Scientific Research Publishing

Open Journal of Anesthesiology, 2026, 16(4), 97-115
https://www.scirp.org/journal/ojanes
ISSN Online: 2164-5558
ISSN Print: 2164-5531

Shift Patterns, Fatigue, and Cognitive Performance in Junior Anesthetists: Evidence from a Stroop Test-Based Pilot Study

Rachel Rui Zhi Lim[1], Sheryl Wei Lin Chee[1], Xuan Han Koh[2], Yanzhi Zeng[1]

[1]Department of Anesthesia and Surgical Intensive Care, Changi General Hospital, Singapore, Singapore
[2]Health Services Research, Changi General Hospital, Singapore, Singapore
Email: Rachel.l.r.z@singhealth.com.sg

How to cite this paper: Lim, R.R.Z., Chee, S.W.L., Koh, X.H. and Zeng, Y.Z. (2026) Shift Patterns, Fatigue, and Cognitive Performance in Junior Anesthetists: Evidence from a Stroop Test-Based Pilot Study. *Open Journal of Anesthesiology*, **16**, 97-115.
https://doi.org/10.4236/ojanes.2026.164010

Received: February 14, 2026
Accepted: April 21, 2026
Published: April 24, 2026

Abstract

Background: Fatigue in anesthesia is a significant concern, impacting patient safety and practitioner well-being due to the demanding nature of the work and irregular, extended hours. This pilot study evaluates the effect of different shift types on fatigue and cognitive performance in junior anesthetists, using the Stroop Color and Word Test (SCWT) as a surrogate for cognitive performance. **Methods:** A prospective observational crossover pilot study was conducted with eight trainees at a tertiary hospital. Participants completed sixty-three post-shift surveys and the SCWT after various shift types, including 8h-Short Day, 12h-Long Day, 12h-Night Shift, 24h-Weekday Call, and 24h-Weekend Call. Subjective fatigue (Relative Perceived Exhaustion, RPE), perceived workload, sleep duration, and SCWT performance (Inverse Efficiency Score, IES) were measured. Linear mixed-effects regression was used to assess associations between shift type and performance, adjusting for learning effects. **Results:** Shift type was independently associated with cognitive performance as measured by IES ($P < 0.001$). Twelve-hour Long Day and 24-hour Weekend Call shifts were linked to significantly higher (worse) IES scores compared to 8-hour Short Day shifts (mean differences: 223 (95% CI 19 - 428) and 809 (95% CI 490 - 1127), respectively), indicating reduced efficiency. Higher RPE correlated with higher IES ($P = 0.005$). No significant association was found between IES and self-reported busyness or sleep duration. **Conclusions:** Extended shifts are associated with increased fatigue and reduced cognitive performance in anesthetists. The SCWT/IES and RPE are feasible tools for assessing fatigue-related performance, supporting the need for evidence-based rostering and fatigue management strategies in anesthesia.

Keywords

Anesthesia, Cognitive Performance, Fatigue, Stroop Color and Word Test, Occupational Health

1. Introduction

Fatigue in anesthesia practice is a critical phenomenon that requires ongoing attention from both practitioners and healthcare systems to maintain patient safety and support practitioner well-being. Indeed, the literature suggests that the prevalence of fatigue among anesthesia providers may be as high as 84%, highlighting the magnitude of this issue within clinical practice [1]. Anesthesia practice is physically and mentally taxing due to factors like long hours, significant responsibility, and the necessity for continuous vigilance. The high-pressure environment in operating rooms and intensive care units demands quick, accurate, and precise actions and decisions. This constant pressure to perform and anticipate complications can result in cognitive overload and mental fatigue, which can be detrimental to practitioner well-being, work performance, and patient safety [2].

Fatigue has been shown to negatively impact work performance across multiple domains—cognitive and psychomotor performance, risk assessment, decision making, performance of complex tasks, communication between colleagues and patients, and empathy towards patients [3]-[5]. Several studies have found a strong association between sleep deprivation and decreased cognitive function and performance in anesthetists [6]-[8]. Of note, Arzalier-Daret *et al.* found that sleep-deprived anesthesia residents had impaired crisis management performance in the form of drug administration errors, delayed recognition of hypotension, and ineffective communication [9]. Non-technical skills such as teamwork, situational awareness, and decision making have also been found to be impacted by sleep deprivation [10].

The scoping review by Scholliers *et al.* summarized the latest evidence on the impact of fatigue on the performance of anesthesia providers. Various cognitive tests were used to study different cognitive domains such as vigilance, reaction time, memory, and executive function. Many studies used proxies for clinical performance, such as psychomotor performance (most commonly using the psychomotor vigilance task) and executive function. Most, but not all, studies concluded that sleep deprivation and fatigue negatively affect cognitive performance [1]. Both Scholliers and Ippolito *et al.* concluded that the available evidence suggests that fatigue among anesthesia providers is associated with clinical performance impairments and poorer outcomes among patients [1] [2].

The effects of fatigue on practitioner well-being must not be ignored as well. Chronic fatigue is associated with an increased risk of cardiovascular disease and sleep disorders, mental health issues like burnout and depression, reduced job satisfaction, and an increased risk of vehicle accidents [11]. In light of this, more attention should be paid to fatigue as a workplace hazard in anesthesia practice.

Aviation and anesthesia share many similarities in that both operate in high-risk, high-reliability environments where safety is critical and constant vigilance is required, and operators are at risk of fatigue, long hours, and shift work. Fatigue risk management systems have long been in place in aviation to mitigate the detrimental consequences of fatigue. These include scientific scheduling with regulated duty hours, as well as a supportive reporting culture that allows staff to report fatigue without punishment [12]. There are many lessons that may be gleaned from this and adapted for healthcare and anesthesia practice.

This pilot study was designed to evaluate the effects of different anesthesia shift types on objective performance metrics in junior anesthetists, using the Stroop Color and Word Test (SCWT) and Inverse Efficiency Score as a surrogate for cognitive performance. It is the first study to correlate subjective levels of fatigue (Relative Perceived Exhaustion or RPE) with objective cognitive performance in junior anesthetists. The primary objective of the study was to assess the feasibility, utility, and validity of this methodological approach for potential application in larger-scale research. The secondary aims were to generate preliminary data to inform sample size calculations for future investigations and to explore the association between extended work shifts and clinician fatigue and cognitive impairment.

The SCWT is a cognitive assessment tool used to evaluate an individual's attentional control, ability to suppress automatic responses, and mental flexibility [13]. During the test, participants are shown color words (like "RED" or "BLUE") printed in either concordant or discordant ink colors and are required to state the ink color, which may differ from the displayed word.

This task creates a challenge because it pits the automatic impulse to read the word against the task of identifying the ink color, a phenomenon known as the Stroop Effect [14]. The test is widely used to assess cognitive control and an individual's ability to resist distractions and suppress habitual responses, and it has been shown to be negatively impacted by fatigue and lack of sleep. Fatigued individuals often take longer and make more mistakes, indicating that the test is a useful indicator of cognitive fatigue, particularly in how it affects attention and self-regulation [15]-[17].

Neuroimaging techniques including magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and positron emission tomography (PET) have shown that there are two main areas in the brain that are involved in the processing of the SCWT—the anterior cingulate cortex and the dorsolateral prefrontal cortex. More specifically, while both are activated when resolving conflicts and catching errors, the dorsolateral prefrontal cortex assists in memory and other executive functions, and the anterior cingulate cortex is used to select an appropriate response and allocate attentional resources [18].

Ultimately, this work contributes toward the long-term goal of identifying an evidence-based rostering system that effectively balances clinician rest, performance, patient care, and operational demands, thereby mitigating fatigue-related risks.

2. Materials and Methods

This was a prospective observational crossover study conducted at Changi General Hospital, a 1,000-bed tertiary hospital in Singapore. At our center, the job scopes within the Department of Anesthesia and Surgical Intensive Care are divided among junior trainees (pre-intermediate examinations), senior trainees (post-intermediate examinations and pre-specialist examinations), and specialists (post-specialist examinations). Participants were junior trainees in the Department, who were evaluated after completing a work shift.

2.1. Eligibility Criteria and Recruitment

All junior doctors working in the Department of Anesthesia and Surgical Intensive Care who were rostered to the Surgical Intensive Care Unit (SICU) or Operating Theatre (OT) duties during the study period were eligible. Inclusion criteria were: i) junior trainees who were pre-intermediate and pre-specialist examinations, ii) participation in the standard shift roster, and iii) ability to complete post-shift assessments via an electronic device. There were no formal exclusion criteria.

2.2. Ethical Considerations

Before study initiation, an application to the SingHealth Centralized Institutional Review Board (CIRB) was submitted (CIRB iSHaRe reference: 202310-00035).

Following review, the CIRB determined that the project did not constitute human subjects research and therefore did not require further ethical deliberation (CIRB determination reference: 2024/2200). Participation was voluntary, and all data were de-identified prior to analysis. The study was conducted in accordance with institutional policy and the Declaration of Helsinki.

2.3. Study Setting and Participants

The department has the following junior manpower allocation system in place:

Junior doctors rostered to the Surgical Intensive Care Unit (SICU) work a variety of the following shifts:

1) 8h-Short Day (8 am to 4 pm)

2) 12-Hour Day (8 am to 8 pm)

3) 12-Hour Night Shift (8 pm to 8 am)

Doctors rostered in the Operating Theatre (OT) work the following shifts:

1) 8h-OT Day (8 a.m. to 4 p.m.)

2) 24-Hour Weekday Call (May have a designated rest period of up to 8 hours depending on manpower)

3) 24-Hour Weekend Call (No designated rest period)

Four junior doctors are rostered to the SICU every month. After completing the month in the SICU, they return to work in the OT, and another four doctors replace them.

Over the course of two months, a total of eight eligible trainees were approached, all of whom consented and were enrolled. They were aged 26 to 32 years old and had less than two years of anesthesia training each. All participants rotated through both SICU and OT postings during the study period and were equally likely to be assigned to any shift type as part of a common pool. Rostering in both SICU and OT was done by a points system to even out the shift types. Each participant was de-identified and allocated a letter from A to H.

Following every shift, each participant filled out a written questionnaire and took the SCWT [19], both administered through links provided to their personal mobile devices, with the data recorded electronically. Both the questionnaire and SCWT were completed as soon as possible after the end of each shift and the completion of handover duties. All participants were given a one-time standardized practice run to familiarize themselves with the test.

As this was a pilot study conducted within the constraints of the departmental rotation structure, the sample size was based on feasibility rather than a priori power calculations. The primary aim was methodological feasibility and effect-size estimation to inform sample-size calculations for future adequately powered studies.

Across the two-month study period, 64 shift assessments were attempted. One SCWT result was unavailable due to a technical device failure, leaving 63 complete post-shift assessments for analysis.

The written questionnaire was developed by the authors de novo and included the following:

1) Shift type

2) Subjective assessment of fatigue—Relative Perceived Exhaustion (RPE)

3) Subjective assessment of workload—difficulty of the shift before/after midnight

4) Subjective assessment of rest—sleep (total and duration of interrupted sleep) they managed during the night shift (or the night before if they were on a day shift)

5) Subjective assessment of their current performance as a percentage of their peak performance

As an objective assessment of their fatigue, the 1-minute SCWT yielded the following results: number of correct answers, number of errors, proportion of errors (PE), and the average reaction time for a correct answer (RT). With these data, the Inverse Efficiency Score (IES) was calculated.

2.4. Inverse Efficiency Score

The Inverse Efficiency Score (IES) in milliseconds (ms) is a measure used in cognitive psychology to quantify the efficiency with which a person performs a task [20]. It is designed to combine both accuracy and reaction time (RT) into a single score, offering a more holistic view of performance efficiency, where a higher performance efficiency is reflected in faster responses with higher accuracy [21] [22].

$$\text{IES} = \text{RT}/[1-\text{PE}] \text{ or } \text{RT}/\text{PC}$$

where PE = proportion of errors, PC = proportion of correct responses.

- A low IES indicates high efficiency, *i.e*., the participant is completing the task quickly and accurately.
- A high IES suggests lower efficiency, *i.e*., the participant is either taking more time to complete the task or making more errors.

The collected data were analyzed by a Senior Analyst from our institution's Department of Data Science, who provided oversight and expertise in the statistical evaluation.

2.5. Statistical Analysis

Descriptive statistics of the characteristics and survey responses of anesthetists' work shifts were reported as number and percent for categorical data, mean ± standard deviation (SD) for normally distributed data, and median and interquartile range (IQR) for non-normally distributed data.

The scatter plot with locally weighted scatterplot smoothing (LOWESS) and a 95% confidence band was used to visualize the relationship between SCWT order and IES. SCWT order denotes the sequence of test administrations for each participant across successive time points. Specifically, SCWT order 1 corresponds to the assessment conducted following the first shift, SCWT order 2 to the assessment following the second shift, and subsequent orders reflect the same pattern.

We generated box plots of the IES by anesthetist and at different values of subjective measures of rest or workload. To determine if shift type was associated with SCWT performance, we performed a multivariable linear mixed-effects regression of the IES on shift type and SCWT order. Random intercepts for anesthetists accounted for clustering within anesthetists and inter-anesthetist variability. The likelihood ratio test of the full model versus a nested model with only SCWT order indicated whether shift type was associated with the IES after adjusting for SCWT order. The adjusted difference in means and 95% confidence interval (CI) were reported. Glass's Δ indicated the adjusted difference expressed in reference group SDs, and the ratio of means (RoM) indicated the relative difference of geometric means. Partial R^2 quantified the incremental variance in IES explained by shift type and SCWT order after accounting for inter-anesthetist variability, with 95% CI estimated via bootstrap resampling with 1000 iterations.

We performed a leave-one-out sensitivity analysis to examine whether the main results were sensitive to any one data point. A cumulative logit mixed-effects regression of the RPE on shift type with random intercepts for anesthetists was used to determine whether shift type was associated with RPE. To determine whether various subjective measures of rest or workload were associated with the IES, we performed Kruskal-Wallis tests. When the Kruskal-Wallis test $P < 0.05$, we performed post-hoc pairwise comparisons using Dunn's test with multiplicity adjustment using the Holm method.

To determine the sample size for a future definitive trial, we performed a power analysis using Monte Carlo simulation based on the baseline mean and variance components from our pilot linear mixed model. We estimated the number of anesthetists required to detect a 5% relative difference in IES with 80% power at α = 0.05. Statistical tests were two-sided with a 0.05 significance level. All analyses were conducted using R version 4.4.2 (R Core Team, 2024), the "simr" R package, and Stata 18 (College Station, TX: StataCorp LLC).

2.6. Reporting Standards

This manuscript was prepared in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for observational studies.

3. Results

A total of eight junior anesthetists contributed 63 complete post-shift assessments. One attempted assessment was excluded due to missing SCWT data. No other post-shift questionnaires or SCWT assessments were incomplete. Table 1 shows the characteristics and survey responses of the anesthetists' work shifts overall, and by shift type.

Table 1. Characteristics and survey responses of anesthetists' work shifts.

Characteristic/ Survey response	All shifts (n = 63)	8h-Short Day (n = 13)	12h-Long Day (n = 15)	12h-Night Shift (n = 28)	24h-Weekday Call (n = 4)	24h-Weekend Call (n = 3)
No. of anesthetists	8	6	7	8	4	3
How tired are you feeling now?						
1 (Not tired)	1 (1.6)	0 (0)	0 (0)	1 (3.6)	0 (0)	0 (0)
2 (Slightly tired)	11 (17.5)	5 (38.5)	2 (13.3)	3 (10.7)	0 (0)	1 (33.3)
3 (Tired)	25 (39.7)	6 (46.2)	4 (26.7)	13 (46.4)	1 (25)	1 (33.3)
4 (Very tired)	20 (31.8)	1 (7.7)	5 (33.3)	10 (35.7)	3 (75)	1 (33.3)
5 (Extremely tired)	6 (9.5)	1 (7.7)	4 (26.7)	1 (3.6)	0 (0)	0 (0)
How busy was your shift before midnight?						
Not busy	27 (42.9)	4 (30.8)	5 (33.3)	17 (60.7)	0 (0)	1 (33.3)
Quite busy	24 (38.1)	7 (53.9)	5 (33.3)	8 (28.6)	4 (100)	0 (0)
Very busy	12 (19.1)	2 (15.4)	5 (33.3)	3 (10.7)	0 (0)	2 (66.7)
How busy was your shift after midnight?						
Not busy	16 (38.1)	1 (25)	0 (0)	14 (50)	0 (0)	1 (33.3)
Quite busy	21 (50)	3 (75)	2 (66.7)	12 (42.9)	3 (75)	1 (33.3)
Very busy	5 (11.9)	0 (0)	1 (33.3)	2 (7.1)	1 (25)	1 (33.3)

Continued

How many hours of sleep did you get during your shift in total?						
0 - 2	18 (28.6)	1 (7.7)	4 (26.7)	8 (28.6)	3 (75)	2 (66.7)
2 - 4	16 (25.4)	1 (7.7)	0 (0)	14 (50)	1 (25)	0 (0)
4 - 6	20 (31.8)	6 (46.2)	7 (46.7)	6 (21.4)	0 (0)	1 (33.3)
>6	9 (14.3)	5 (38.5)	4 (26.7)	0 (0)	0 (0)	0 (0)
How interrupted was your sleep?						
1 (At least 4 hours of continuous sleep)	24 (38.1)	10 (76.9)	12 (80)	2 (7.1)	0 (0)	0 (0)
2 (At least 2 hours of continuous sleep)	14 (22.2)	2 (15.4)	1 (6.7)	9 (32.1)	1 (25)	1 (33.3)
3 (At least 1 hour of continuous sleep)	8 (12.7)	0 (0)	0 (0)	6 (21.4)	1 (25)	1 (33.3)
4 (At least 30 minutes of continuous sleep)	12 (19.1)	0 (0)	0 (0)	9 (32.1)	2 (50)	1 (33.3)
5 (Longest sleep period < 30 minutes)	5 (7.9)	1 (7.7)	2 (13.3)	2 (7.1)	0 (0)	0 (0)
What percentage relative to your peak performance do you feel you are at?						
≤20%	8 (12.7)	0 (0)	5 (33.3)	3 (10.7)	0 (0)	0 (0)
21% - 40%	16 (25.4)	3 (23.1)	2 (13.3)	8 (28.6)	2 (50)	1 (33.3)
41% - 60%	17 (27.0)	2 (15.4)	3 (20)	10 (35.7)	1 (25)	1 (33.3)
61% - 80%	18 (28.6)	7 (53.9)	4 (26.7)	5 (17.9)	1 (25)	1 (33.3)
81% - 100%	4 (6.4)	1 (7.7)	1 (6.7)	2 (7.1)	0 (0)	0 (0)
SCWT order						
1st test	7 (11.1)	5 (38.5)	1 (6.7)	0 (0)	0 (0)	1 (33.3)
2nd test	8 (12.7)	1 (7.7)	2 (13.3)	3 (10.7)	2 (50)	0 (0)
3rd test	7 (11.1)	1 (7.7)	1 (6.7)	5 (17.9)	0 (0)	0 (0)
4th test	7 (11.1)	1 (7.7)	1 (6.7)	4 (14.3)	0 (0)	1 (33.3)
≥5 tests	34 (54.0)	5 (38.5)	10 (66.7)	16 (57.1)	2 (50)	1 (33.3)
Inverse Efficiency Score (ms)						
Mean ± SD	1264 ± 384	1202 ± 213	1359 ± 277	1169 ± 187	1259 ± 246	1961 ± 1502
Median (IQR)	1176 (1053 - 1392)	1176 (1052 - 1302)	1349 (1132 - 1667)	1132 (1052 - 1263)	1229 (1077 - 1441)	1173 (1017 - 3692)
Min, max	856, 3692	923, 1621	968, 1899	856, 1662	1000, 1577	1017, 3692

Continued

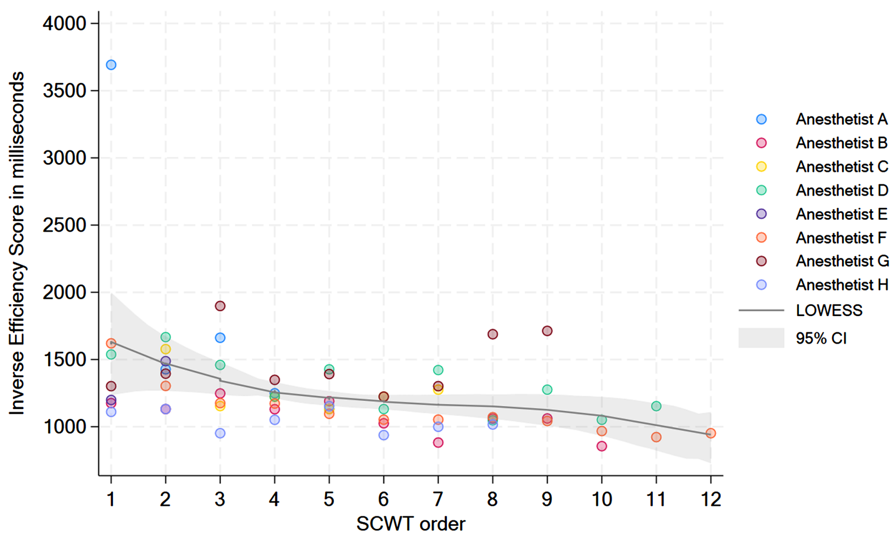
			Relative Perceived Exhaustion (RPE)			
Mean ± SD	3.3 ± 0.9	2.8 ± 0.9	3.7 ± 1.0	3.3 ± 0.8	3.8 ± 0.5	3 ± 1
Median (IQR)	3 (3 - 4)	3 (2 - 3)	4 (3 - 5)	3 (3 - 4)	4 (3.5 - 4)	3 (2 - 4)
Min, max	1, 5	2, 5	2, 5	1, 5	3, 4	2, 4

IQR = Interquartile range. SD = Standard deviation.

3.1. Association between Shift Type and SCWT Performance Calculated in IES

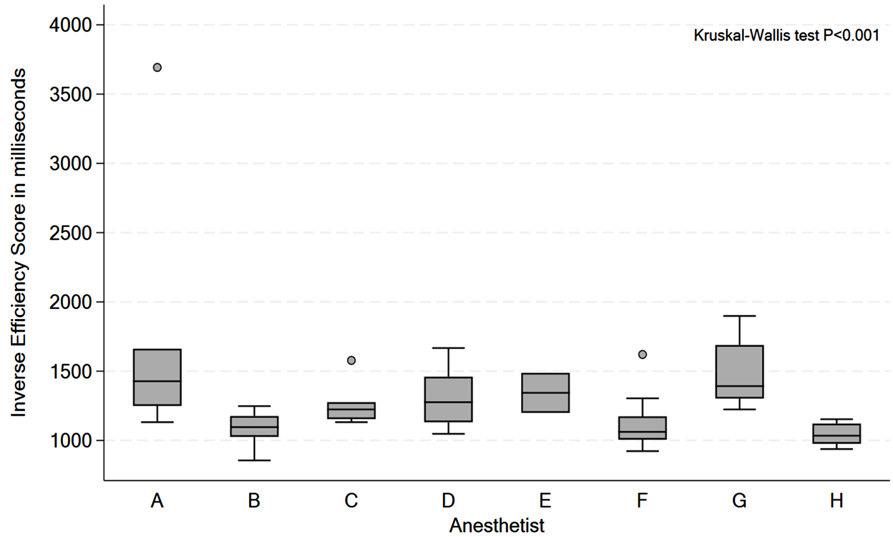
A negative association between SCWT order and IES was taken to suggest a learning effect of the SCWT, which appeared to be non-linear, where the effect was greatest from the 1st to the 4th or 5th test, before plateauing thereafter (Figure 1). In addition, there appeared to be variability in SCWT performance across anesthetists (Figure 2).

After adjusting for SCWT order, 12h-Long Day was associated with a 233 ms higher mean IES compared to 8h-Short Day (95% CI 19 to 428, P = 0.032; Table 2). This corresponded to a 1.05 SD (95% CI 0.00 to 2.10) and 15% (RoM 95% CI 1.02 to 1.31) higher mean IES. 24h-Weekend Call was associated with an 809 ms higher mean IES compared to 8h-Short Day (95% CI 490 to 1127, P < 0.001; Table 2). This corresponded to a 3.80 SD (95% CI 2.17 to 5.42) and 44% (RoM 95% CI 1.19 to 1.75) higher mean IES. The other shifts did not show a difference in IES compared to 8h-Short Day (Table 2). Work shift type and SCWT order uniquely explained 18.6% (95% CI 6.4 to 39.5) and 16.4% (95% CI 3.8 to 37.5) of the variance in IES, respectively.



A locally weighted scatterplot smoothing curve was used to visualize any non-linearity in the data. Inverse Efficiency Score (ms) = (average reaction time in ms)/(1 − proportion of errors), where proportion of errors = (no. of errors)/(total no. of responses). LOWESS = Locally weighted scatterplot smoothing.

Figure 1. Scatterplot of the inverse efficiency score (IES) against SCWT order.



Inverse Efficiency Score (ms) = (average reaction time in ms)/(1 − proportion of errors), where proportion of errors = (no. of errors)/(total no. of responses).

Figure 2. Box plots of inverse efficiency score by anesthetist.

Table 2. Multivariable linear mixed model of inverse efficiency score (IES) on anesthetists' work shift type and SCWT order.

Variable	Difference in means (95% CI)	Glass's Δ (95% CI)	Ratio of means (95% CI)	Partial R^2 (95% CI)	P value
Work shift				0.186 (0.064 to 0.395)	<0.001
8h-Short Day	0 (ref.)	-	-		-
12h-Long Day	223 (19 to 428)	1.05 (0.00 to 2.10)	1.15 (1.02 to 1.31)		0.032
12h-Night Shift	93 (−97 to 283)	0.44 (−0.54 to 1.41)	1.04 (0.93 to 1.17)		0.338
24h-Weekday Call	170 (−128 to 468)	0.80 (−0.71 to 2.31)	1.10 (0.92 to 1.32)		0.262
24h-Weekend Call	809 (490 to 1127)	3.80 (2.17 to 5.42)	1.44 (1.19 to 1.75)		<0.001
SCWT order				0.164 (0.038 to 0.375)	<0.001
1st test	0 (ref.)	-	-		-
2nd test	−267 (−546 to 11)	−0.29 (−0.62 to 0.04)	0.90 (0.76 to 1.06)		0.060
3rd test	−273 (−563 to 16)	−0.30 (−0.64 to 0.05)	0.88 (0.74 to 1.05)		0.065
4th test	−540 (−821 to −258)	−0.59 (−0.92 to −0.26)	0.76 (0.64 to 0.90)		<0.001
≥5 tests	−500 (−738 to −263)	−0.55 (−0.83 to −0.26)	0.74 (0.64 to 0.86)		<0.001

All 63 work shifts had available work shift and SCWT order data and were included in the multivariable linear mixed model.

One work shift had a much larger IES (3692 ms) compared to all other work shifts (856 - 1899 ms). After excluding this work shift in a sensitivity analysis, the observation of a higher mean IES for 12h-Long Day persisted (Table 3).

Table 3. Multivariable linear mixed model of inverse efficiency score on anesthetists' work shift type and SCWT order, excluding one work shift with an inverse efficiency score of 3692.

Variable	Difference in means (95% CI)	Glass's Δ (95% CI)	Ratio of means (95% CI)	Partial R^2 (95% CI)	P value
Work shift				0.079 (0 to 0.297)	0.034
8h-Short Day	0 (ref.)	-	-		-
12h-Long Day	155 (36 to 273)	0.73 (0.12 to 1.33)	1.12 (1.02 to 1.24)		0.010
12h-Night Shift	8 (−102 to 119)	0.04 (−0.53 to 0.60)	1.01 (0.92 to 1.11)		0.882
24h-Weekday Call	70 (−102 to 242)	0.33 (−0.55 to 1.21)	1.06 (0.91 to 1.22)		0.425
24h-Weekend Call	85 (−138 to 308)	0.40 (−0.74 to 1.54)	1.07 (0.89 to 1.30)		0.454
SCWT order				0.205 (0.110 to 0.422)	<0.001
1st test	0 (ref.)	-	-		-
2nd test	12 (−155 to 179)	0.06 (−0.81 to 0.93)	1.00 (0.87 to 1.16)		0.889
3rd test	12 (−162 to 186)	0.06 (−0.86 to 0.97)	0.99 (0.85 to 1.15)		0.893
4th test	−163 (−338 to 12)	−0.78 (−1.70 to 0.13)	0.88 (0.76 to 1.02)		0.068
≥5 tests	−224 (−369 to −78)	−1.07 (−1.84 to −0.31)	0.83 (0.73 to 0.94)		0.003

All 62 work shifts had available work shift and SCWT order data and were included in the multivariable linear mixed model. Statistical analysis: We performed a multivariable linear mixed-effects regression of the IES on shift type and SCWT order, excluding one work shift with an IES of 3692 ms. Random intercepts for anesthetists accounted for clustering within anesthetists and inter-anesthetist variability. The likelihood ratio test of the full model versus a nested model with only SCWT order indicated if shift type was associated with the IES after adjusting for SCWT order. The adjusted difference in means and 95% confidence interval (CI) were reported. Glass's Δ indicated the adjusted difference expressed in reference group SDs and the ratio of means (RoM) indicated the relative difference of geometric means. Partial R^2 quantified the incremental variance in IES explained by shift type and SCWT order after accounting for inter-anesthetist variability, with 95% CI estimated via bootstrap resampling with 1000 iterations.

3.2. Association between Shift Type and Relative Perceived Exhaustion (RPE)

The cumulative logit mixed model of RPE on shift type (Table 4) showed that 12h-Long Day may be associated with 6.8 times the odds of higher RPE compared to 8h-Short Day (95% CI 1.4 to 31.7, P = 0.015), while 24h-Weekday Call with a designated rest period may be associated with 10.1 times the odds of higher RPE compared to 8h-Short Day (95% CI 1.2 to 81.7, P = 0.031).

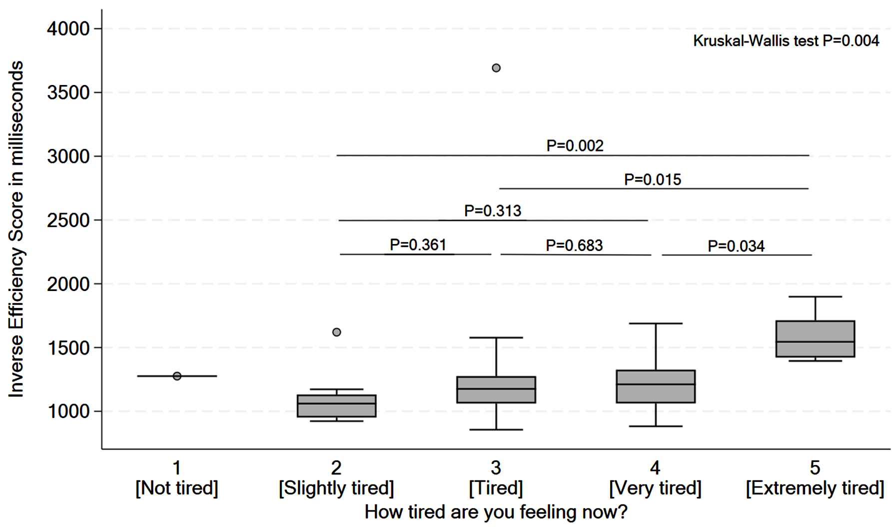
3.3. Association between Subjective Measures of Rest or Workload and SCWT Performance

Work shifts with higher RPE were associated with a higher IES (Kruskal-Wallis test P = 0.004; Figure 3). Work shifts where the anesthetist reported feeling extremely tired were associated with a higher IES compared to work shifts where the anesthetist reported feeling not or slightly tired (P = 0.002), tired (P = 0.015), or very tired (P = 0.034).

Table 4. Cumulative logit mixed model of Relative Perceived Exhaustion (RPE) on anesthetists' work shift type.

Work shift type	Odds ratio (95% CI)	P value
8h-Short Day	1 (ref.)	-
12h-Long Day	6.76 (1.44 to 31.66)	0.015
12h-Night Shift	2.95 (0.82 to 10.57)	0.097
24h-Weekday Call	10.07 (1.24 to 81.68)	0.031
24h-Weekend Call	1.86 (0.16 to 21.76)	0.622

All 63 work shifts had available work shift data and were included in the cumulative logit mixed model. Random intercepts for anesthetists allowed the Relative Perceived Exhaustion (RPE) for the reference work shift (8h-Short Day) to vary across anesthetists.



Inverse Efficiency Score (ms) = (average reaction time in ms)/(1 − proportion of errors), where proportion of errors = (no. of errors)/(total no. of responses).

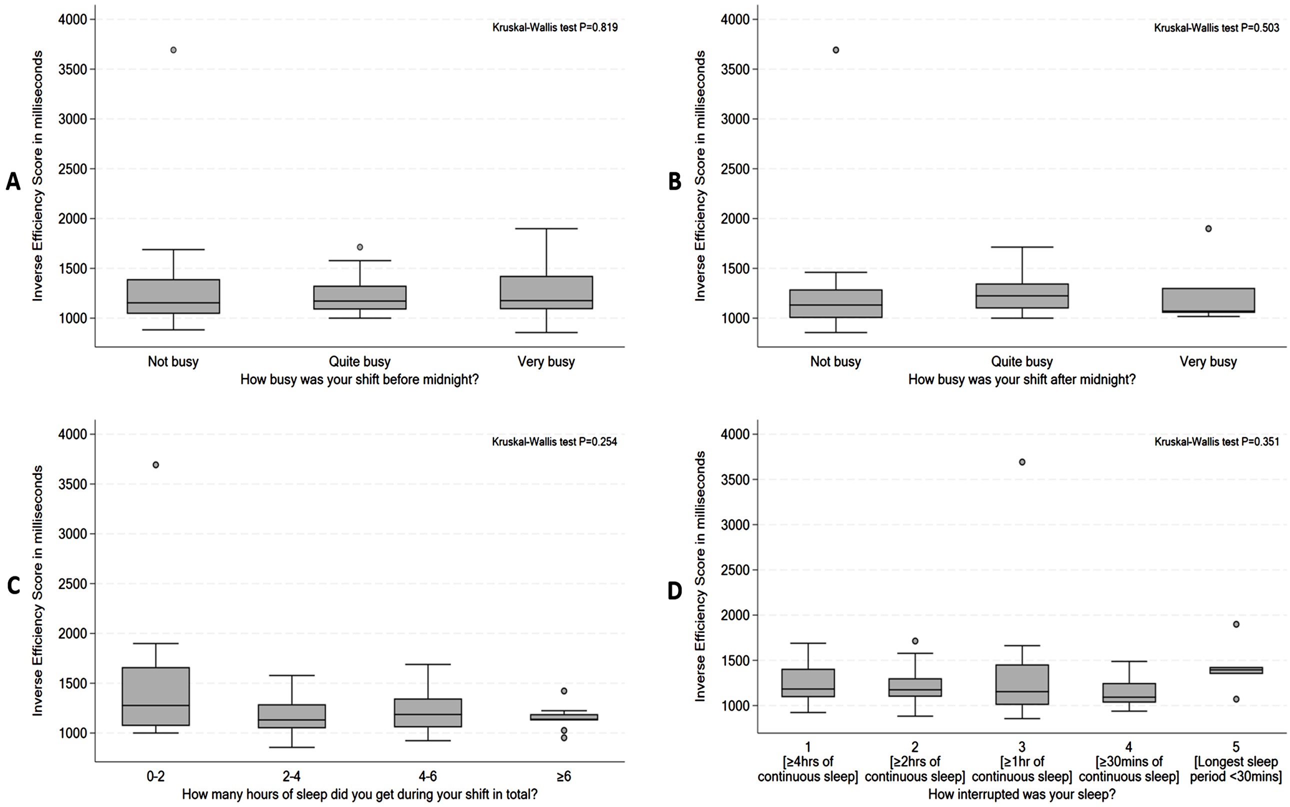
Figure 3. Box plots of inverse efficiency score (IES) by anesthetists' relative perceived exhaustion (RPE) following work shift.

There was no evidence that self-reported busyness of work shifts before (P = 0.819; **Figure 4(A)**) or after midnight (P = 0.503; **Figure 4(B)**), self-reported total hours of sleep an anesthetist had during their shift (P = 0.254; **Figure 4(C)**), or extent of sleep interruption (P = 0.351; **Figure 4(D)**) was associated with the IES.

The self-reported performance level relative to peak performance may be associated with the IES (P = 0.058). Of note, work shifts with the lowest self-reported relative performance level of ≤ 20% had the highest median IES of 1425 ms and the largest variability in scores (IQR 1189 - 1690; **Figure 5**).

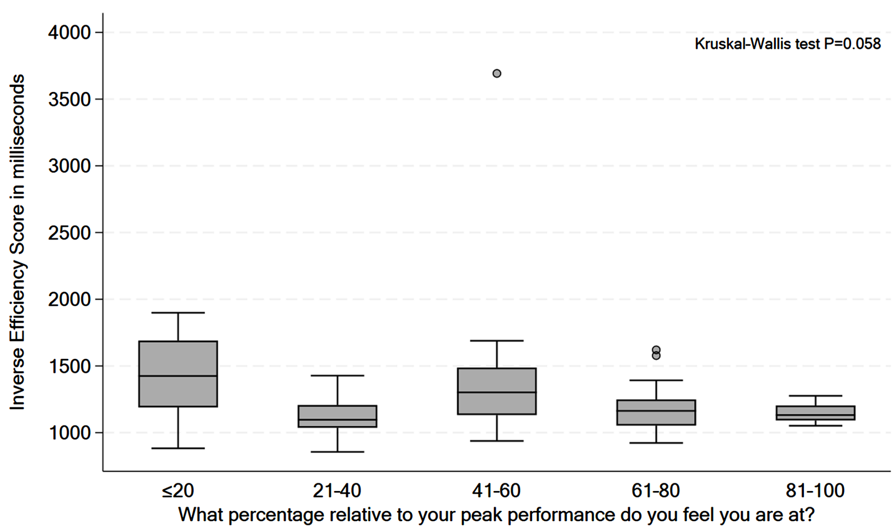
3.4. Sample Size Calculation for the Future SCWT IES Study

In a future definitive trial, to detect a 5% difference in the mean IES (equivalent to 76.2 ms) between a work shift type and an 8h-Short Day with 80% power at a two-tailed 5% significance level, assuming eight SCWTs per anesthetist, at least 34 anesthetists (272 SCWTs) are required (**Figure 6**).



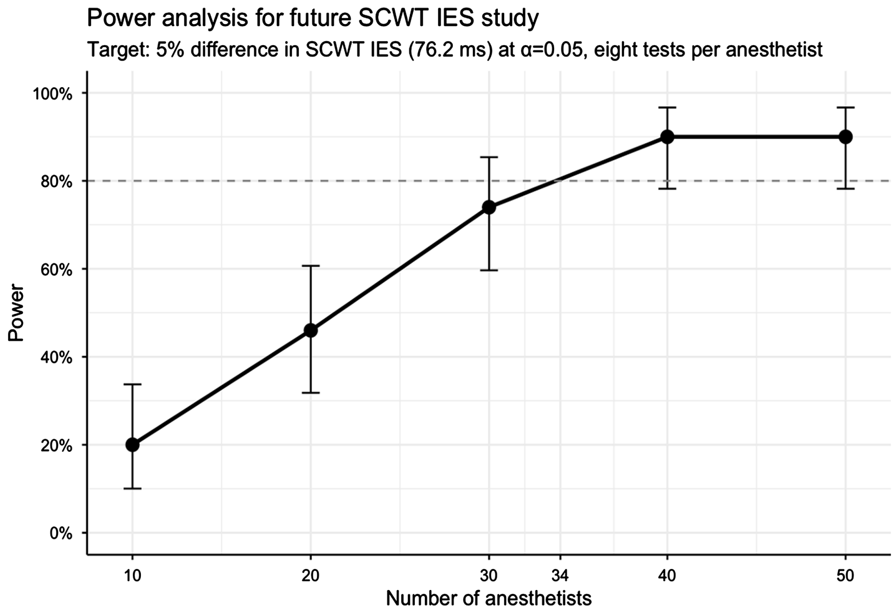
Inverse Efficiency Score (ms) = (average reaction time in ms)/(1 − proportion of errors), where proportion of errors = (no. of errors)/(total no. of responses).

Figure 4. (A) Box plots of inverse efficiency score by anesthetists' self-reported busyness of shift before midnight. (B) Box plots of inverse efficiency score by anesthetists' self-reported busyness of shift after midnight. (C) Box plots of inverse efficiency score by anesthetists' self-reported total hours of sleep during shift. (D) Box plots of inverse efficiency score by anesthetists' self-reported sleep interruption.



Inverse Efficiency Score (ms) = (average reaction time in ms)/(1 − proportion of errors), where proportion of errors = (no. of errors)/(total no. of responses).

Figure 5. Box plots of Inverse Efficiency Score by anesthetists' self-reported relative performance level.



IES = Inverse Efficiency Score. SCWT = Stroop Color and Word Test. Inverse Efficiency Score (ms) = (average reaction time in ms)/(1 − proportion of errors), where proportion of errors = (no. of errors)/(total no. of responses).

Figure 6. Power analysis for a future study investigating the effect of work shift type on SCWT IES.

4. Discussion

Junior doctors are frequently required to work through the biological night and are subject to chronic partial sleep deprivation. A substantial body of literature has documented the adverse effects of prolonged working hours and insufficient sleep on the physical and mental well-being of healthcare professionals, with consequential impacts on patient care [1] [2] [11]. Despite this growing evidence, most departments within Singapore's restructured hospitals continue to implement 24- to 30-hour on-call shifts for junior doctors, likely due to manpower limitations and the logistical complexities of scheduling.

Current guidelines from the Singapore Medical Council (SMC) [23] permit shifts of up to 24 hours, with an additional allowance of up to 6 hours for administrative or educational activities. These recommendations exceed international standards. For example, the Australian and New Zealand College of Anesthetists advises that anesthetists work no longer than 12 hours per shift, with an absolute maximum of 16 hours under exceptional circumstances [24]. Prior local surveys have demonstrated a high prevalence of fatigue and burnout among junior doctors in Singapore [25], yet no study to date has objectively correlated fatigue with functional performance in this population.

To our knowledge, this study is among the first to explore the association between subjective fatigue and objective cognitive performance in junior anesthetists, with the SCWT and IES employed as validated proxies for executive function. Our results indicate that shifts exceeding 12 hours in duration are associated with a measurable decline in objective performance, as reflected in SCWT outcomes. Notably, this impairment occurred regardless of subjective workload intensity or whether rest periods were available during the shift. Furthermore, a positive correlation was observed between subjective fatigue ratings, as measured by the RPE, and cognitive inefficiency (as indicated by the IES), suggesting that perceived fatigue meaningfully influences performance.

Importantly, even when trainees reported shifts as relatively undemanding or having had rest breaks, objective performance was still negatively impacted. This discrepancy raises concerns about the ability of clinicians to self-assess their cognitive limitations, which could have implications for patient safety.

Our findings suggest the need to re-examine existing rostering practices in Singapore, with consideration toward fostering a more sustainable and supportive work environment for junior doctors. More balanced work schedules may have the potential to improve trainee well-being, reduce burnout and attrition, and, in turn, contribute to the quality of patient care. It is noteworthy that the shift structure examined in this study is currently employed by only one restructured hospital's ICU; elsewhere, 24-hour calls or consecutive night float systems remain the norm.

Our study was limited by a small sample size (n = 8), due to the constrained rotation system in a tertiary academic center where only 4 junior doctors are posted to the ICU monthly. Efforts were made to enhance data robustness, including repeated measures collected from each participant to allow within-subject comparisons across different shift types, and short shifts were used as control conditions. Additionally, potential confounders such as the identity of the anesthetist and learning effects were accounted for in the analysis.

Cognitive performance is usually at its peak during two periods of the day (10:00 - 14:00 and 16:00 - 22:00) [26]. Due to the nature of the study, with tests conducted at varied times of the day post-shift, it was not possible to control for circadian rhythm influences and individual chronotype. As cognitive performance and alertness vary across the 24-hour cycle, the differences in IES may partly reflect circadian influences rather than fatigue alone. It is, however, important to note that the majority of the tests following the 12h-Long Day were carried out at around 8 - 10 pm, which, despite being within a typical peak performance time, had poorer SCWT outcomes. This is likely explained by the fact that fatigue has a stronger detrimental effect on cognitive efficiency than circadian-based alertness peaks can compensate for.

Nevertheless, the study was underpowered to detect some differences, particularly between 24-hour shifts and shorter day shifts in terms of RPE and IES, as supported by the leave-one-out sensitivity analysis. This was further compounded by data loss on certain longer shifts such as the 24h-Weekend Call, which could have reflected the excessive fatigue following such shifts, resulting in an inability to complete the study protocol. From our results, the 12h-Long Day had higher IES and RPE scores, and the 24h-Weekday Call had higher RPE. Therefore, it is unlikely that there is not a similar (if not higher) increase in RPE and IES for the 24h-

Weekend Call. This can be addressed by future studies with larger sample sizes.

While participant assignment to shifts was not based on any planned pattern, it was naturally not practical or feasible to randomize the shift allocations due to the presence of scheduling and leave requirements. To address potential confounders arising from roster sequence, unequal distribution of workload, or individual-specific factors such as level of experience and caffeine consumption, propensity score matching could be considered in a future larger study. Given that this was a pilot study, we did not have the sample size to allow for this.

The generalizability of our findings is limited by the single-center design, the small number of trainees, and the specific rostering structure of our institution. Workload intensity, staffing ratios, and rest provisions vary across hospitals, and therefore the observed effects of shift type on cognitive performance may not fully represent other clinical environments. Larger, multi-center studies involving different specialties and hospital systems are warranted to improve external validity and provide a more comprehensive understanding of fatigue in junior medical staff.

Another limitation is the challenge of interpreting whether statistically significant differences in IES translate into clinically meaningful deficits in performance. Although significant associations were observed, the minimal clinically important difference (MCID) for IES in the context of anesthesia practice is not established. Nevertheless, studies have shown that blood alcohol levels corresponding to a legally drunk standard can impair performance on the SCWT [27]; thus, it is conceivable that when performance on the SCWT is demonstrably affected, the degree of cognitive impairment may not be insignificant.

The SCWT is a validated, quick, and practical tool for assessing domains relevant to anesthesia, such as attention, cognitive flexibility, and processing speed. While it would be difficult to translate findings from SCWT performance into objective measures of performance in the anesthesia context, given the highly varied nature of the role, the above domains are highly relevant in anesthesia practice, particularly when handling emergency situations after hours.

We recognize that there are many other important aspects and competencies for good anesthetic care beyond those listed above. The SCWT does not evaluate other crucial competencies like clinical reasoning, psychomotor skills, or memory, all of which are vital in high-stakes environments. To evaluate the clinical impact of fatigue on anesthesia performance, future studies should consider integrating a broader range of assessment tools, including validated fatigue and sleepiness questionnaires (for example, the Epworth Sleepiness Scale), short-term memory tasks, and simulations that assess decision-making and procedural proficiency under fatigue. Future institutional research could also integrate clinical outcomes or near-miss reports with shift scheduling and measured fatigue levels, offering valuable real-world insight into how fatigue affects healthcare performance. Achieving adequate power for such analyses would likely require large sample sizes or extended longitudinal data collection.

5. Conclusion

The study's design demonstrates the feasibility of measuring the effect of extended shifts on cognitive performance objectively (via SCWT/IES) and subjectively (via RPE). The findings suggest an association between extended working hours, increased fatigue, and diminished cognitive performance among anesthesia trainees. As this was a pilot study, the results should be interpreted as preliminary and require confirmation in larger, multi-center studies before being used to inform institutional policy decisions regarding work scheduling. With further validation, such findings may contribute to efforts aimed at improving trainee well-being and patient safety.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Scholliers, A., Cornelis, S., Tosi, M., Opsomer, T., Shaproski, D., Vanlersberghe, C., *et al*. (2023) Impact of Fatigue on Anaesthesia Providers: A Scoping Review. *British Journal of Anaesthesia*, **130**, 622-635. https://doi.org/10.1016/j.bja.2022.12.011

[2] Ippolito, M., Einav, S., Giarratano, A. and Cortegiani, A. (2024) Effects of Fatigue on Anaesthetist Well-Being and Patient Safety: A Narrative Review. *British Journal of Anaesthesia*, **133**, 111-117. https://doi.org/10.1016/j.bja.2024.03.017

[3] Sinha, A., Singh, A. and Tewari, A. (2013) The Fatigued Anesthesiologist: A Threat to Patient Safety? *Journal of Anaesthesiology Clinical Pharmacology*, **29**, 151-159. https://doi.org/10.4103/0970-9185.111657

[4] Gander, P.H., Merry, A., Millar, M.M. and Weller, J. (2000) Hours of Work and Fatigue-Related Error: A Survey of New Zealand Anaesthetists. *Anaesthesia and Intensive Care*, **28**, 178-183. https://doi.org/10.1177/0310057x0002800209

[5] Kayser, K.C., Puig, V.A. and Estepp, J.R. (2022) Predicting and Mitigating Fatigue Effects Due to Sleep Deprivation: A Review. *Frontiers in Neuroscience*, **16**, Article ID: 930280. https://doi.org/10.3389/fnins.2022.930280

[6] Saadat, H., Bissonnette, B., Tumin, D., Raman, V., Rice, J., Barry, N., *et al*. (2017) Effects of Partial Sleep Deprivation on Reaction Time in Anesthesiologists. *Pediatric Anesthesia*, **27**, 358-362. https://doi.org/10.1111/pan.13035

[7] Murray, D. and Dodds, C. (2003) The Effect of Sleep Disruption on Performance of Anaesthetists—A Pilot Study. *Anaesthesia*, **58**, 520-525. https://doi.org/10.1046/j.1365-2044.2003.03131.x

[8] Ramier, M., Clavier, T., Allard, E., Lambert, M., Dureuil, B. and Compère, V. (2024) Examining the Impact of Sleep Deprivation on Medical Reasoning's Performance among Anaesthesiology Residents and Doctors: A Prospective Study. *BMC Anesthesiology*, **24**, Article No. 356. https://doi.org/10.1186/s12871-024-02712-5

[9] Arzalier-Daret, S., Buléon, C., Bocca, M., Denise, P., Gérard, J. and Hanouz, J. (2018) Effect of Sleep Deprivation after a Night Shift Duty on Simulated Crisis Management by Residents in Anaesthesia. A Randomised Crossover Study. *Anaesthesia Critical Care & Pain Medicine*, **37**, 161-166. https://doi.org/10.1016/j.accpm.2017.05.010

[10] Neuschwander, A., Job, A., Younes, A., Mignon, A., Delgoulet, C., Cabon, P., *et al*. (2017) Impact of Sleep Deprivation on Anaesthesia Residents' Non-Technical Skills:

A Pilot Simulation-Based Prospective Randomized Trial. *British Journal of Anaesthesia*, **119**, 125-131. https://doi.org/10.1093/bja/aex155

[11] Sutherland, C., Smallwood, A., Wootten, T. and Redfern, N. (2023) Fatigue and Its Impact on Performance and Health. *British Journal of Hospital Medicine*, **84**, 1-8. https://doi.org/10.12968/hmed.2022.0548

[12] Olaganathan, R., Holt, T.B., Luedtke, J. and Bowen, B.D. (2021) Fatigue and Its Management in the Aviation Industry, with Special Reference to Pilots. *Journal of Aviation Technology and Engineering*, **10**, Article 45. https://doi.org/10.7771/2159-6670.1208

[13] Scarpina, F. and Tagini, S. (2017) The Stroop Color and Word Test. *Frontiers in Psychology*, **8**, Article ID: 557. https://doi.org/10.3389/fpsyg.2017.00557

[14] Stroop, J.R. (1935) Studies of Interference in Serial Verbal Reactions. *Journal of Experimental Psychology*, **18**, 643-662. https://doi.org/10.1037/h0054651

[15] Mullins, K.M. and Reynolds, A.M. (2020) 0123 Sleep Duration and Cognitive Performance on the Stroop Color-Word Task and Simple Reaction Time Task. *Sleep*, **43**, A48-A49. https://doi.org/10.1093/sleep/zsaa056.121

[16] Khullar, S., Singh, P. and Kaur, J. (2020) Effects of Sleep Deprivation on Stroop Performance in Young Adults. *Indian Journal of Physiology and Pharmacology*, **64**, 123-9.

[17] You, Y., Liu, J., Li, X., Wang, P., Liu, R. and Ma, X. (2024) Relationship between Accelerometer-Measured Sleep Duration and Stroop Performance: A Functional Near-Infrared Spectroscopy Study among Young Adults. *PeerJ*, **12**, e17057. https://doi.org/10.7717/peerj.17057

[18] Harrison, B.J., Shaw, M., Yücel, M., Purcell, R., Brewer, W.J., Strother, S.C., *et al*. (2005) Functional Connectivity during Stroop Task Performance. *NeuroImage*, **24**, 181-191. https://doi.org/10.1016/j.neuroimage.2004.08.033

[19] (2023) Stroop Effect Test. https://psycho-tests.com/test/stroop-test

[20] Townsend, J.T. and Ashby, F.G. (1983) Stochastic Modeling of Elementary Psychological Processes. Cambridge University Press.

[21] Bruyer, R. and Brysbaert, M. (2011) Combining Speed and Accuracy in Cognitive Psychology: Is the Inverse Efficiency Score (IES) a Better Dependent Variable than the Mean Reaction Time (RT) and the Percentage of Errors (PE)? *Psychologica Belgica*, **51**, 5-13. https://doi.org/10.5334/pb-51-1-5

[22] Statsenko, Y., Habuza, T., Gorkom, K.N., Zaki, N. and Almansoori, T.M. (2020) Applying the Inverse Efficiency Score to Visual-Motor Task for Studying Speed-Accuracy Performance While Aging. *Frontiers in Aging Neuroscience*, **12**, Article ID: 574401. https://doi.org/10.3389/fnagi.2020.574401

[23] Ministry of Health Singapore (2015) Working Hours for Housemen on Night Shifts (Internet). Ministry of Health. https://www.moh.gov.sg/newsroom/working-hours-for-housemen-on-night-shifts

[24] Australian and New Zealand College of Anaesthetists (2020) PG43(A) Guideline on Fatigue Risk Management in Anaesthesia Practice (Internet). ANZCA https://www.anzca.edu.au/getContentAsset/8b589f42-3b0b-40c5-9b6d-f32379652603/80feb437-d24d-46b8-a858-4a2a28b9b970/PG43(A)-Guideline-on-fatigue-risk-management-in-anaesthesia-practice-2020.PDF

[25] Tan, K.H., Lim, B.L., Foo, Z., Tang, J.Y., Sim, M., Lee, P.T., *et al*. (2022) Prevalence of Burnout among Healthcare Professionals in Singapore. *Annals of the Academy of Medicine, Singapore*, **51**, 409-416.

https://doi.org/10.47102/annals-acadmedsg.2021338

[26] Valdez, P., Ramírez, and García, A. (2012) Circadian Rhythms in Cognitive Performance: Implications for Neuropsychological Assessment. *ChronoPhysiology and Therapy*, **2**, 81-92. https://doi.org/10.2147/cpt.s32586

[27] Riedel, P., Wolff, M., Spreer, M., Petzold, J., Plawecki, M.H., Goschke, T., *et al*. (2021) Acute Alcohol Does Not Impair Attentional Inhibition as Measured with Stroop Interference Scores but Impairs Stroop Performance. *Psychopharmacology*, **238**, 1593-1607. https://doi.org/10.1007/s00213-021-05792-0