Exploring the operator experience in automated shuttles: Fatigue, attention, and gaze behaviour

Citation for the original published paper (version of record):

Ahlström, C., Weidel, M., Sjörs, A. et al (2025). Exploring the operator experience in automated shuttles: Fatigue, attention, and gaze behaviour. Transportation Research Interdisciplinary Perspectives, 29. http://dx.doi.org/10.1016/j.trip.2025.101332

Transportation Research Interdisciplinary Perspectives 29 (2025) 101332

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives

Exploring the operator experience in automated shuttles: Fatigue, attention, and gaze behaviour

Christer Ahlström [a,b,*], My Weidel [a,c], Anna Sjörs Dahlman [a,d], Ashleigh Filtness [e], Anna Anund [a,f]

[a] Swedish National Road and Transport Research Institute, 581 95 Linköping, Sweden
[b] Department of Biomedical Engineering, Linköping University, 581 83 Linköping, Sweden
[c] Department of Behavioural Sciences and Learning, Linköping University, 581 83 Linköping, Sweden
[d] Department of Electrical Engineering and SAFER, Chalmers University of Technology, 412 96 Gothenburg, Sweden
[e] Transport Safety Research Centre, Loughborough University, Loughborough, Leicestershire, LE11 3TU, UK
[f] Rehabilitation Medicine, Linköping University, 581 83 Linköping, Sweden

ARTICLE INFO

Keywords:
Autonomous shuttle
Professional driver
Driver fatigue
Driver sleepiness
Situation awareness
Driver vigilance

ABSTRACT

Automated shuttles provide a first look at future transportation, however, they still require on-board human operators for regulatory compliance and safety assurance. This paper examines the experiences of eight shuttle operators through two studies. The aim was to investigate their alertness throughout their working shifts. Study A used a controlled experimental methodology to compare fatigue and gaze behaviour at complex road sections during the first and last hours of one shift. Study B involved naturalistic observations over two months, examining sleep, sleepiness and stress.

Study B found that 27% of work shifts occurred following less than six hours of sleep. However, only 1% of shifts resulted in a Karolinska Sleepiness Score of 7 or higher, suggesting that insufficient sleep was rare. Stress was also infrequently reported. Notable individual differences suggested the potential value of personalized approaches to fatigue management. Study A revealed that while overall alertness was generally adequate, gaze patterns often deviated from safe expectations. Operators paid less attention to their surroundings than would be expected (21% not looking left, 38% not looking right, 58% not looking to the rear of the vehicle, in situations where this would have been appropriate).

The results are important for safety operators and their employers, highlighting the shared responsibility of having well-prepared and well-rested operators who are fit to effectively monitor the automated shuttle for an entire driving period. Further research is needed to develop effective strategies to maintain operators' situational awareness over time, especially as their confidence in the vehicles' capabilities increases.

1. Introduction

Automated vehicles, electrification, and ridesharing have the potential to improve road safety, reduce traffic congestion, lower emissions, reduce the need for parking spaces, and promote sustainable urban mobility. Automated shuttles (Fig. 1) offer a peek into what the future of automated mobility may hold. These shuttles aim to enhance accessibility for everyone, including individuals with disabilities, the elderly, and those without access to private vehicles (Iclodean et al., 2020). They also address the last-mile problem in urban environments with Bus Rapid Transit-solutions as well as cases where narrow streets are not easily served by traditional buses (Bucchiarone et al., 2021; Thorhauge et al., 2022; Whitmore et al., 2022). The use of automated shuttles is expanding rapidly as cities seek innovative transportation solutions and technology advances. Deployment projects are emerging globally in forms of demonstrations, pilots and showcases, including in Europe (e.g., Sweden, Norway, France, Germany, and Spain), North America (e.g., USA and Canada), Asia (e.g., Singapore, Japan, and South Korea), and Australia.

The technological maturity of automated shuttles is continuously improving. However, further development is necessary to ensure smooth operations in complex traffic situations, considering surrounding traffic,

---

* Corresponding author at: VTI, Olaus Magnus väg 35, 581 95 Linköping, Sweden.
E-mail address: christer.ahlstrom@vti.se (C. Ahlström).

https://doi.org/10.1016/j.trip.2025.101332
Received 20 November 2024; Received in revised form 8 January 2025; Accepted 12 January 2025
Available online 23 January 2025

Fig. 1. Shared automated shuttle bus.

shared spaces, sensor noise, and weather conditions with precipitation (Anund et al., 2022). Snow, dust, rain, and leaves often interfere with surround sensing which triggers emergency braking manoeuvres, jeopardizing the safety of passengers who risk falling (Backhaus, 2020). For safety reasons, shuttles are therefore (i) driving at low speeds, and (ii), being monitored by an on-board operator who is ready to take over control whenever needed.

In addition to their technical capabilities, the success of automated shuttle services depends on passengers' willingness to use them (Piatkowski, 2021). The perceived safety when traveling by automated shuttles is generally high (Hilgarter & Granig, 2020; Mouratidis & Serrano, 2021; Salonen & Haavisto, 2019). Flexibility, efficiency, number of departures, and appropriate transportation connections are key aspects if people are to use shuttles instead of other modes of transportation (Hilgarter & Granig, 2020; Mouratidis & Serrano, 2021). These expectations conflict with today's shuttle buses' state of technology, which may lead to negative user experiences. Especially the low vehicle speed (Schuß et al., 2022), about 13 km/h at our test site, is problematic as users expect the shuttle services to save them time. Ongoing research is currently investigating on-demand functionalities to accommodate some of the expectations. However, to overcome the users' demands for higher speed and fewer abrupt braking events (Mouratidis & Serrano, 2021), the underlying automation technology must first be improved.

While much research has been devoted to automation capability and passenger experience, the role of safety operators remains under-researched. Safety operators have multiple roles in operation of the shuttle. They support passengers, check that the vehicle functions properly, maintain the safety of passengers inside the vehicle as well as surrounding road users outside. They also monitor the automated systems and intervene in case of automation failures (Schrank et al., 2024; Schuß et al., 2022; Sherry et al., 2020). Monitoring and intervention for hazardous rare events is a straining task that human operators are not well-suited to perform (Greenlee et al., 2019; Solís-Marcos et al., 2018; Stapel et al., 2019; Warm et al., 2008). Extensive fatigue risk management programmes have therefore been developed to prevent, monitor, and mitigate fatigue-induced risks among safety operators in automated driving systems (Favaro et al., 2022). These programmes include components dealing with continued education, awareness and reporting guidelines, real-time vigilance assessment, supplemental engagement, and adaptive scheduling. There are no technical quick fixes that makes safety operators more vigilant. The full range of components is needed, something which requires trust and collaboration between the safety operators and their employers.

The general aim of this study was to investigate work-related fatigue, stress, and inattention among automated shuttle operators. Two experiments were conducted: Study A was a controlled field study investigating sleepiness and attention during one work shift, and Study B was a long-term naturalistic study where sleep patterns and fatigue/stress levels were investigated over a period of two months. Together, the two studies provide an overview of the working conditions (in terms of fatigue, attention, and stress) faced by automated shuttle operators in their daily work.

2. Methods

The study was conducted at the test site "Ride the Future" in Linköping, Sweden (https://www.ridethefuture.se). This research platform aims to explore how small electrified and automated shuttles can complement existing public transportation, thereby increasing its competitiveness against privately owned cars and making it easier for more people to use public transportation. Three shuttles operate on a 3.7-km route that crosses a university campus, which has many pedestrians and bicyclists, and a nearby suburb with mixed traffic and narrow streets. A safety operator is on board for each trip, responsible for monitoring the vehicle's driving and taking over manual control if necessary. The safety operator also acts as a customer host.

The vehicles utilized during the data collection are of an early generation and lack advanced technological solutions. Due to the presence of protected objects along the vehicles' routes, the cameras are disabled. The navigation sensors are LiDAR and an advanced GPS system. The safety of individuals outside the shuttles is prioritized, which can result in sudden braking events that may cause discomfort for the safety operator and the passengers.

This paper reports on two separate data collections. Study A was conducted in May 2021 and focused on the safety operators' sleepiness and attention levels during one work shift. Study B, conducted between August and October 2023, examined the drivers' fatigue and stress levels over an extended period. While study A could make use of obtrusive sensors such as electrodes and eye-tracking glasses, study B, that lasted for two months, was conducted with less obtrusive wearable devices and before/after shift questionnaires and tests.

2.1. Participants

Eight drivers operate the shuttles at the test site and all of them were included in the two studies. All eight participants have, in addition to a bus driver's license including a professional competence training certificate, also received specific training to work as shuttle safety operators. The two vehicle suppliers (EasyMile and Navya) require this specific training, which have been developed by the manufacturers themselves. The training lasts four days and includes both theory and practice. The test site has been operational since March 2020 and seven of the drivers have been working at the site throughout this period. One driver was replaced in 2022, meaning that the participant population is close to, but not identical, in the two studies.

The participants worked full time, with 60 % of their working days as shuttle operators. For the remaining time, they drive a city bus or tram in a nearby city. When working in the shuttle, unlike driving a city bus or tram, they can take short breaks as needed. The shuttle operation is a special shift, and there are no general guidelines on how many days per week or in a row they must work in the shuttles. Their working day begins at the depot in the other city, and they commute to the shuttle site. The commuting time, approximately 45 min, is included in their working hours.

Population characteristics for the participants in the two studies are reported in Table 1. The body mass index is lower in study B, which is not only due to the replaced driver, but also due to a general decrease in body weight over the course of the two studies. The participants' self-reported sleep hygiene was generally good, but two participants in both study A and study B felt that they did not get enough sleep, and one participant in study B reported clearly insufficient sleep. The Bordeaux

Table 1
Participant characteristics.

	Study A (controlled)	Study B (naturalistic)
Gender	7 men, 1 woman	6 men, 2 women
Age (M ± SD, range years)	47.8 ± 8.2 [30 – 58]	47.8 ± 7.7 [31 – 59]
BMI (M ± SD, range kg/m^2)	28.8 ± 4.3 [21.8 – 34.7]	25.0 ± 2.8 [19.7 – 29.3]
Years as a bus driver (M ± SD, range years)	16.6 ± 7.3 [6 – 30]	14.8 ± 2.8 [3 – 24]
Do you feel stressed (tense, restless, nervous or anxious)? (1 Not at all – 5 Very stressed)	1.4 ± 0.7 [1 – 3]	2.1 ± 1.0 [1 – 4]
How many hours do you usually sleep per 24 h period? (M ± SD, range hours)	7.1 ± 0.6 [6 – 8]	7.1 ± 0.7 [6 – 8]
Are you getting enough sleep? (1 Yes, definitely – 5 No, far from enough)	2.2 ± 0.4 [2 – 3]	2.4 ± 0.9 [1 – 4]
In general how would you like to rate your sleep? (1 Very good – 5 Very bad)	2.0 ± 0.5 [1 – 3]	2.0 ± 0.5 [1 – 3]
Bordeaux Sleepiness Scale (combined score 0 – 8)		2.4 ± 1.1 [1 – 5]

Sleepiness Scale (Philip et al., 2023) was used to evaluate sleep-related driving risk in study B. If the total score is 3 or higher, the risk of having a sleep-related near-miss or crash is defined as positive. One participant scored 3 and one participant scored 5. The elevated risk score originated from the question “Have you experienced in the previous year at least one episode of severe sleepiness at the wheel that made driving difficult or forced you to stop?”, where both participants answered “Yes, at least once a month”. The question was asked in the context of their role as professional drivers, encompassing both city/tram driving and shuttle operations.

2.2. Design and procedure

2.2.1. Study A

The data collection in study A was conducted during one shift, within normal operation during normal work shifts. An afternoon shift begins after lunch and lasts approximately eight hours (including commute time). Data collection in study A covers the first and last hour of the afternoon work shift. Since the shuttles can complete about three laps on their route in one hour, a total of 2 × 3 = 6 laps are included in the analysis. The first three laps occur sometime between 1:30 PM and 3:00 PM, collectively these laps will be referred to as “Drive 1”, and the last three laps occur between 4:30 PM and 6:00 PM, collectively these laps will be referred to as “Drive 2”. The instructions were to drive as they normally would, except that they were not allowed to take any breaks during the three laps, instead being required to take a break between each set of three laps. Study A had a within-subjects design with a factor for the first versus second drive.

Before participating in study A, the drivers kept a diary for three days to track their sleep and wakefulness as well as their working hours. The night before data collection, the drivers had slept between 5 and 9 h. During the three preceding days, they had slept between 4 h and 15 min and 9 h. When arriving to the laboratory, the participants received instructions and signed an informed consent form. Electrodes for the physiological measurements were attached and eye tracking glasses were mounted and calibrated.

2.2.2. Study B

The data collection in Study B was designed as a naturalistic driving study and conducted over two months (62 days) in real-life operations. Participants were informed about the study during an introductory meeting with all eight shuttle drivers. After receiving information about the study, participants signed an informed consent form and filled in a background questionnaire. They were also provided with a wearable sleep and activity tracker to be worn at all times for the entire data collection period.

Throughout the data collection period, participants reported their subjective sleepiness, stress, and fitness levels on a daily basis. The Karolinska Sleepiness Scale (KSS) was used for sleepiness ratings, the VTI Acute Stress Scale (VSS; Sjörs Dahlman et al., 2024) was used for stress ratings, and general fitness was rated on a scale from 0 to 100. All participants performed an alcohol test before each shift (Senseair Wall, Västerås, Sweden) and underwent two random drug tests for benzodiazepines and opioids (Leitat, Barcelona, Spain).

Study B is part of the EU funded project PANACEA. The overall aim of the project is to create a holistic monitoring and assessment system which detects professional drivers who are not fit to drive and support them and their employers to manage the situation and put measures in place to prevent it happening again. This paper includes the baseline data collection done before implementing the fitness-to-drive support system.

2.3. Measurements

2.3.1. Study A

Every fifth minute during the drives the participants rated their sleepiness level on the Karolinska Sleepiness Scale (KSS; Åkerstedt & Gillberg, 1990). The scale has nine anchored levels: (1) extremely alert (2) very alert (3) alert (4) rather alert (5) neither alert nor sleepy (6) some signs of sleepiness (7) sleepy, no effort to stay awake (8) sleepy, some effort to stay awake; and (9) very sleepy, great effort to keep awake, fighting sleep. Participants were instructed to report a value corresponding to their average feeling over the past 5 min.

An electrocardiogram (ECG, lead II) and an electrooculogram (EOG, electrodes placed vertically across the left eye) were recorded at a sample rate of 256 Hz using a Vitaport 3 bio-amplifier (Temec Instruments BV, the Netherlands). The EOG was bandpass filtered between 0.3 and 11.5 Hz, while the ECG was filtered between 0.3 and 30 Hz. Both filters were zero-phase 3rd order band-pass Butterworth filters. Heart rate, heart rate variability, and blink behaviour were extracted from the ECG and EOG. Increasing levels of sleepiness is expected to decrease heart rate, increase heart rate variability (Lu et al., 2022; Persson et al., 2020), and lengthen blink durations (Schleicher et al., 2008; Wierwille & Ellsworth, 1994). *Heart beats* (R-peaks) were detected in the ECG using the filter bank approach by Afonso et al. (1999), and outliers were removed by filtering out RR-intervals deviating more than 3 standard deviations from the mean (Aubert et al., 1999). Additionally, the root mean square of successive differences (RMSSD) was calculated according to Shaffer and Ginsberg (2017).Blink duration was extracted from the EOG using the filtering and differentiation scheme developed by Jammes et al. (2008). The mean heart rate and blink duration were calculated in five-minute windows to match the KSS ratings. Similarly, RMSSD was calculated in the same five-minute windows.

The participants situational awareness was assessed with camera-based eye tracking glasses (Pupil Invisible, Pupil Labs GmbH, Germany). The glasses record a scene video with a gaze overlay. The videos were analysed manually by encoding each glance based on its direction and relevance. The directions or objects coded were forward, backward, right, left, the bus interface (GUI), mobile phone, and other. Each glance was also coded as necessary, useful, or irrelevant for driving based on the Minimum Required Attention framework (Ahlström et al., 2021a; Kircher & Ahlstrom, 2017).

2.3.2. Study B

In Study B, the KSS was used to rate sleepiness before and after each shift. A KSS score of 7 or higher was used to identify cases of sleepiness. In addition, stress was rated using the VSS, which has nine levels matching the KSS levels: (1) completely relaxed, feeling entirely calm and relaxed (2) very relaxed (3) relaxed (4) rather relaxed (5) neither relaxed nor stressed (6) slightly stressed, (7) stressed, feeling some tension and pressure (8) very stressed (9) extremely stressed, feeling very tense and under high pressure, on the verge of what I can handle. A VSS score of 7 or higher indicated stress. After each shift, participants also answered the question, "How was your fitness to drive today?" on a scale from 0 to 100 %, with ratings below 50 % indicating they felt unfit to drive.

Objective measures of sleep quality and stress were obtained from a wrist-worn wearable device (Fitbit Charge 5, San Francisco, US). The device provides a stress score, ranging from 0 to 100, which is calculated daily based on heart rate variability, exertion levels, and sleep patterns. Higher scores indicate fewer physical signs of stress. A score of 60 or below was used to identify stress in the data. The Fitbit also tracks sleep duration and provides a sleep score, which ranges from 0 to 100. This score is based on sleep duration, sleep quality, and heart rate variability, with higher scores indicating better sleep quality. Sleep durations below 6 h (Cirelli et al., 2017) and Fitbit sleep scores below 60 (https://enterprise.fitbit.com/blog/track-sleep/) were used to identify insufficient sleep.

Due to the naturalistic design, participants were not asked for their KSS ratings during drives as this would have been too obtrusive. Instead, sleepiness levels were estimated using a biomathematical fatigue model (BMM). The model is based on the sleep/wake predictor model by Åkerstedt et al. (2008) but further developed to incorporate time on task (Ahlström et al., 2021b). The BMM predicts fatigue levels on a continuous scale that matches the 9-level KSS scale by considering time of day (circadian rhythm), time awake (sleep homeostasis, from the wearable device), and time on task (from participants' shift schedules). Sleep and shift schedule information from the past three days are used as input to the BMM. The maximum BMM value in each working shift, along with a threshold of 7, was used to identify shifts with sleepiness.

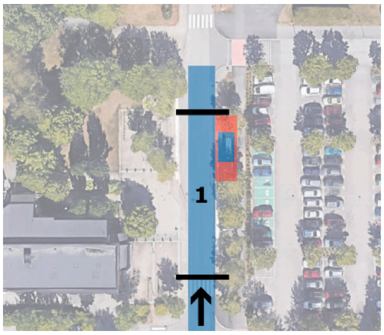
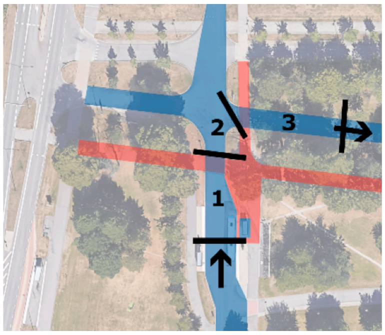
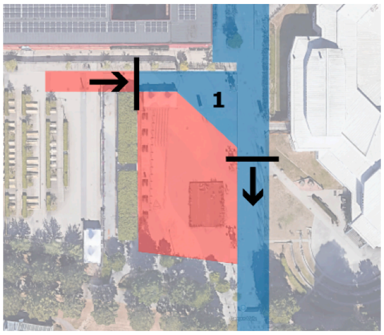
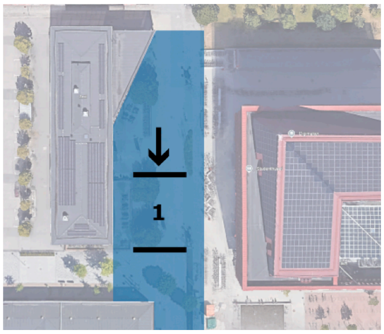
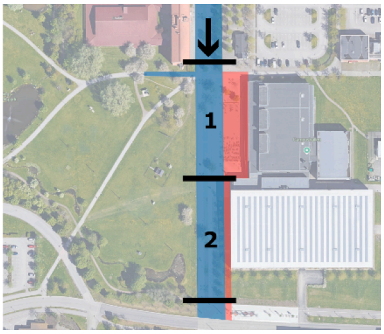
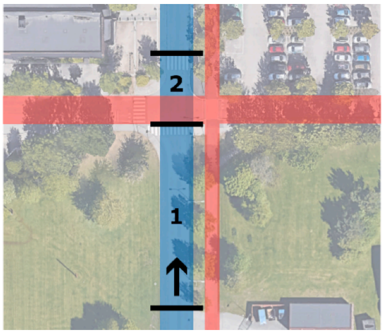
2.4. *Statistical analyses*

The sleepiness indicators in study A were analysed with separate analyses of covariance (ANCOVA) with a random effect. Both the input data (time on task, 5 – 55 min in 5-minute intervals) and the response variable (sleepiness indicator) were treated as continuous variables whereas the grouping variable (Drive 1 versus Drive 2) was treated as a categorical fixed factor. Participant (1–8) was included as a random factor to account for inter-individual variations. The model assumes a linear relationship between time on task and the sleepiness indicators.

A descriptive approach was used to analyse the participants' situational awareness. Data from the eye tracker was categorised according

Table 2

Visualisation and specification of locations/zones and where it is deemed necessary (blue areas) or useful (red areas) to look to ensure sufficient situational awareness. Black lines indicate start and end of the zones. The arrows represent the direction of travel. Map data was obtained from Google Earth (Google LLC).

Location 1	Location 2	Location 3
		
Zone 1: Dock at bus stop • Necessary: Forward, Backward, Right (at the bus stop) • Useful: Area around the bus stop	Zone 1: Exit from bus stop, zebra crossing • Necessary: Forward, Backward, Right (at the bus stop) • Useful: Right/Left (bicycle path) • Zone 2: Yield obligation, turn right • Necessary: Forward, right, left • Useful: Forward (bicycle path) • Zone 3: Turn right, bicycle lane • Necessary: Forward, backward • Useful: Right/Left (bicycle path)	Zone 1: Crossing shared space, enter thoroughfare for pedestrians and cyclists • Necessary: Forward, Right, Left • Useful: Right, Backward
Location 4	Location 5	Location 6
		
Zone 1: Exit from bus stop in shared space • Necessary: Forward, Backward, Right, Left	Zone 1: Exit from bus stop, yield obligation • Necessary: Forward, Backward, Right (bicycle path) • Useful: Left • Zone 2: Dock at bus stop • Necessary: Forward, Backward, Right (bus stop)	Zone 1: Go straight, zebra crossing • Necessary: Forward (including zebra crossing), Backward • Useful: Right, Left • Zone 2: Intersection (no yield obligation), zebra crossing • Necessary: Forward, Backward • Useful: Right, Left

to glance direction/target and whether the glance was deemed necessary, useful, or irrelevant. The analysis includes not only objects the participants looked at but also the proportion of necessary directions/objects the driver did not look at. This was done in six geographical locations of particular interest along the route. The locations include both conventional bus stops (on the sidewalk, with a glass shelter) and unconventional bus stops (bus stop with only a sign). For some locations, the geographical area has been divided into zones due to changes in the vehicle's direction and the attentional requirements in the area. For each location/zone it was then predefined *which* directions that needed to be scanned, and *when* they needed to be scanned, to ensure sufficient situational awareness. For example, in cases where the shuttle is obligated to give way, it is necessary for the driver to check all directions from which other road users may come. Such directions are classified as *necessary* for safe driving. Yield obligation means the shuttle stops automatically before the intersection if the driver does not confirm that nothing is in the way. For the EasyMile shuttle, yield obligation means the bus stops before the driver can approve the exit, while the Navya only stops if any object is detected in the surroundings of the shuttle. In addition to necessary attentional requirements, there are also "requirements" that are classified as *useful*. These useful areas of interest are not required by law, but can facilitate a smooth ride, and may also increase the safety of surrounding road users. Details about their locations and how they were analysed are found in Table 2. It should be noted that glance behaviour is just a proxy of situational awareness.

Study B aimed to determine the prevalence of insufficient sleep before work shifts and the frequency of sleepiness and stress during shifts. The analyses were exploratory and descriptive, focusing on the distribution of these variables and counting occurrences of insufficient sleep, sleepiness, and stress. Linear regression models were used to identify trends throughout the study.

3. Results

3.1. Results study A

The development of fatigue throughout the study sessions is illustrated in Fig. 2. Subjective sleepiness is at a low level indicating that most participants are alert throughout. The highest reported value was KSS 7 (sleepy but not difficult to stay awake) and the lowest was 2 (very alert). There was a significant increase in KSS with time on task (slope 0.02 KSS units/min), but no significant difference was found between Drive 1 and Drive 2 (Table 3). Similarly, there was an increase in blink durations over time in both Drive 1 and Drive 2 (slope 0.25 ms/min), indicting the participants became more fatigued over time. Blink durations were slightly longer in Drive 2. A time on task effect was also seen for heart rate, with decreasing heart rate over time in both Drive 1 and Drive 2 (slope −0.08 bpm/min). This could not be seen in the heart rate variability metric RMSSD. Clear individual differences were seen in all

Table 3
F-values and degrees of freedom from mixed-model ANCOVAs for drive 1/2 and time on task (5 – 55 min in 5-minute intervals). Participant is included as a random factor. Significant differences at the 0.05 level (0.0125 after Bonferroni correction) are marked with a *.

	Time on task (5–55 min)	Drive [1,2]	Participant	Drive*Time on task
KSS	F(1,161) = 45.8, p < 0.001*	F(1,161) = 2.7, p = 0.10	F(7,161) = 47.3, p < 0.001*	F(1,161) = 2.9, p = 0.09
Blink duration	F(1,144) = 12.3, p < 0.001*	F(1,144) = 9.7, p = 0.002*	F(6,144) = 70.3, p < 0.001*	F(1,144) = 5.6, p = 0.02
Heart rate	F(1,154) = 18.4, p < 0.001*	F(1,154) = 3.5, p = 0.06	F(7,154) = 70.2, p < 0.001*	F(1,154) = 4.8, p = 0.03
RMSSD	F(1,154) = 0.5, p = 0.48	F(1,154) = 3.2, p = 0.07	F(7,154) = 43.7, p < 0.001*	F(1,154) = 0.4, p = 0.52

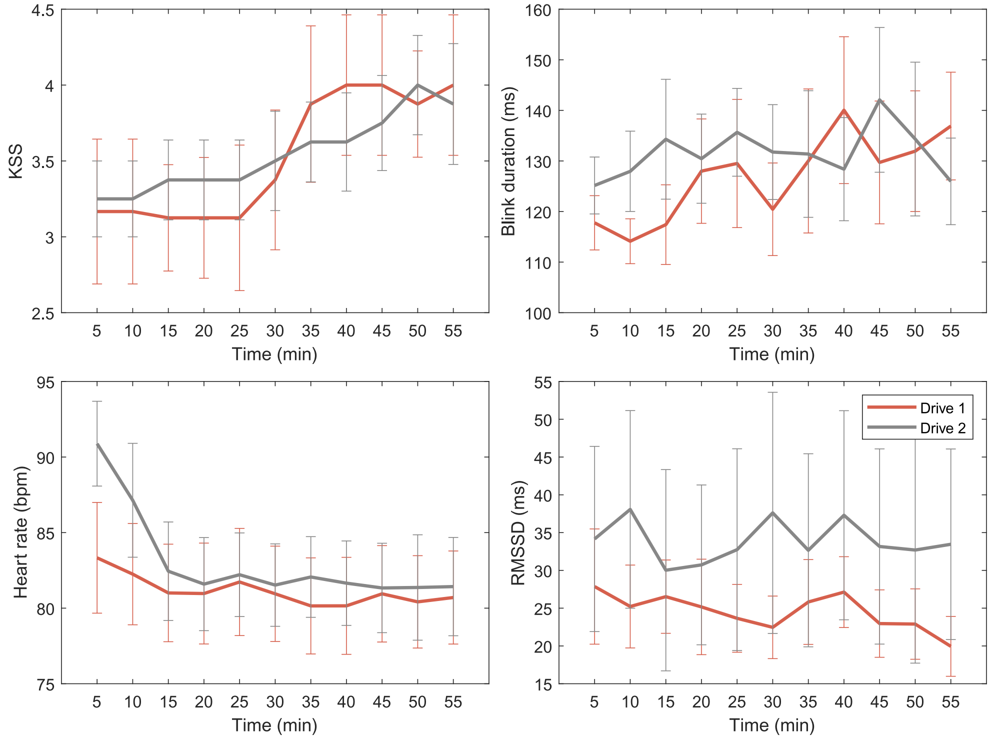


Fig. 2. Fatigue indicators as a function of time on task. The error bars represent the standard error of mean.

variables.

The participants directed 39 % of their glances in the forward direction, corresponding to 43 % of the time, see Fig. 3. Approximately 9 % of the glances were directed backward, 19 % to the left, and about 18 % to the right. The drivers also looked at their mobile phone for 3 % of the time, at the shuttle interface (GUI) for 5 % of the time, and at other things for 9 % of the time. The drivers distributed their glances in almost the same way during Drive 1 and Drive 2.

Table 4 shows the percentage of time the participants looked in different directions, subdivided according to the locations and zones defined in Table 2. The participants glanced extensively at the shuttle's GUI at location 2, zone 2, for up to 10 % of the time, which may be because the driver uses the interface to acknowledge that the shuttle can continue driving there. It can also be seen that the participants spent about 10 % of their time interacting with their mobile phone at location 1.

The directions, per location and zone, deemed necessary to look in to obtain sufficient situational awareness, have been compared with the participants' glance data, Table 5. Consistently across all locations, participants did not glance backwards in on average 58 ± 24 % of the locations or passages where it was deemed necessary to do so, Table 5. The corresponding percentage of neglected glances to the left and to the right was 21 ± 9 % and 38 ± 15 %, respectively.

3.2. *Results study B*

Fig. 4 presents the Fitbit sleep score, hours slept, KSS ratings before and after shifts, and the maximum BMM value during the workday, of all working days in the shuttle over the two study months. The mean ± std (range) of the Fitbit sleep score was 78.7 ± 6.8 (57 – 97) %, hours slept = 6.7 ± 1.3 (3.0 – 11.8), KSS before = 3.1 ± 1.1 (1 – 7), KSS after = 3.2 ± 1.2 (1 – 7), and maximum BMM = 5.0 ± 0.9 (3.1 – 8.2). Linear regression models indicated no significant changes in sleep or sleepiness metrics during the data collection period. Summary statistics and histograms show that participants were generally well-rested and alert. Counting the number of events indicating insufficient sleep, 0.65 % of the nights before a work shift had a Fitbit sleep score below 60, 27.45 % of the recorded sleep durations were shorter than 6 h, 1.27 % of the KSS ratings before the shift started were rated 7 or higher (1.34 % after the shift), and 3.92 % of the work shifts had a BMM fatigue score above 7.

Fig. 5 illustrates the stress and fitness metrics over time. The subjective fitness ratings increased significantly throughout the study (F-statistic versus constant model = 6.16, p = 0.01). In contrast, the stress metrics remained stable. Histograms reveal that participants generally reported high fitness and low stress levels. The mean ± std (range) rated fitness level = 88.6 ± 13.6 (49 – 100) %, the Fitbit stress score = 78.2 ± 5.8 (65 – 90) %, VSS before shift = 2.9 ± 0.8 (1 – 6), and VSS after shift = 3.0 ± 0.9 (1 – 5). Only once was fitness rated below 50 %, and there were no instances of reported stress.

During the two-month data collection period, two drivers reported drinking alcohol within 12 h before their shift, on three separate occasions. However, none of the participants tested positive for alcohol before starting their shift. One driver reported taking drugs within 12 h before their shift on one occasion. None of the participants tested positive for drugs in the random drug tests.

4. Discussion

This paper provides insights into fatigue, attention and stress experienced by operators of automated shuttles. As an emerging transport mode, little is currently known about operators work experience. By combining focused experimental investigation with naturalistic observation, this study offers a first-time exploration of the unique challenges and demands faced by these operators. Generally, participants reported high fitness, low stress, low subjective sleepiness and high sleep scores. In 27.45 % of the recorded nights, the sleep duration was shorter than 6 h, but this rarely resulted in high levels of sleepiness during their shifts. Thus, the safety operators appear well-prepared for the task of monitoring the automated shuttle. Although, there were notable individual differences in all fatigue-related measures highlighting the personal nature of fatigue experience. A particularly concerning observation was that operators frequently neglected to monitor areas deemed necessary for safe driving, especially those to the rear of the vehicle, and to the left and right when approaching intersections. This behaviour could indicate overreliance on the automated system, potentially leading to gaps in operators' situational awareness.

Shuttle operators are required to maintain situational awareness at all times, with an expectation that they will take over control at a moments notice if it is deemed that there is a risk to safety. Analysing situational awareness based on eye-tracking data has its limitations. The eye tracker measures where you are looking, not what you have seen and comprehended (Mack, 2003; Simons, 2000), nor will it account for peripheral vision (Vater et al., 2022). Results from study A show that participants often did not look left and right in situations where it was deemed necessary to do so (Table 2). There was also a high number of missed backward glances. This could indicate that the participants relied on peripheral vision to cover the sides and mirrors, or that they believed it was unnecessary to monitor the surroundings, trusting the vehicle to do so.

Much attention was directed towards a display in the shuttle that

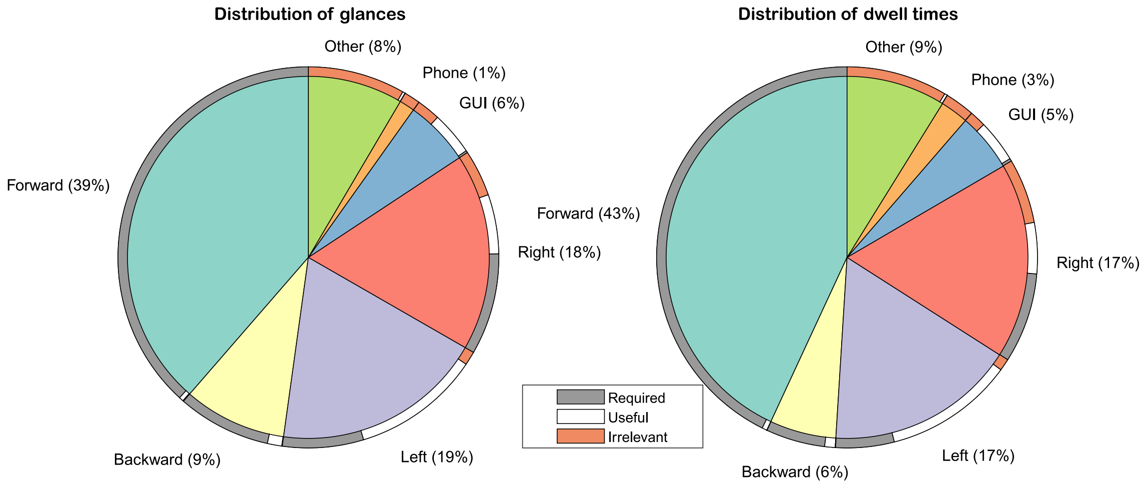


Fig. 3. Pie chart with glance distribution and dwell time distribution towards different glance directions/targets. The outer circle shows the distribution required, useful, or irrelevant glances per direction/target.

Table 4
Percentage of time that the participants looked at different glance directions/targets, subdivided by location, zone and first/second drive. The locations and zones were defined in Table 2.

		Forward		Backward		Left		Right		GUI		Phone		Other	
	Drive	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Location	Zone														
1	1	34 %	39 %	4 %	8 %	8 %	1 %	27 %	27 %	5 %	2 %	11 %	8 %	10 %	14 %
2	1	17 %	27 %	12 %	11 %	27 %	15 %	24 %	22 %	10 %	15 %	1 %	3 %	8 %	8 %
2	2	29 %	28 %	3 %	5 %	17 %	15 %	37 %	39 %	11 %	9 %	1 %	2 %	4 %	3 %
2	3	60 %	44 %	2 %	6 %	8 %	7 %	11 %	21 %	7 %	5 %	0 %	2 %	12 %	15 %
3	1	45 %	52 %	10 %	7 %	20 %	20 %	12 %	10 %	8 %	6 %	2 %	0 %	3 %	4 %
4	1	60 %	47 %	9 %	8 %	11 %	21 %	7 %	6 %	5 %	2 %	0 %	1 %	9 %	15 %
5	1	35 %	41 %	7 %	3 %	23 %	24 %	21 %	15 %	6 %	7 %	0 %	4 %	7 %	5 %
5	2	49 %	50 %	3 %	6 %	16 %	22 %	16 %	11 %	2 %	2 %	1 %	0 %	12 %	10 %
6	1	44 %	39 %	5 %	3 %	18 %	18 %	20 %	19 %	2 %	6 %	1 %	7 %	9 %	8 %
6	2	59 %	45 %	4 %	2 %	9 %	8 %	18 %	32 %	3 %	3 %	0 %	3 %	8 %	7 %

Table 5
Percentage of cases where the participants did **not** look in the required direction deemed necessary for safe driving as defined in Table 2. Empty cells represent glance directions where the driver is not obliged to sample visual information.

		Forward		Backward		Left		Right	
	Drive	1	2	1	2	1	2	1	2
Location	Zone								
1	1	4 %	0 %	83 %	88 %			26 %	33 %
2	1	4 %	4 %	43 %	33 %			39 %	17 %
2	2	13 %	21 %			22 %	38 %	43 %	25 %
2	3	0 %	0 %	79 %	67 %				
3	1	0 %	0 %			21 %	13 %	58 %	63 %
4	1	0 %	0 %	38 %	38 %	19 %	14 %	52 %	48 %
5	1	0 %	0 %	27 %	42 %			23 %	21 %
5	2	0 %	0 %	36 %	30 %			41 %	48 %
6	1	0 %	0 %	72 %	65 %				
6	2	6 %	6 %	88 %	94 %				

shows what the vehicle knows about its surroundings. Monitoring this display is reasonable as it allows operators to anticipate and avoid hard braking manoeuvres by manually slowing down. While it is very reasonable to do so, this behaviour diverts attention from the outside of the bus limiting the potential to identify a hazard which the vehicle has not detected. Additionally, the participants occasionally looked at their mobile phones. Although some work-related tasks are managed via the phone, these tasks should not be prioritized in the analysed locations/zones where high situational awareness is needed. Overall, the glance results suggest a high level of reliance on the vehicle's capabilities, an inability to sustain attention over long periods, or a combination of both. It should be noted that at the time of data collection in Study A, the shuttles had been in operation for over a year. As such, operators had had noticeable opportunity to get used to the vehicles' behaviour. It is likely that gaze patterns changed as operators became more confident and used to shuttle operation, potentially anticipating likely scenarios where the shuttle would have an error. This may also explain the higher-than-expected glance proportion at a mobile phone, as drivers are known to regulate mobile phone use depending on the road conditions being faced (Oviedo-Trespalacios et al., 2018). The high percentage of phone glances, 10 %, at location 1, while there were hardly any phone glances in other locations, also indicates self-regulation, and that spare visual capacity was available at location 1. It is possible that ongoing training may be beneficial to sustain appropriate gaze allocation as confidence in the vehicles' ability increases.

Although there were no adverse events recorded during Study A, outside of the direct data collection period several anecdotal observations of rare events have been reported. These include red light violations, operators falling during hard braking events caused by oncoming traffic, and running over obstacles at "high" speed, resulting in dislodged sensors. The occurrence of these events evidence that the shuttle operator still has an important safety role to play, as yet the vehicles are not safe enough to drive unsupervised. The occurrence of these events also indicates that sometimes the human operator does not step in to resolve a safety issue. There are many reasons why an operator may not take over control at a safety critical time, for example, inappropriate gaze direction, low expectancy, habituation, inattention blindness, and the above-mentioned overreliance on the shuttle's capabilities. It may be unreasonable to expect that safety operators should stay attentive during tasks that are largely passive in nature (Alambeigi & McDonald, 2023; Greenlee et al., 2024), and a reassessment of the demands placed on safety operators may be needed to ensure safe operations.

In terms of sleep and fatigue, both objective and subjective metrics in both studies showed that participants generally had sufficient sleep and low sleepiness scores. However, individual variations existed. One driver in study A reported elevated KSS values, especially during Drive 1, but also during Drive 2, although at lower levels. The elevated fatigue level can likely be attributed to long and demanding workdays prior to data collection. Individual differences in fatigue development are well-known, even in situations where fatigue is more related to low or high workload rather than to physiological sleepiness (Szalma, 2012). Fatigue is a common problem in city bus drivers (Anund et al., 2016; Miller et al., 2020), but it does not appear to be a regular workplace hazard for shuttle operators. During Study B there were some occasions (27 % of shifts) when the objective measure of sleep durations from Fitbit suggested insufficient sleep, but the KSS scores did not reflect this. As there is strong evidence that individuals have good insight into their sleepiness (Cai et al., 2021), this discrepancy should be investigated further.

The study was conducted on the entire population of automated shuttle operators at the test site in Linköping. The drivers responsible for shuttle operations were handpicked from a pool of bus drivers based on their good health and dedication. The finding that they predominantly showed good sleep and health metrics over the two-month data collection period is therefore not surprising. This should not only be attributed to the safety operators though, but also to their employment environment. The employer has a mature fatigue management culture with sustainable scheduling and opportunities for rest and recovery, facilitated by teamwork between employees and employers. Although this works well on a small-scale test site, it remains to be seen whether safety operators will be prioritized and well-treated also when shuttle operations are scaled up to a city level. The results from this study should therefore be treated as a best-case scenario. For example, it has been

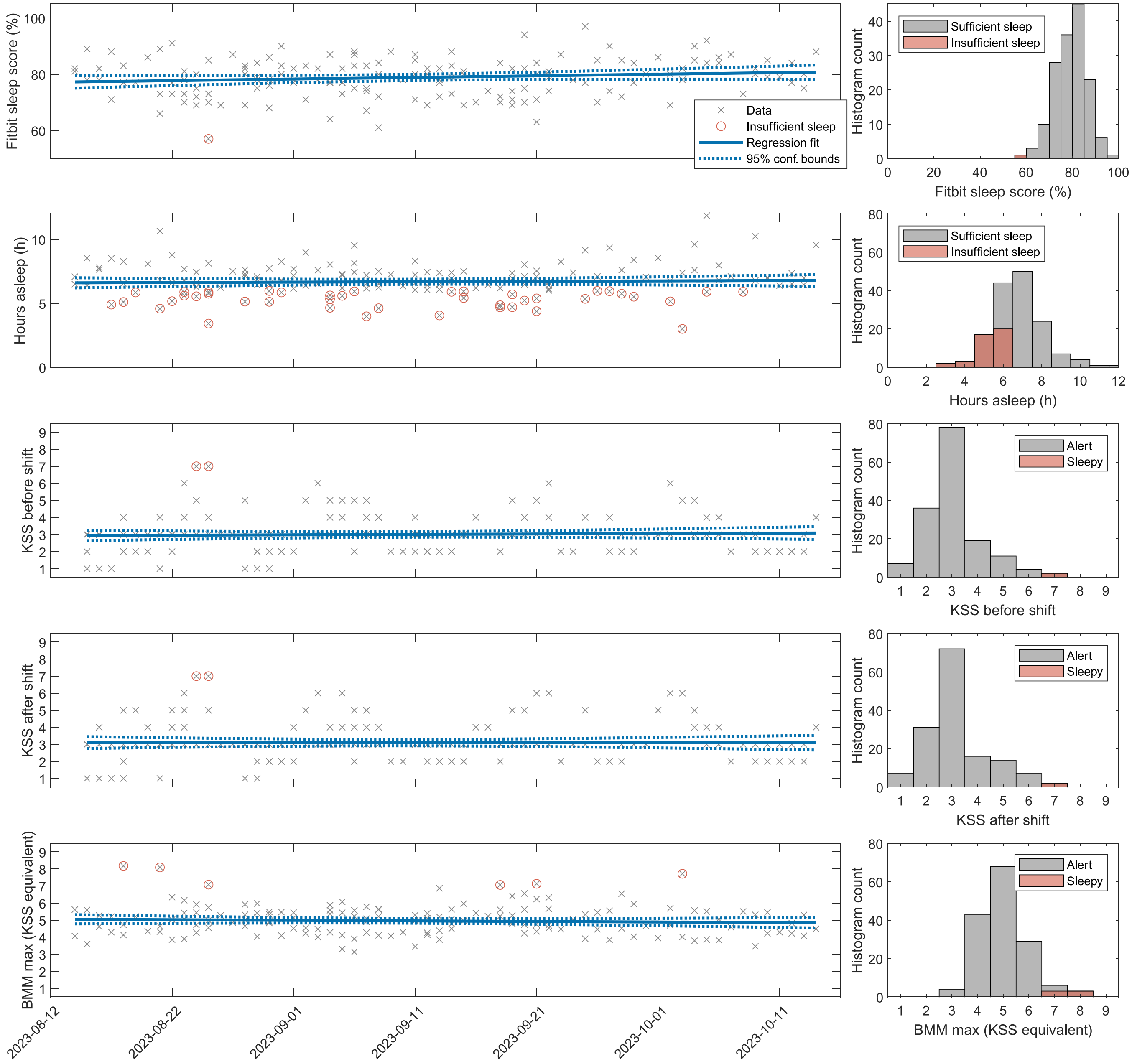


Fig. 4. Linear regression model plots (left) of various sleep hygiene metrics as a function of time. Data points indicating insufficient sleep are marked in red. Corresponding histograms are provided to the right. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

reported in London bus drivers that work related factors including less than 11 h off duty time between shifts and working six or more days without a rest day predicts sleepiness at work (Miller et al., 2020). A plausible explanation for the low sleepiness levels in this study is that these risk factors were not present during shuttle operation. Ensuring that professional drivers have sufficient time for rest and recuperation between shifts is likely to reduce problems with sleepiness at work (Zhang et al., 2023).

The launch of automated or autonomous vehicles is progressing rapidly, with advanced robotaxi technologies developed by companies such as Waymo, Baidu Apollo, WeRide, Zoox, and Cruise. These robotaxis aim for high levels of autonomy, often operating without a safety operator, and are designed for higher speeds and longer ranges compared to shuttles. They are also more versatile and suitable for both personal and commercial purposes. In addition to small-sized robotaxis, there are also larger robobuses focused on autonomous public transport solutions (e.g. Otokar and WeRide). The role of the safety operator may thus soon be obsolete, but until then, we must treat them well to ensure safe operations.

5. Conclusions

Employing both controlled experimental methods and naturalistic observation, the eight participating automated shuttle safety operators participating in the study generally had sufficient sleep, low sleepiness, and low stress levels across the two-month study period. However, the gaze patterns deviated from safe expectations with operators paying significantly less attention to the sides and rear of the vehicle. This could be attributed to several factors, such as a perceived need to focus on the forward direction for braking readiness, reliance on peripheral vision for

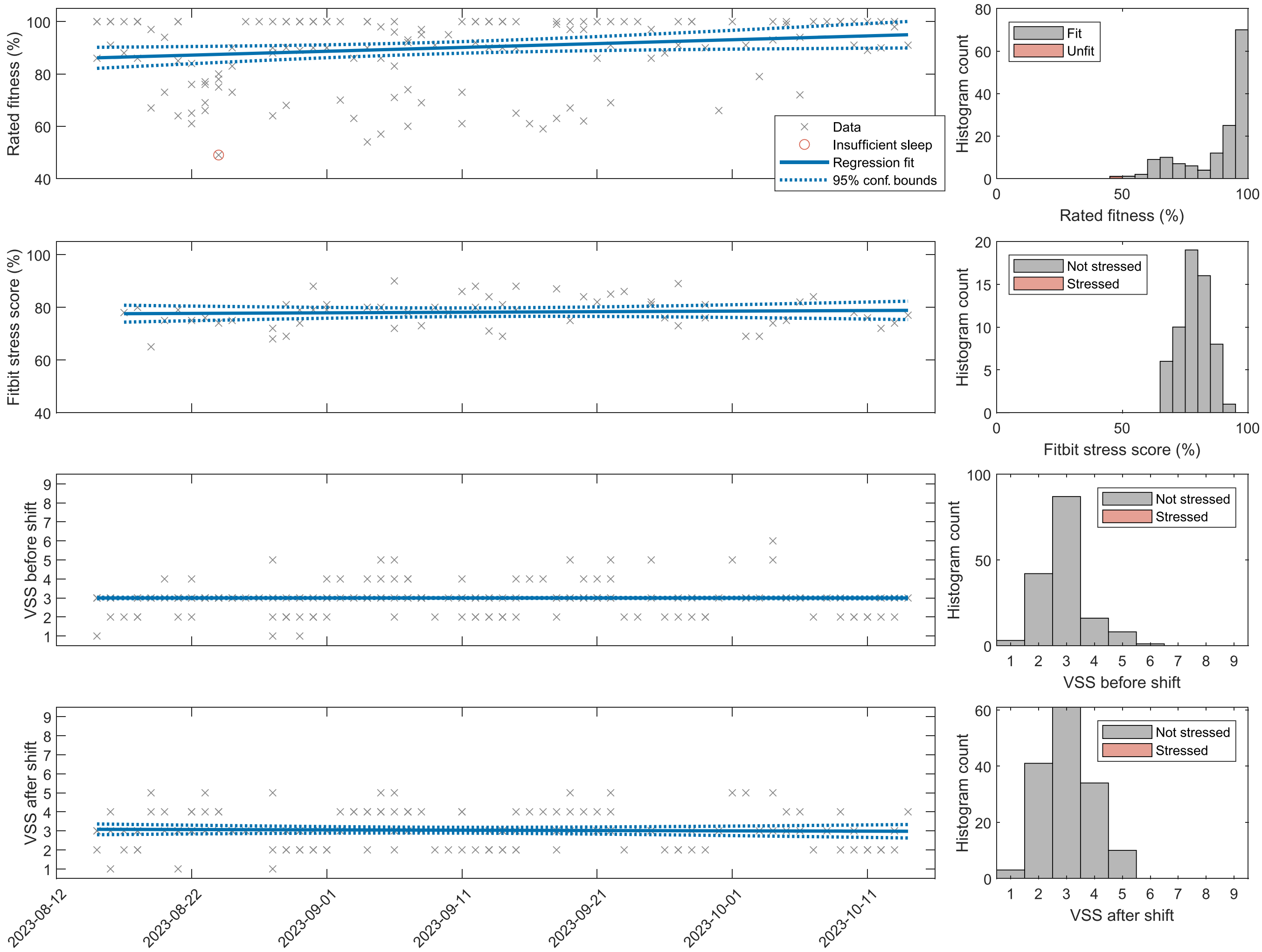


Fig. 5. Linear regression model plots (left) of various stress metrics as a function of time. No data points indicating high stress levels were found in the dataset. Corresponding histograms are provided to the right.

sideward information, or complacency regarding shuttle's capabilities.

The findings highlight the importance of ensuring all safety operators are well-prepared to avoid fatigue during work. Scheduling, rest, and recovery are crucial factors to ensure safe operations. Importantly, this is true for on-board operators but also applies to remote operators, which is the likely next step in supervising automated shuttles in public transportation.

CRediT authorship contribution statement

Christer Ahlström: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation. **My Weidel:** Writing – review & editing, Investigation, Data curation. **Anna Sjörs Dahlman:** Writing – review & editing, Methodology, Investigation. **Ashleigh Filtness:** Writing – review & editing, Methodology. **Anna Anund:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Ethics approval

The study was conducted in accordance with the declaration of Helsinki, and informed consent was obtained by all participants. The study protocols were approved by the Swedish Ethical Review Authority (Dnr 2022–06398-01 and Dnr 2020–04089).

Funding

This work was funded by European Union's Horizon 2020 research and innovation programme under grant agreement number 953426 (the PANACEA project), and by the Swedish Strategic Vehicle Research and Innovation Programme (VINNOVA FFI; Grant No. 2019-03104).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful to all PANACEA project partners for feedback and support in the preparation of this study. We are also very grateful to the data collection team at VTI, to the shuttle operators who participated in the study, and to the Ride the future consortium who runs the shuttle test site in Linköping.

Data availability

The authors do not have permission to share data.

References

Afonso, V.X., Tompkins, W.J., Nguyen, T.Q., Luo, S., 1999. ECG beat detection using filter banks. IEEE Trans. Biomed. Eng. 46 (2), 192–202.

Ahlström, C., Kircher, K., Nyström, M., Wolfe, B., 2021a. Eye Tracking in Driver Attention Research—How Gaze Data Interpretations Influence What We Learn [Perspective]. Front. Neuroergon. 2. https://doi.org/10.3389/fnrgo.2021.778043.

Ahlström, C., van Leeuwen, W., Krupenia, S., Jansson, H., Finér, S., Anund, A., Kecklund, G., 2021b. Real-time adaptation of driving time and rest periods in automated long-haul trucking: development of a system based on biomathematical modelling, fatigue and relaxation monitoring. IEEE Trans. Intell. Transport. Syst.

Åkerstedt, T., Connor, J., Gray, A., Kecklund, G., 2008. Predicting road crashes from a mathematical model of alertness regulation—The Sleep/Wake Predictor. Accid. Anal. Prev. 40 (4), 1480–1485. https://doi.org/10.1016/j.aap.2008.03.016.

Åkerstedt, T., Gillberg, M., 1990. Subjective and objective sleepiness in the active individual. Int. J. Neurosci. 52 (1–2), 29–37. http://www.scopus.com/inward/record.url?eid=2-s2.0-0025429130&partnerID=40&md5=3e11e548db13c5bce74c6bec09964fed.

Alambeigi, H., McDonald, A.D., 2023. A Bayesian regression analysis of the effects of alert presence and scenario criticality on automated vehicle takeover performance. Hum. Factors 65 (2), 288–305.

Anund, A., Ihlström, J., Fors, C., Kecklund, G., Filtness, A., 2016. Factors associated with self-reported driver sleepiness and incidents in city bus drivers. Ind. Health 54 (4), 337–346. https://doi.org/10.2486/indhealth.2015-0217.

Anund, A., Ludovic, R., Caroleo, B., Hardestam, H., Dahlman, A., Skogsmo, I., Arnone, M., 2022. Lessons learned from setting up a demonstration site with autonomous shuttle operation–based on experience from three cities in Europe. J. Urban Mobility 2, 100021.

Aubert, A.E., Ramaekers, D., Beckers, F., Breem, R., Denef, C., Van de Werf, F., Ector, H., 1999. The analysis of heart rate variability in unrestrained rats. Validation of method and results. Comput. Methods Progr. Biomed. 60 (3), 197–213.

Backhaus, R., 2020. Automatic shuttle buses - from the test track to scheduled services. ATZelectronics Worldwide 15 (11), 34–39. https://doi.org/10.1007/s38314-020-0296-x.

Bucchiarone, A., Battisti, S., Marconi, A., Maldacea, R., Ponce, D.C., 2021. Autonomous Shuttle-as-a-Service (ASaaS): challenges, opportunities, and social implications. IEEE Trans. Intell. Transp. Syst. 22 (6), 3790–3799. https://doi.org/10.1109/TITS.2020.3025670.

Cai, A.W.T., Manousakis, J.E., Lo, T.Y.T., Horne, J.A., Howard, M.E., Anderson, C., 2021. I think I'm sleepy, therefore I am – Awareness of sleepiness while driving: A systematic review. Sleep Med. Rev. 60. https://doi.org/10.1016/j.smrv.2021.101533, 101533.

Cirelli, C., Benca, R., Eichler, A., 2017. Insufficient sleep: Definition, epidemiology, and adverse outcomes. Nat. Sleep Found. Consensus Rep. https://www.uptodate.com/contents/insufficient-sleep-definition-epidemiology-and-adverse-outcomes.

Favaro, F., Hutchings, K., Nemec, P., Cavalcante, L., & Victor, T. (2022). Waymo's Fatigue Risk Management Framework: Prevention, Monitoring, and Mitigation of Fatigue-Induced Risks while Testing Automated Driving Systems. https://doi.org/10.48550/arXiv.2208.12833.

Greenlee, E.T., DeLucia, P.R., Newton, D.C., 2019. Driver vigilance in automated vehicles: Effects of demands on hazard detection performance. Hum. Factors 61 (3), 474–487.

Greenlee, E.T., DeLucia, P.R., Newton, D.C., 2024. Driver vigilance decrement is more severe during automated driving than manual driving. Hum. Factors 66 (2), 574–588.

Hilgarter, K., Granig, P., 2020. Public perception of autonomous vehicles: A qualitative study based on interviews after riding an autonomous shuttle. Transport. Res. F: Traffic Psychol. Behav. 72, 226–243.

Iclodean, C., Cordos, N., Varga, B.O., 2020. Autonomous shuttle bus for public transportation: A review. Energies 13 (11), 2917.

Jammes, B., Sharabty, H., Esteve, D., 2008. Automatic EOG analysis: a first step toward automatic drowsiness scoring during wake-sleep transitions. Somnologie - Schlafforschung Und Schlafmedizin 12 (3), 227–232. https://doi.org/10.1007/s11818-008-0351-y.

Kircher, K., Ahlstrom, C., 2017. Minimum required attention: a human-centered approach to driver inattention. Hum. Factors 59 (3), 471–484. https://doi.org/10.1177/0018720816672756.

Lu, K., Sjörs Dahlman, A., Karlsson, J., Candefjord, S., 2022. Detecting driver fatigue using heart rate variability: A systematic review. Accid. Anal. Prev. 178, 106830. https://doi.org/10.1016/j.aap.2022.106830.

Mack, A., 2003. Inattentional blindness: Looking without seeing. Curr. Dir. Psychol. Sci. 12 (5), 180–184.

Miller, K.A., Filtness, A.J., Anund, A., Maynard, S.E., Pilkington-Cheney, F., 2020. Contributory factors to sleepiness amongst London bus drivers. Transport. Res. F: Traffic Psychol. Behav. 73, 415–424. https://doi.org/10.1016/j.trf.2020.07.012.

Mouratidis, K., Serrano, V.C., 2021. Autonomous buses: Intentions to use, passenger experiences, and suggestions for improvement. Transport. Res. F: Traffic Psychol. Behav. 76, 321–335.

Oviedo-Trespalacios, O., Haque, M.M., King, M., Demmel, S., 2018. Driving behaviour while self-regulating mobile phone interactions: A human-machine system approach. Accid. Anal. Prev. 118, 253–262. https://doi.org/10.1016/j.aap.2018.03.020.

Persson, A., Jonasson, H., Fredriksson, I., Wiklund, U., Ahlström, C., 2020. Heart rate variability for classification of alert versus sleep deprived drivers in real road driving conditions. IEEE Trans. Intell. Transp. Syst. 22 (6), 3316–3325.

Philip, P., Micoulaud-Franchi, J.-A., Taillard, J., Coelho, J., Tisserand, C., Dauvilliers, Y., Sagaspe, P., 2023. The Bordeaux Sleepiness Scale (BOSS): a new questionnaire to measure sleep-related driving risk. J. Clin. Sleep Med. 19 (5), 957–965.

Piatkowski, D.P., 2021. Autonomous shuttles: what do users expect and how will they use them? J. Urban Technol. 28 (3–4), 97–115.

Salonen, A.O., Haavisto, N., 2019. Towards autonomous transportation. Passengers' experiences, perceptions and feelings in a driverless shuttle bus in Finland. Sustainability 11 (3), 588.

Schleicher, R., Galley, N., Briest, S., Galley, L., 2008. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? Ergonomics 51 (7), 982–1010. https://doi.org/10.1080/00140130701817062.

Schrank, A., Kettwich, C., Oehl, M., 2024. Aiding automated shuttles with their driving tasks as an on-board operator: a case study on different automated driving systems in three living labs. Appl. Sci. 14 (8), 3336. https://www.mdpi.com/2076-3417/14/8/3336.

Schuß, M., Rollwagen, A., Riener, A., 2022. Understanding operator influence in automated urban shuttle buses and recommendations for future development. Multimodal Technol. Interact. 6 (12), 109.

Shaffer, F., Ginsberg, J., 2017. An overview of heart rate variability metrics and norms. Front. Public Health 5, 258.

Sherry, L., Shortle, J., Donohue, G., Berlin, B., & West, J. (2020, 8-10 Sept. 2020). Autonomous Systems Design, Testing, and Deployment: Lessons Learned from The Deployment of an Autonomous Shuttle Bus. 2020 Integrated Communications Navigation and Surveillance Conference (ICNS).

Simons, D.J., 2000. Attentional capture and inattentional blindness. Trends Cogn. Sci. 4 (4), 147–155.

Sjörs Dahlman, A., Karlsson, K., Candefjord, S., Anund, A., 2024. Validation of a one-item acute stress scale for driving tasks. Adv. Human Factors Transport. 148, 390–396. https://doi.org/10.54941/ahfe1005230.

Solís-Marcos, I., Ahlström, C., Kircher, K., 2018. Performance of an additional task during Level 2 automated driving: an on-road study comparing drivers with and without experience with partial automation, 0018720818773636 Hum. Factors.

Stapel, J., Mullakkal-Babu, F.A., Happee, R., 2019. Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. Transport. Res. F: Traffic Psychol. Behav. 60, 590–605. https://doi.org/10.1016/j.trf.2018.11.006.

Szalma, J.L., 2012. Individual Differences in Stress, Fatigue and Performance. Taylor & Francis Group http://ebookcentral.proquest.com/lib/linkoping-ebooks/detail.action?docID=883348.

Thorhauge, M., Fjendbo Jensen, A., Rich, J., 2022. Effects of autonomous first- and last mile transport in the transport chain. Transp. Res. Interdiscip. Perspect. 15, 100623. https://doi.org/10.1016/j.trip.2022.100623.

Vater, C., Wolfe, B., Rosenholtz, R., 2022. Peripheral vision in real-world tasks: A systematic review. Psychon. Bull. Rev. 29 (5), 1531–1557. https://doi.org/10.3758/s13423-022-02117-w.

Warm, J.S., Parasuraman, R., Matthews, G., 2008. Vigilance requires hard mental work and is stressful. Hum. Factors 50 (3), 433–441.

Whitmore, A., Samaras, C., Hendrickson, C.T., Scott Matthews, H., Wong-Parodi, G., 2022. Integrating public transportation and shared autonomous mobility for equitable transit coverage: A cost-efficiency analysis. Transp. Res. Interdiscip. Perspect. 14, 100571. https://doi.org/10.1016/j.trip.2022.100571.

Wierwille, W.W., Ellsworth, L.A., 1994. Evaluation of driver drowsiness by trained raters. Accid. Anal. Prev. 26 (5), 571–581. http://www.scopus.com/inward/record.url?eid=2-s2.0-0028525249&partnerID=40&md5=058796dc23a34a2a226512f8ffe2d5c7.

Zhang, H., Ni, D., Ding, N., Sun, Y., Zhang, Q., Li, X., 2023. Structural analysis of driver fatigue behavior: A systematic review. Transp. Res. Interdiscip. Perspect. 21, 100865. https://doi.org/10.1016/j.trip.2023.100865.