukey-type adjustment was used for post hoc pairwise comparisons. Linear mixed model analyses of covariance (ANCOVA) were also undertaken in SAS 9.3. To account for individual differences, participant ID was included as a random effect using the 'variance components' covariance structure. The Kenward-Roger adjustment was applied to the degrees of freedom estimation. For each model, normality of residuals was evaluated visually in addition to using the Shapiro-Wilk statistic. The equality of variance between groups was tested using Levene's test. If variances were not constant, then a more conservative P-value threshold was used ($P < 0.01$ instead of $P < 0.05$). Where outliers were identified, details are provided with the results. Where main effects were statistically significant, the level of significance of post hoc t-tests was adjusted for multiple comparisons using the Holm method. Model structures are described in detail with their findings in the Results section below.

RESULTS

From December 2014 to May 2015, 617 surveys were completed on 131 of the 133 targeted study trips (outbound-inbound) by 298 pilots (1 pilot completed a survey for the outbound flight only; 9 pilots completed 4 surveys on 2 consecutive trips; and 1 pilot completed 6 surveys on 3 consecutive trips), resulting in a participation rate of 77%. Of these 617 surveys, 23 were excluded from the analyses reported here because they had 4-pilot crews instead of 3-pilot crews (including training and line check flights) and 8 surveys were excluded because they were completed on aircraft with Class 1 rest facilities with lie-flat bunks (all other surveys were completed on aircraft with Class 2 rest facilities, which is a seat in the aircraft cabin that allows for a flat or near-flat sleeping position and is separated from passengers by a minimum of a curtain to provide darkness and some sound mitigation). The remaining 586 surveys completed on flights flown between 53 different city pairs included 163 surveys completed on eastward outbound flights, 162 completed on westward inbound flights, 130 completed on southward outbound flights, and 129 completed on northward inbound flights. Surveys completed on one westward outbound flight and one eastward inbound flight were excluded from the analyses because of their unique flight directions. Flight durations ranged between 6.3–11.9 h and crossed 1–7 time zones (adjusted for daylight saving where appropriate).

The number of observations for each domicile time bin is summarized in **Table I**. Since there were insufficient data in break start time bin 14:00–17:59 and in arrival time bin 18:00–21:59, these time bins were excluded from the respective analyses. Flight direction (eastward outbound; southward outbound; westward inbound; northward inbound) was confounded with domicile departure time bin and therefore not considered in the present analyses.

As shown in **Table II**, the majority of pilots attempted sleep during their scheduled rest break. While there was no difference by rest break number in terms of the proportion of pilots attempting sleep [$\chi^2(2) = 1.5804$, $P = 0.4538$], this proportion did vary depending on the domicile break start time [$\chi^2(4) = 13.8684$, $P = 0.0077$]. Post hoc pairwise comparisons showed that the proportion of pilots attempting sleep during breaks starting between 18:00–21:59 was significantly smaller than the proportion of pilots attempting sleep during breaks starting between 22:00–01:59 and 02:00–05:59.

To identify independent predictors of self-reported sleep duration and sleep quality, mixed model ANCOVAs were run which included the following factors: participant ID (random factor); domicile break start time (02:00–05:59, 06:00–09:59, 10:00–13:59, 18:00–21:59, and 22:00–01:59), rest break number (first, second, or third), and flight duration (continuous variable). For the sleep duration model, crew who did not attempt sleep ($N = 28$) were included (sleep duration = 0 min). Since sleep quality could only be rated if sleep was attempted, pilots who did not attempt sleep were excluded from the sleep quality model.

Sleep duration varied with flight duration, rest break number, and domicile break start time (**Table III**). For every 1-h increase in flight duration, sleep duration increased by 10.2 min. Post hoc pairwise comparisons showed that pilots taking the first rest break obtained significantly less sleep than pilots taking the second break [$t(346) = -7.19$, $P < 0.0001$] or third break [$t(431) = -5.05$, $P < 0.0001$]. Sleep durations were not significantly different between the second and third breaks. As shown in **Fig. 1**, pilots obtained significantly more sleep during breaks starting between 02:00–05:59 (domicile time) than in breaks starting between 06:00–09:59 [$t(442) = 3.90$, $P = 0.001$], 10:00–13:59 [$t(464) = 3.49$, $P = 0.0042$], and 18:00–21:59 [$t(376) = 5.14$, $P < 0.0001$]. Pilots also obtained significantly more sleep during breaks starting between 22:00–01:59 (domicile time) in comparison to breaks starting between 18:00–21:59 [$t(469) = -2.87$, $P = 0.03$].

Sleep quality varied with rest break number and domicile break start time (Table III). Pilots taking the first rest break reported poorer sleep quality than pilots taking the second break [$t(321) = 6.18$, $P < 0.0001$] or the third break [$t(394) = 3.85$, $P = 0.0003$]. Pilots taking the third break also reported poorer sleep quality than pilots taking the second break [$t(344) = -2.13$, $P = 0.0340$]. As shown in **Fig. 2**, sleep quality was rated as significantly better for breaks starting between 02:00–05:59 (domicile time) than for breaks starting between 18:00–21:59 [$t(346) = -3.96$, $P = 0.0009$] and 22:00–01:59 [$t(408) = -2.93$, $P = 0.0322$).

The mixed model ANCOVAs for subjective sleepiness and fatigue at TOD included the following factors: participant ID (random factor), domicile arrival time bin (02:00–05:59, 06:00–09:59, 10:00–13:59, 14:00–17:59, and 22:00–01:59), rest break number (first, second, or third), and sleep duration (continuous). Both sleepiness and fatigue ratings varied with sleep duration, rest break number, and domicile arrival times (Table III). With every 1-h increase in sleep duration, sleepiness ratings decreased by 0.6 points and fatigue ratings decreased by 0.4 points. Post hoc pairwise comparisons showed that pilots taking the first break were significantly sleepier at TOD than pilots taking the second break [$t(352) = 7.31$, $P < 0.0001$] or third break [$t(406) = 4.69$, $P < 0.0001$]. Pilots taking the second break were also less sleepy at TOD than pilots taking the third break [$t(366) = -2.47$, $P = 0.0141$]. Similarly, pilots taking the first break were significantly more fatigued at TOD than pilots taking the second break [$t(356) = 6.92$, $P < 0.0001$] or the third break [$t(411) = 3.99$, $P = 0.0002$]. Pilots taking the second break were also less fatigued at TOD than pilots taking the third break [$t(370) = -2.82$, $P = 0.0051$]. As shown in **Fig. 3**, pilots were significantly sleepier at TOD on flights arriving between 02:00–05:59 domicile time than on flights arriving between 10:00–13:59 [$t(462) = 4.95$, $P < 0.0001$], 14:00–17:59 [$t(427) = 5.92$, $P < 0.0001$], and 22:00–01:59 [$t(506) = 4.48$, $P < 0.0001$]. Pilots were also significantly sleepier at TOD on flights arriving between 06:00–09:59 than on flights arriving between 10:00–13:59 [$t(452) = 6.21$, $P < 0.0001$], 14:00–17:59 [$t(470) = 6.89$, $P < 0.0001$], and 22:00–01:59 [$t(482) = 6.04$, $P < 0.0001$]. Post hoc findings for fatigue ratings at TOD were identical to the sleepiness ratings and are shown in **Fig. 4**.

A second set of mixed model ANCOVAs was run for subjective sleepiness and fatigue ratings

Table I. Number of Observations in Each 4-h Time Bin.

	N, DOMICILE BREAK START TIME		*N*, DOMICILE ARRIVAL TIME	
	SLEEP DURATION*	SLEEP QUALITY[‡]	KSS	SAMN-PERELLI
02:00-05:59	121	117	187	186
06:00-09:59	68	60	126	126
10:00-13:59	50	44	64	64
14:00-17:59	28[†]	24[†]	78	79
18:00-21:59	138	118	20[†]	20[†]
22:00-01:59	176	169	107	107

* Including 28 pilots who did not attempt sleep (sleep = 0 min).

[†] Excluded from analyses because *N*s too small.

[‡] Excluding pilots who did not attempt sleep.

Table II. Number of Pilots Who Reported Attempting and Obtaining Sleep.

DOMICILE BREAK START TIME	ATTEMPTED SLEEP YES	ATTEMPTED SLEEP NO	% WHO ATTEMPTED	OBTAINED SLEEP YES	OBTAINED SLEEP NO	% WHO SLEPT
02:00–05:59	120	2	98.3	112	10	91.8
06:00–09:59	64	4	94.1	61	7	89.7
10:00–13:59	46	5	90.2	41	10	80.4
14:00–17:59						
18:00–21:59	125	13	90.6	100	38	72.5
22:00–01:59	173	4	97.7	158	19	89.3
1st break	157	11	93.5	123	45	73.2
2nd break	185	7	96.4	173	19	90.1
3rd break	186	10	94.9	176	20	89.8

at TOD that included flight duration instead of sleep duration (these two variables were colinear). Flight duration was not a significant predictor of either subjective sleepiness or fatigue at TOD.

DISCUSSION

This large-scale survey (*N* = 584) study examined relationships between subjective sleep duration and key operational factors on long range flights (flight duration = 6.3–11.9 h, crossing 1–7 time zones, 1-d layovers between outbound and inbound flights), as well as examining relationships between subjective sleep duration and pilots' ratings of sleepiness and fatigue at TOD. The reliability of these subjective sleep estimates for monitoring pilot fatigue during flight operations is considered below in relationship to: 1) expected circadian variation in these measures, from laboratory studies; and 2) findings from comparable analyses of data from a large multi-airline database with 730 long range and ultra-long range flights (flight duration = 9.8–18.3 h, crossing 5–12 time zones, 1–3 d layovers between outbound and inbound flights).[6,7] Pilots on these longer flights had access to Class 1 rest facilities with lie-flat bunks, whereas pilots in the present study had Class 2 rest facilities (a partitioned-off seat in the passenger cabin that allows horizontal or near-horizontal rest). Both studies used time in the pilots' domicile time zone as a surrogate measure of circadian phase, assuming minimal circadian adaptation to the layover time zone.

Findings in the present study highlight that sleep was more difficult during breaks that started in the 18:00–21:59 time window, which corresponds with the expected time of the evening wake maintenance zone.[14,28] In this time-window, the proportion of pilots who reported obtaining sleep was lowest (73%), the amount of sleep reported was significantly less than for pilots on breaks starting later (Fig. 1), and sleep quality ratings were poorest (Fig. 2). Conversely for breaks starting in the 02:00–05:59 time-window, the proportion of pilots who reported obtaining sleep was highest (92%), the reported amount of sleep obtained was highest (Fig. 1), and sleep quality was rated as best (Fig. 2). This corresponds to the expected time of the window of circadian low (WOCL), when sleepiness is expected to be greatest.[24] In the large multi-airline database, in-flight sleep measurements were analyzed with respect to the domicile time from which flights departed rather than the start time of rest breaks, since on these longer flights, the majority of pilots had two in-flight rest breaks each.[6] Despite the differences between studies, the independent contribution of domicile time of day to variation in total (actigraphic) in-flight sleep in the multi-airline study was similar to that seen for subjective in-flight sleep in the present study.

In the present study, mixed model ANCOVA found that for every 1-h increase in flight duration, reported total in-flight sleep increased by an estimated 10.2 min after controlling for rest break number and break start time (domicile time). In the large multi-airline database, the comparable analysis found that for every 1-h increase in flight duration, total actigraphic in-flight sleep increased by an estimated 10.1 min after controlling for flight direction, domicile departure time, and crew position (landing crew vs. relief crew at landing).[6]

In the present study, subjective ratings of sleepiness and fatigue at TOD were highest

Table III. Effect of Domicile Break Start Time, Rest Break Number, and Flight Duration on Self-Reported Sleep Duration and Sleep Quality.

MEASUREMENT	*N* (USED)/*N* (TOTAL)	FIXED EFFECTS	DF	F-VALUE	*P*-VALUE
Self-reported sleep duration	552/556*	Break start time (domicile)	4, 449	8.92	<0.0001
		Rest break	2, 378	26.81	<0.0001
		Flight duration	1, 482	31.72	<0.0001
Self-reported sleep quality	507/528†	Break start time (domicile)	4, 411	4.57	0.0013
		Rest break	2, 348	19.17	<0.0001
		Flight duration	1, 462	0.89	0.3460
KSS rating	559/564‡	Arrival time (domicile)	4, 419	18.77	<0.0001
		Rest break	2, 372	27.01	<0.0001
		Sleep duration	1, 550	36.20	<0.0001
SP fatigue rating	559/564¶	Arrival time (domicile)	4, 423	16.97	<0.0001
		Rest break	2, 376	23.95	<0.0001
		Sleep duration	1, 551	41.15	<0.0001

* Residual distribution slight negatively skewed and variance not constant for rest break and flight duration; therefore a more conservative alpha level of 0.01 was used. However, results did not change when residual distribution was normalized with reflect and square-root transformation.

† Variance not constant for rest break.

‡ Includes 5 outliers, variance not constant for sleep duration, rest break, and arrival time.

¶ Includes 4 outliers, variance not constant for sleep duration.

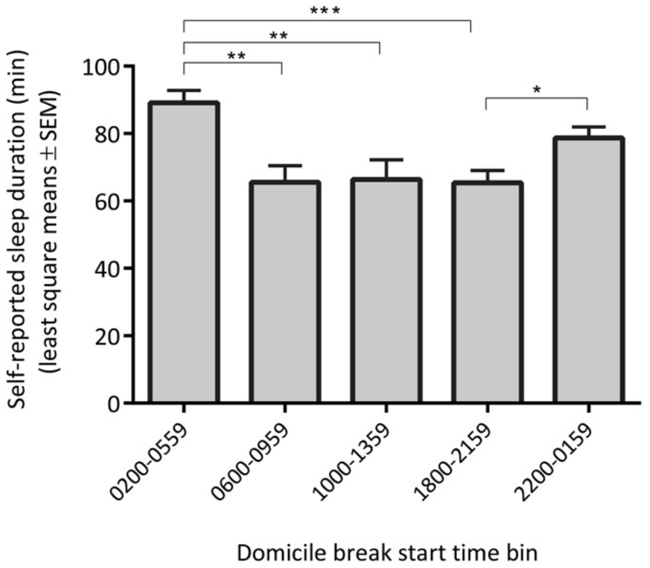

Fig. 1. Least square means (± 1 SEM) self-reported sleep duration (min) by domicile break start time bin; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

during flights arriving between 02:00–05:59 and 06:00–09:59, which is consistent with previous laboratory[3,24] and field studies.[6,7] For every 1-h increase in self-reported sleep duration, sleepiness ratings at TOD decreased by an estimated 0.6 points and fatigue ratings at TOD by 0.4 points. By comparison, in the large multi-airline database, for every 1-h increase in actigraphically recorded total in-flight sleep, sleepiness ratings decreased by an estimated 0.3 points and fatigue ratings by an estimated 0.2 points after controlling for domicile arrival time, flight direction, and crew position.[7] Taking into consideration the operational differences between studies (three-pilot versus four-pilot crews, rest break patterns, crew rest facilities, and flight duration), these comparisons suggest that subjective sleep duration is a useful predictor of pilot sleepiness and fatigue at TOD.

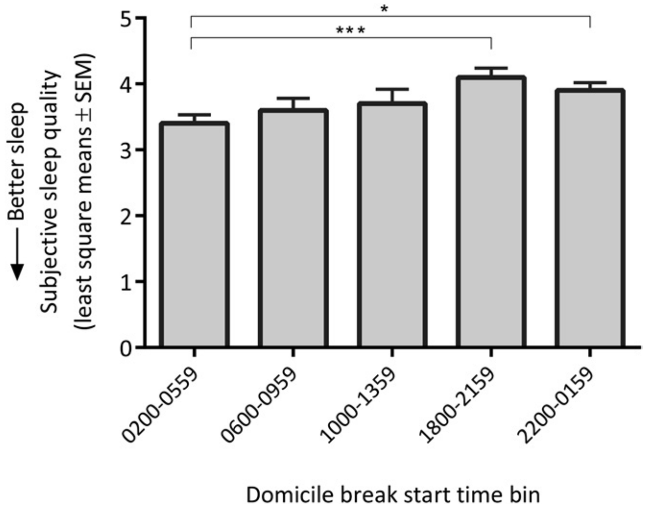

Fig. 2. Least square means (± 1 SEM) self-reported sleep quality by domicile break start time bin; *$P < 0.05$; ***$P < 0.001$.

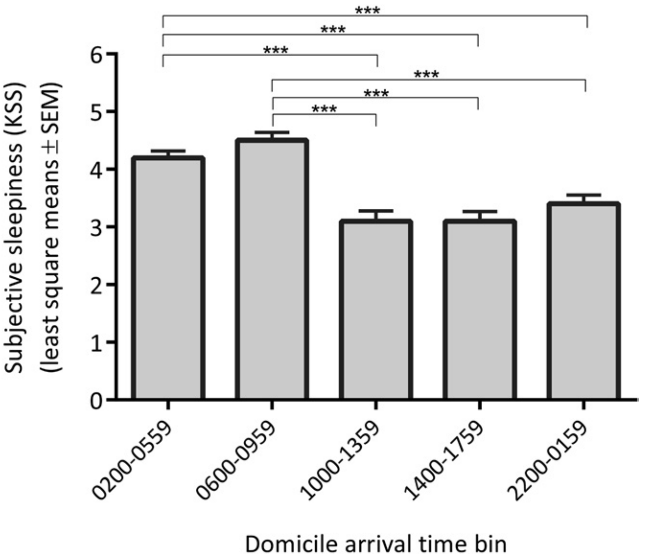

Fig. 3. Least square means (± 1 SEM) Karolinska Sleepiness Scale ratings (KSS) by domicile arrival time bin; ***$P < 0.001$.

Other findings that support the usefulness of subjective measures are the independent relationships between which rest break pilots took (first, second, or third) and their reported in-flight sleep, and sleepiness and fatigue at TOD after controlling for the circadian timing of the break or flight arrival. Pilots taking the first break obtained less sleep and reported poorer quality sleep than those taking the second or third break. This is consistent with previous studies and is possibly attributable to pilots taking the first rest break being awake a shorter time since their last sleep episode on the ground.[15,16] Sleep quality was also rated as poorer in the third rest break than the second rest break. Pilots' comments indicate that this may be related to disturbances associated with the passenger meal service during the third break (the Class 2 rest facility is located in the main

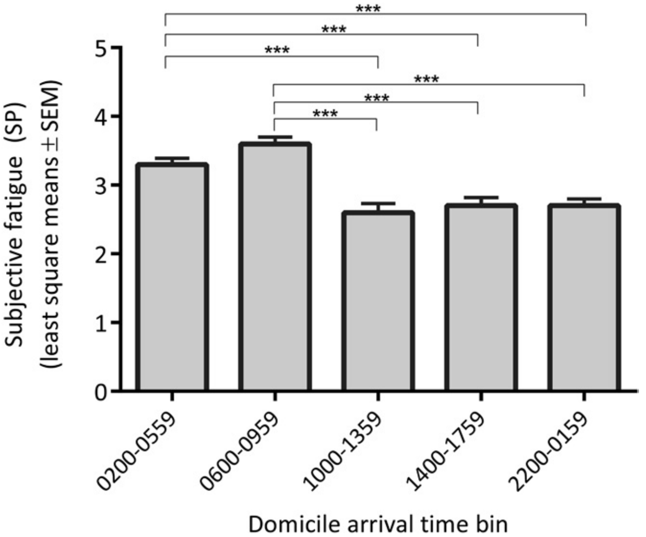

Fig. 4. Least square means (± 1 SEM) Samn-Perelli fatigue ratings (SP) by domicile arrival time bin; ***$P < 0.001$.

passenger cabin). The finding that pilots taking the first break felt more sleepy and fatigued at TOD compared to crew taking the second or third break is consistent with the findings of the multi-airline study, which showed that the duration of time awake at TOD was a significant predictor of subjective fatigue and sleepiness after controlling for flight duration, flight type, total sleep in the 24 h prior to TOD, and domicile arrival time.[7]

In conclusion, the present findings suggest that self-reported sleep duration can be a reliable alternative to actigraphically recorded sleep in field studies monitoring pilot fatigue if they have sufficiently large sample sizes, use validated measures, and where operational fatigue risk is expected to be low. A common concern about subjective measures is that pilots may be influenced by industrial motivations, frustrations with the airline, or other factors when providing subjective data. The relationships seen in the present study with operational factors and circadian variation suggest that the influence of these factors is sufficiently strong to be evident despite potential individual biases in reporting, provided that a large enough group of pilots is studied. Commercial airline pilots in the United States are required to have recurrent fatigue risk management education/training, which includes discussion of the reliability of self-assessment. Pilots' confidence in the airline's safety culture would also be expected to play a role in the reliability of subjective measures, as would their understanding of the importance of their personal role in reporting fatigue hazards.[11,12]

ACKNOWLEDGMENTS

The authors are indebted to the pilots who participated in this study. They also gratefully acknowledge the support and assistance from Jim Mangie at Delta Air Lines, Adrienne Phillips at Delta Air Lines for coordinating the data collection, and Hannah Timms at the Sleep/Wake Research Centre, Massey University, for all data entry.

The project was funded by Delta Air Lines.

Authors and affiliations: Margo J. van den Berg, B.A., Lora J. Wu, M.S., Ph.D., and Phillipa H. Gander, M.Sc.(Hons.), Ph.D., Sleep/Wake Research Centre, Massey University, Wellington, New Zealand.

REFERENCES

1. Åkerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. Int J Neurosci. 1990; 52(1-2):29–37.
2. Federal Aviation Administration. Flight crewmember duty and rest requirements. Oklahoma City (OK): FAA; 2011. FAA-2009-1093; Amendment Nos. 117-1, 119-16, 121-357.
3. Ferguson SA, Paech GM, Sargent C, Darwent D, Kennaway DJ, Roach GD. The influence of circadian time and sleep dose on subjective fatigue ratings. Accid Anal Prev. 2012; 45(Suppl.):50–54.
4. Flight Safety Foundation. Fourth workshop yields insights into early ultra-long-range operations. Flight Saf Dig. 2005; 24(8-9):1–15.
5. Gander PH, Mangie J, van den Berg MJ, Smith AAT, Mulrine HM, Signal TL. Crew fatigue safety performance indicators for Fatigue Risk Management Systems. Aviat Space Environ Med. 2014; 85(2):139–147.
6. Gander PH, Mulrine HM, van den Berg MJ, Smith AAT, Signal TL, et al. Pilot fatigue: relationships with departure and arrival times, flight duration and direction. Aviat Space Environ Med. 2014; 85(8):833–840.
7. Gander PH, Mulrine HM, van den Berg MJ, Smith AAT, Signal TL, et al. Effects of sleep/wake history and circadian phase on proposed pilot fatigue safety performance indicators. J Sleep Res. 2015; 24(1):110–119.
8. Gander PH, Wu LJ, van den Berg MJ, Lamp A, Hoeg L, Belenky G. Fatigue Risk Management Systems. In: Kryger MH, Roth T, editors. Principles and practice of sleep medicine, 6th ed. St. Louis (MO): Elsevier; 2016.
9. Gillberg M, Kecklund G, Åkerstedt T. Relations between performance and subjective ratings of sleepiness during a night awake. Sleep. 1994; 17(3):236–241.
10. Härmä M, Sallinen M, Ranta R, Mutanen P, Muller K. The effect of an irregular shift system on sleepiness at work in train drivers and railway traffic controllers. J Sleep Res. 2002; 11(2):141–151.
11. International Air Transport Association, International Civil Aviation Organisation, International Federation of Airline Pilots Associations. Fatigue management guide for airline operators, 2nd ed. Montreal (Canada): ICAO; 2015.
12. International Civil Aviation Organisation. Manual for the oversight of fatigue management approaches. Montreal (Canada): ICAO; 2016.
13. Kaida K, Takahashi M, Åkerstedt T, Nakata A, Otsuka Y, et al. Validation of the Karolinska sleepiness scale against performance and EEG variables. Clin Neurophysiol. 2006; 117(7):1574–1581.
14. Lavie P. Ultrashort sleep-waking schedule. III.'Gates' and 'forbidden zones' for sleep. Electroencephalogr Clin Neurophysiol. 1986; 63(5):414–425.
15. Pascoe PA, Johnson MK, Montgomery JM, Roberston KA, Spencer MB. Sleep in rest facilities onboard aircraft: questionnaire studies. Farnborough (UK): RAF Institute of Aviation Medicine; 1994. Report No.: IAM 778.
16. Pascoe PA, Johnson MK, Roberston KA, Spencer MB. Sleep in rest facilities on board aircraft: field studies. Farnborough (UK): DERA; 1995. Report No.: DERA/CHS/A&N/CR/95/002.
17. Powell DM, Spencer MB, Petrie KJ. Automated collection of fatigue ratings at the top of descent: a practical commercial airline tool. Aviat Space Environ Med. 2011; 82(11):1037–1041.
18. Roach GD, Darwent D, Dawson D. How well do pilots sleep during long-haul flights? Ergonomics. 2010; 53(9):1072–1075.
19. Robertson K, Spencer M, Stone BM, Akerstedt T, Gillberg M, et al. Predicting alertness in future ultra long-range operations: a validation study by ECASS. Farnborough (UK): QinetiQ; 2002. Report No.: QINETIQ/KI/CHS/CR021119/2.0.
20. Robertson KA, Spencer MB, Stone BM, Johnson MK. Scheduling the on-board rest of aircrew. Farnborough (UK): DERA; 1997. Report No.: DERA/CHS/PP5/CR97095/1.0.
21. Roehrs T, Carskadon M, Dement W, Roth T. Daytime sleepiness and alertness. In: Kryger M, Roth T, Dement W, editors. Principles and practice of sleep medicine, 5th ed. St. Louis: Elsevier Saunders; 2011:42–53.
22. Samel A, Wegmann H-M, Vejvoda M. Aircrew fatigue in long-haul operations. Accid Anal Prev. 1997; 29(4):439–452.
23. Samn SW, Perelli LP. Estimating aircrew fatigue: a technique with implications to airlift operations. Brooks AFB (TX): USAF School of Aerospace Medicine; 1982. Report No.: SAM-TR-82-21.
24. Sargent C, Darwent D, Ferguson SA, Roach GD. Can a simple balance task be used to assess fitness for duty? Accid Anal Prev. 2012; 45(Suppl.):74–79.
25. Signal TL, Gale J, Gander PH. Sleep measurement in flight crew: comparing actigraphic and subjective estimates to polysomnography. Aviat Space Environ Med. 2005; 76(11):1058–1063.
26. Signal TL, Gander PH, van den Berg MJ, Graeber RC. In-flight sleep of flight crew during a 7-hour rest break: implications for research and flight safety. Sleep. 2013; 36(1):109–115.
27. Signal TL, Mulrine HM, van den Berg MJ, Smith AAT, Gander PH. Mitigating and monitoring flight frew Fatigue on a westward ultra long range flight. Aviat Space Environ Med. 2014; 85(12):1199–1208.
28. Strogatz SH, Kronauer RE, Czeisler CA. Circadian pacemaker interferes with sleep onset at specific times each day: role in insomnia. Am J Physiol. 1987; 253(1, Pt. 2):R172–R178.