Quantifying Workload and Stress in Intensive Care Unit Nurses: Preliminary Evaluation Using Continuous Eye-Tracking

Nima Ahmadi, Center for Outcomes Research, Houston Methodist, Houston, TX, USA, **Farzan Sasangohar**, Center for Outcomes Research, Houston Methodist, Houston, TX, USA and Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA, **Jing Yang** and **Denny Yu**, School of Industrial Engineering, Purdue University, West Lafayette, IN, USA, **Valerie Danesh**, Baylor Scott & White Health, Center for Applied Health Research, Dallas, TX, USA and University of Texas at Austin, School of Nursing, Austin, TX, USA, **Steven Klahn** and **Faisal Masud**, Center for Critical Care, Houston Methodist Hospital, Houston, TX, USA

Objective: (1) To assess mental workloads of intensive care unit (ICU) nurses in 12-hour working shifts (days and nights) using eye movement data; (2) to explore the impact of stress on the ocular metrics of nurses performing patient care in the ICU.

Background: Prior studies have employed workload scoring systems or accelerometer data to assess ICU nurses' workload. This is the first naturalistic attempt to explore nurses' mental workload using eye movement data.

Methods: Tobii Pro Glasses 2 eye-tracking and Empatica E4 devices were used to collect eye movement and physiological data from 15 nurses during 12-hour shifts (252 observation hours). We used mixed-effect models and an ordinal regression model with a random effect to analyze the changes in eye movement metrics during high stress episodes.

Results: While the cadence and characteristics of nurse workload can vary between day shift and night shift, no significant difference in eye movement values was detected. However, eye movement metrics showed that the initial handoff period of nursing shifts has a higher mental workload compared with other times. Analysis of ocular metrics showed that stress is positively associated with an increase in number of eye fixations and gaze entropy, but negatively correlated with the duration of saccades and pupil diameter.

Conclusion: Eye-tracking technology can be used to assess the temporal variation of stress and associated changes with mental workload in the ICU environment. A real-time system could be developed for monitoring stress and workload for intervention development.

Keywords: eye movements, naturalistic study, intensive care unit, mental workload

Address correspondence to Farzan Sasangohar, Industrial and Systems Engineering, Texas A&M University, College Station, 3131 TAMU, Houston, TX 77842, USA; e-mail: sasangohar@tamu.edu

HUMAN FACTORS
Vol. 0, No. 0, ■■ ■, pp. 1-15
DOI:10.1177/00187208221085335
Article reuse guidelines: sagepub.com/journals-permissions

High workload is one of the key job stressors for intensive care unit (ICU) nurses (Malacrida et al., 1991; Neill, 2011; Oates & Oates, 1996) and is linked to burnout (Aiken et al., 2002), increased patients' length of stay (Amaravadi et al., 2000), and higher mortality rates (Carayon & Gürses, 2005; Cho et al., 2003; Kiekkas et al., 2008; Tarnow-Mordi et al., 2000). Workload in the context of the clinical care delivery in the ICU setting includes both physically and cognitively demanding work (Young et al., 2015). Cognitive workload can be defined as demand on human mental or attentional resources (Jeffri & Awang Rambli, 2021). ICU nursing has been associated with high cognitive workload. For example, ICU nurses are exposed to frequent alarms arising from the continuous physiologic monitoring and care demands associated with critically ill patients (Donchin & Seagull, 2002). The frequent stimuli of physiologic alarms and complex care coordination is concurrent with performing critical reasoning with clinical decisions tied to patient safety and outcomes (Higgs et al., 2008; Stubbings et al., 2012). To make informed decisions, high levels of situational awareness are required to critically evaluate evolving physical assessments, laboratory results, medication administration, and care plans to deliver patient- and family-centered care (Davidson et al., 2017), as well as to be responsive to fluctuating physiologic status and needs (Endsley, 1995; Stubbings et al., 2012).

Situational awareness could be compromised by cognitive overloading arising from the parallel activities and multifaceted demands of patient care, including assembling patient care supplies (e.g., wound care, suctioning equipment; Koch et al., 2012); receiving communication from multiple sources (e.g., audio, written, visual changes in patient condition) with sensitive interpersonal communication between clinicians and/or family members; real-time electronic documentation requirements; and time-sensitive interactions with complex technologies (e.g., infusion pump programming, ventilator settings). To prevent reaching a "red zone" region, defined as an unacceptable level of mental workload when one is overloaded and cognitive demand exceeds cognitive capacity (Pickup et al., 2005; D. Yu et al., 2016), techniques for assessing ICU nurses' cognitive workload are required.

In prior nursing ICU studies, patient-based approximations were used to evaluate workload, including nurse-to-patient staffing ratios, patient acuity measures, and/or workload scoring systems such as the Therapeutic Intervention Scoring System-76 (TISS-76; Lee et al., 2017) and Nursing Activities Score (NAS; Armstrong et al., 2015; Debergh et al., 2012; Moghadam et al., 2021; Nogueirade et al., 2014; Stafseth et al., 2011), quantitatively derived from 76 therapeutic items and 22 nursing activities, respectively. However, previous approximations do not capture the temporal fluctuations of workload within a patient care shift, and cannot be distilled to assess workload specific to demanding events or periods (e.g., handoffs periods; Carayon & Gürses, 2005; Hoonakker et al., 2011). In an alternative human factors paradigm, instead of using patient characteristics, workload is estimated from nurses' perspective using subjective self-reported tools and objective measures extracted from the autonomic nervous system activity (Carayon & Gürses, 2005; Hoonakker et al., 2011; Wilbanks & McMullan, 2018) and could be utilized to address this shortcoming. Self-reported tools such as NASA-Task Load Index (NASA-TLX) are easy to use, but their applications are limited to perceived workload (Byrne et al., 2010; Hoonakker et al., 2011) and can be susceptible to response bias (Matthews et al., 2015). In an effort to address the subjectivity in workload assessment, patient conditions have been uses to develop workload scoring systems (Debergh et al., 2012; Moghadam et al., 2021; Reis Miranda et al., 1996). For instance, the APACHE II Score assess workload based on patients' conditions such as temperature, metal arterial pressure, pH, PaO_2, mmHg, and Glasgow Coma Scale.

Objective measures have the potential of temporal workload assessment based on documented correlations between cognitive load and physiological responses. Of these responses, eye movement metrics—fixational, saccadic, and pupillary—could be used as unobtrusive markers of cognitive load or mental stress with implications for integration into the clinical work environment (Skaramagkas et al., 2021). Pupillary dilations have been shown to be relevant to the levels of cognitive workload (Castner et al., 2020; Hess & Polt, 1964; Hyönä et al., 1995; Mandrick et al., 2016; Menekse Dalveren & Cagiltay, 2018; Schulz et al., 2011; Szulewski et al., 2015; Wu et al., 2020, 2021; Zhang et al., 2017) with applications in real-time workload monitoring (Appel et al., 2018, 2019). Fixation duration and frequency are also sensitive to task-specific characteristics (Chapman & Underwood, 1998; de Greef et al., 2009; De Rivecourt et al., 2008; Foy & Chapman, 2018; Marandi et al., 2018; Recarte & Nunes, 2000; Schulz et al., 2011) and associated with difficulties in information processing (Ehmke & Wilson, 2007; Holmqvist et al., 2011). Saccades—combination of durations (Marandi et al., 2018; Szewczyk et al., 2020) and frequency (Kataoka et al., 2011; Szewczyk et al., 2020)—are also markers of mental workload. Furthermore, the randomness in the visual search—via either entropy (a metric of fixation and gaze points randomness; Di Stasi et al., 2016; Diaz-Piedra et al., 2019; Harris et al., 1986; Tole et al., 1982; Wu et al., 2020, 2021) or as nearest neighbor index (NNI; a marker of fixation points dispersion; Camilli et al., 2008; Di Nocera et al., 2006, 2007)—is associated with the mental workload.

In the majority of eye-tracking studies in healthcare settings, the application of eye tracking technology was limited to simulated settings to explore diagnostic visual search and clinical decision making (Brunyé et al., 2019), visual attention (Herrick et al., 2020; Hofmaenner et al., 2021; Koh et al., 2011; Law et al., 2020; Marquard et al., 2011; Weinberg et al., 2020), skill assessments (Harezlak & Kasprowski, 2018; Tien et al., 2014), or to perform cognitive task analysis (Zehnder et al., 2021). Simulated settings specific to healthcare have included the cognitive workload of anesthetists (Schulz et al., 2011), surgeons (Menekse Dalveren & Cagiltay, 2018; Wu et al., 2020, 2021; Zhang et al., 2017), and nurses (Kataoka et al., 2011). In naturalistic studies of critical care nurses, eye tracking technology was primarily used for audio and video recording of ICU activities from nurses' point of view, with retrospective analysis of 1–3 hour segments to investigate visual attention (Hofmaenner et al., 2021), to perform cognitive task analysis (Zehnder et al., 2021), and to examine the impact of interruption on nursing activities (Grundgeiger et al., 2010).

The majority of research conducted in this domain is limited to interactions with specific interfaces and isolated tasks. To the best of our knowledge, no naturalistic study has been conducted in ICUs to holistically assess nurses' mental workload in a cognitively demanding digitalized environment for the entire shift consisting of combination of tasks using eye movement metrics. The objective of this research is to explore nurses' workload during patient care using fixational, saccadic, and pupillary responses in the naturalistic setting of the ICU. In addition, as part of a larger study, we evaluated the eye movement metrics during periods of elevated stress—which was defined as physiological responses to life challenges that trigger fight or flight responses (Acerbi et al., 2017). While most prior literature on eye movement metrics has examined the efficacy of such measures for workload assessment, their utility for stress measurement remains a research gap. Characteristics of visual search under high workload could be divergent from stressful episodes when participants are anxious due to usability problems of ICU devices or time pressure; elevated mental workload could present via narrowed visual attention and longer durations with a fewer number of fixations, while stress manifests itself by shorter fixation duration, a higher number of fixations (Holmqvist et al., 2011) and more random visual search.

To address the gap in naturalistic evaluation of nurses' workload and its correlations with stress, as grounded in previous studies (Debergh et al., 2012; Moghadam et al., 2021; Reis Miranda et al., 1996; Yu et al., 2020)—which reported a higher workload of nurses in the day and evening shifts (7 a.m. – 11 p.m.) compared to night shifts (11 p.m. – 7 a.m.) and higher levels of physical activity in 12-hour day shifts than night shifts (F. Yu et al., 2020)—we tested the following hypotheses:

Hypothesis 1: Changes of ocular metrics correlate with higher mental workload of nurses in day shift compared to night shift nurses.

Hypothesis 2: Changes of ocular metrics indicate higher mental workload at the start of shifts than middle and end of shifts.

Hypothesis 3: Changes of ocular metrics such as fixational metrics, randomness in visual search (e.g., entropy), and pupil diameter are indicative of acute stress.

METHOD

The research was approved by the Houston Methodist Research Institute Institutional Review Board (Pro00019025), including the implementation of special ethical considerations specific to eye tracking technology in the ICU work environment (Larsen et al., 2020).

Participants and Setting

A convenience sample of 21 registered nurses were recruited from a 40-bed cardiovascular intensive care unit at Houston Methodist Hospital, a large metropolitan tertiary care hospital in Southeastern Texas. The 12-hour shifts began from 7:00 p.m./7:00 a.m. to 7:00

a.m./7:00 p.m. and the first and last hours of shifts were considered as handoff periods. While the eye tracking data from all 21 participants were used for comparison of workloads between day and night shifts, complete data—one 12-hour shift per participant—from both the physiologic and eye-tracking sensors was only available for 15 nurse participants, representing approximately 180 observation hours. Demographic information is presented in Table 1.

Procedure

After obtaining written informed consent and within the first 10 min of the work shift, the study administrator equipped participants with two wearable technologies: (a) Empatica E4—a wristband technology (which was worn on the self-reported non-dominant hand) with electrodermal activity, skin temperature, photoplethysmography, and accelerometer sensors that captured heart rate at 32 Hz; and (b) the Tobii Pro Glasses 2—a head-mounted eye tracking device that is equipped with one microphone and four cameras to reliably track subjects' eye movements. The data collection began by calibrating the eye-tracker and proceeded throughout the 12-hour nursing shifts; the study administrator was on-site to change the eye-trackers' batteries, recalibrate the eye tracking device, address technical troubleshooting, and administer end-of-shift surveys.

Analysis

The captured eye movements data (at 50 Hz) were exported to Tobii Pro Lab for detection of eye movement elements using a velocity-threshold identification fixation filter. A python script processed Tobii Pro Lab's generated files and one-minute intervals were used for analysis to align with the stress data that used the same timeframe. To measure physical and cognitive stress, the interbeat interval (IBI) data from the Empatica device was used to compute the Baevsky Stress Index (Baevsky & Berseneva, 2008) for each minute. Low, normal, and high-stress zones that represent stress intensities were defined based on values of the Baevsky stress index (Baevsky & Berseneva, 2008) where stress index between 80 and 150 was considered normal. Physiological stress correlations with the Empatica on this population and a semi-overlapping dataset are reported elsewhere (Ahmadi et al., 2022). Nurses' mental workload was estimated based on industry standards of correlations between cognitive load and ocular metrics: pupil diameter, fixation duration, number of fixations, saccade duration, number of saccades, entropy, and NNI. Visual entropy (in bits) was measured per Shannon entropy (Shannon, 1948; Wu et al., 2020) and NNI was computed in accordance with the Di Nocera et al. (2006) method. All eye movement metrics were measured for each minute of data.

Cognitive Workload at Handoff Periods and Mid-shifts. For hypotheses 1 and 2, the

TABLE 1: Demographic information of study participants

Demographic Information	Eye movement data, *n* = 21	Physiologic + Eye movements data, *n* = 15
Gender, *n* (%)		
Male	4 (19%)	3 (20%)
Female	17 (81%)	12 (80%)
Age, years (mean ± std. Dev.)	34.1 ± 7.6	35.1 ± 8.2
Shift, *n* (%)		
Day shift	14 (67%)	10 (67%)
Night shift	7 (33%)	5 (33%)
ICU experience, years (mean ± std. Dev.)	7.5 ± 6.7	8.7 ± 7.4

Note. ICU = Intensive Care Unit.

TABLE 2: Post-hoc comparisons and effect sizes of independent variables (working shift and handoff periods) on each eye tracking response metric

	Periods of ICU nursing workload comparisons											
	Day versus night shifts				Start versus end of shifts				Start versus mid-shifts			
Ocular Metrics	Day	Night	*p*-value	Effect size	Start	End	*p*-value	Effect size	Start	Mid	*p*-value	Effect size
NNI	0.6 ± 0.1	0.6 ± 0.2	0.78	0.51	0.6 ± 0.2	0.6 ± 0.1	<0.01	0.85	0.6 ± 0.1	0.6 ± 0.1	<0.01	0.27
NF	92.7 ± 4.2	85.2 ± 24.1	0.06	0.90	88.1 ± 25.4	91.7 ± 23.7	<0.01	0.02	88.1 ± 25.4	90.5 ± 24.2	0.03	0.12
NS	99.0 ± 39.0	89.3 ± 35.3	0.12	0.70	89.1 ± 37.5	100.7 ± 38.3	<0.01	0.05	89.1 ± 37.5	97.1 ± 37.4	<0.01	0.26
DS	4.1 ± 0.7	4.1 ± 0.7	0.24	0.50	4.1 ± 0.7	4.17 ± 0.7	<0.01	0.08	4.1 ± 0.7	4.1 ± 0.7	<0.01	0.34
PD log	0.04 ± 0.0	0.05 ± 0.0	0.75	0.14	0.04 ± 0.0	0.05 ± 0.0	<0.01	0.04	0.04 ± 0.0	0.05 ± 0.0	0.8	0.07

Note. DIF, mean metric of the first group – mean metric of the second group); DS = Duration of Saccades; NF = number of fixations; NNI = Nearest Neighbor Index; NS = number of saccades; PD = pupil diameter; effect size of Cohen's d: small = 0.2, medium = 0.5, large = 0.8.
Gray cells indicate statistical significance.

impacts of working shifts (day and night), handoff periods (first and last hour of shifts), and mid-shifts (1 hour of mid-shifts) on eye movement metrics were estimated through mixed effect models. One model was developed for each eye movement metric where the response variable was an eye movement, the fixed effects were either working shifts, handoff periods, or mid-shift, and the random effect was research participants. The effect size was calculated using Cohen's *d*, and the Tukey post hoc test of mixed effect models was used to find differences between eye movements measures in different working shifts and handoff periods.

Correlation Between Stress and Eye Movement Metrics. For hypothesis 3, an ordinal regression model with a random effect was used (Cumulative Link Mixed Models) to identify correlations between eye movement metrics with stress levels which were defined as the response variable with three levels: low, normal, and high stress. The eye metrics and their interactions with shifts were represented as fixed variables in the model. This model also included a random effect which was research participants. To find the optimal subset of independent variables included in the model, the stepwise regression algorithm was performed; the Akaike information criterion was used as a criterion for forward stepwise selection to determine the optimal subset of variables. As for the effect of each identified fixed variable, the odds ratio was calculated, which measures changes in odds of stress level for a unit increase in a predictor while holding constant all other variables included in the model.

RESULTS

Working Shift and Handoff Period versus Eye-tracking

The average of eye movement values in the day shift were compared with the night shift (hypothesis #1), but no significant difference was found between the shifts (all p-values > 0.05) as eye movements metrics were not significantly different among day and night shift nurses. The analysis of eye movement data showed that all eye movement metrics were significantly higher (p-values < 0.05) at the end of working shifts compared to the start of shifts (hypothesis #2). All eye movement metrics suggested the beginning of the shifts were more demanding than the end of shifts except pupil diameter. Similar results were obtained when the start of shifts was compared to the middle of shifts (hypothesis #2); all eye movement metrics except pupil diameter were significantly higher at the middle of shifts compared to the start of shifts. The results of mixed effect models are summarized in Table 2. The outcomes of averaged fixation duration and entropy were excluded, as those metrics did not satisfy the normality assumption of the model and failed to pass the Shapiro–Wilk test. The pupil diameter was transferred to log form to ensure the validation of normality assumption of the model.

Stress and Eye Movement Metrics

Table 3 presents the results of the cumulative link mixed-model analysis carried out using identified optimal subset eye tracking metrics to assess association with different levels of stress. The optimal subset of independent variables that were identified by the stepwise regression algorithm includes: shift, pupil diameter, averaged duration of saccades, number of fixations, number of saccades, NNI, fixation duration, and entropy. There were no interaction terms identified by the algorithm. In Table 3, the positive coefficients indicated that as the number of fixations (p < 0.001) and entropy (p = 0.01) increased, nurses were more likely experiencing higher stress levels (hypothesis #3). In contrast, the pupil diameter (p < 0.001) and duration of saccades (p < 0.001) were significantly decreased as participants' stress index increased. Moreover, the shift was found to be correlated with stress level with a p-value less than 0.001; in particular, the night shift is associated with increased stress which has the largest effect size with an odds ratio of 3.8, indicating that the odds of being stressful is multiplied 3.8 times in the night shifts compared to day shifts.

Table 3: Summary of cumulative link mixed model of the stress level response variable

Independent Variables	Coeff (Std. Err)	*p*-value	Odds Ratio
Shift	1.35 (0.11)	<0.01	3.80
Pupil diameter	−0.46 (0.04)	<0.01	0.62
Averaged DS	−0.34 (0.03)	<0.01	0.70
Number of fixations	0.18 (0.03)	<0.01	1.19
Entropy	0.13 (0.05)	<0.01	1.14
Number of saccades	−0.06 (0.07)	0.22	0.87
NNI	0.04 (0.18)	0.17	1.04
Averaged DF	−0.03 (0.049)	0.54	0.97

Note. DF = duration of fixations; DS = duration of saccades; PD = pupil diameter; NF = number of fixations; NNI = nearest neighbor index; NS = number of saccades.
Gray cells indicate statistical significance. Log-likelihood −3594, AIC = 7212.

DISCUSSION

We studied ocular metrics (number of fixations, duration of fixations, number of saccades, duration of saccades, entropy, NNI, and pupil diameter) to estimate changes in critical care nurses' mental workload during 12-hour day and night shifts as well as handoff periods based on the documented correlations between eye movements and cognitive load in a naturalistic study.

Mental Workload of Day Shift Nurses versus Night Shift Nurses (Hypothesis #1)

The results of mixed effect models showed no significant difference between day shifts and night shifts in terms of eye movement metrics; it indicates that overall mental workload during the 12-hour day and night shifts were roughly equivalent, in contrast with some previous work suggesting that day and evening shift nurses experience a higher workload compared to 8-hour night shift nurses (Debergh et al., 2012; Moghadam et al., 2021; F. Yu et al., 2020; see also Reis Miranda et al., 1996). A potential reason for such divergence is differences associated with the methods used for workload measurement, including comparisons between 8-hour shifts of days/evenings (7 a.m. – 11 p.m.) and nights (11 p.m. – 7 a.m.) that vary from 12-hour shifts predominant in the United States (7 a.m. – 7 p.m., 7 p.m. – 7 a.m.). When workload was estimated using workload scoring systems, both physical and cognitive tasks are included. For example, Debergh et al. (2012) used NAS to estimate nurses' workload based on 23 nursing activities, including positioning patients, medication administration, family communication, and specific procedural interventions such as nurse workload to support endotracheal intubation (Miranda et al., 2003). Results of such workload measures could be misleading when estimating the mental workload of nurses since a task's level of demand depend on skill-, rule-, and/or knowledge-based components (Rasmussen, 1983). Skill- (e.g., routine tasks) and rule-based tasks are governed by structured patterns of thoughts and deterministic rules. Conversely, knowledge-based tasks require analytical processes, reasoning, and novel solutions which make tasks more demanding (Armitage, 2009; Henneman et al., 2010). Instead, NAS and other similar methods use time-on-task or categorize task difficulty without accounting for different levels of experience and expertise. Similarly, studies that reported higher physical intensity (F. Yu et al., 2020), energy expenditure, and numbers of steps (Kwiecień-Jaguś et al., 2019) in day shifts than night shifts measured only physical demand and activity classifications (e.g., standing, sitting).

The inherent uncertainty in ICU activities and processes entail that nursing workload could fluctuate within the 12-hour ICU shifts and regardless of the time of day or night due to the dynamic demands of patient admissions and discharges (Nogueira et al., 2014; Romano et al., 2019), ICU occupancy (Kim et al., 2017), fluctuating patient acuity (Padilha et al., 2008), therapeutic interventions (Romano et al., 2019), and the organization-specific cadence of routine nursing activities (Romano et al., 2019); these variables are common among day and night ICU shifts. The said factors have similar impacts on the workload of the day and night shifts and should be considered in the workload assessment.

Mental Workload at Start versus End of shifts and Mid-shifts (Hypothesis #2)

Per hypothesis #2, higher mental workload at the start of shifts (of incoming nurses) compared to the end of shifts (of outgoing nurses) at handoff periods was predicted. A handoff period occurs at shift change and is characterized by transition of authority, responsibility, and information from an outgoing nurse (at end of shift) to an incoming nurse (at start of shift; Friesen et al., 2008). During shift change, a considerable amount of information is communicated to the incoming nurses including patients' information and care plan. The handoff communication period is approximately 30 min. Intaking information raises the incoming nurses' cognitive load and subsequently affects eye movements, attributable to dynamic reviews of the electronic health record, patient assessments, and equipment inspections.

In this study, nurses had a lower frequency of fixations and saccades, and a shorter duration of saccades at the start-of-shift handoff period compared to the end- or middle-of-shifts. Previous studies that examined correlations between fixational and saccadic metrics with mental workload reported mixed results (Table 4); these metrics are indicative of operators' visual search (e.g., drivers, pilots, clinicians) and reflect their visual search strategies.

For example, in some cases, drivers must scan right or left to receive information; similarly, pilots in emergency flight scenarios explore the flight environment to detect faulty equipment. In other words, scanning patterns differ under time pressure which subsequently could reverse the direction of change in fixational and saccadic eye movements in demanding situations. For example, in demanding situations, drivers (Yang et al., 2012) and harvester operators (Szewczyk et al., 2020) had a higher number of saccades, while another study reported a reduction in saccade rate under elevated workload (Nakayama et al., 2002).

The outcomes of this study are aligned with the latter study (Nakayama et al., 2002). Our results are also in line with Kataoka et al. (2011) who reported a reduction in number of saccades while programming an infusion pump under time pressure (elevated mental workload), and with a study that examined the reliability of

TABLE 4: Summary of changes in ocular metrics under elevated mental workload

Ocular metrics	Correlation with an elevated workload
Pupil diameter	▲ (Lohani et al., 2019; Marquart et al., 2015; Wu et al., 2020)
Duration of saccades	▼ (Marandi et al., 2018; Szewczyk et al., 2020)
Number of fixations	▲ (Szewczyk et al., 2020), ▼(Marandi et al., 2018)
Entropy	Depending on the events and objectives of operators ▼ (Diaz-Piedra et al., 2019; Harris et al., 1982), ▲ (Wu et al., 2020)
Number of saccades	▲ (Szewczyk et al., 2020), ▼ (Kataoka et al., 2011; Marandi et al., 2018; Nakayama et al., 2002)
NNI	▲ (Di Nocera et al., 2006, 2007)
Duration of fixations	▲ (Kataoka et al., 2011; Marandi et al., 2018; Marquart et al., 2015), ▼ (De Rivecourt et al., 2008)

Note. NNI = Nearest Neighbor Index.

ocular metrics for cognitive workload in which elevated mental load reduced saccade duration and increased fixation duration in young (21–42 years old) and older (51–70 years old) participants, but the saccade frequency went down only in the young participants (Marandi et al., 2018). Researchers reported mixed outcomes when studying fixation duration and mental workload as well; while a decrease in fixation duration was noticed in simulated instrument flight (De Rivecourt et al., 2008), this was not the case for driving studies (Marquart et al., 2015). Higher fixation duration indicates that more time is needed for information processing. NNI—a measure of randomness in visual search—may decrease or increase when cognitive load rises. The direction of change depends on the visual search which could become more random (higher NNI) or stereotyped (lower NNI) when mental workload rise (Camilli et al., 2008; Di Nocera et al., 2006; Di; Diaz-Piedra et al., 2019; Di Stasi et al., 2016; Harris et al., 1982; Wu et al., 2020). In most cases, a positive correlation between NNI (Di Nocera et al., 2006, 2007), as well as gaze entropy (Di Stasi et al., 2016; Wu et al., 2020) with the workload, was reported. In this study, NNI was lower at the start of shifts than middle and end of shifts. We believe that results follow the interpretation of Harris et al. on the effect of mental workload on visual search (Harris et al., 1982). Nurses at the start of shifts narrowed visual attention instead of exploring the visual environment to manage workload; once the mental workload of shift handoff concluded (mid-shifts and end of shifts), nurses used more time to visually assess the physical environment. This means that NNI decreases by increasing mental workload.

In this naturalistic study, pupil diameter had an upward trend during entire shifts. Nurses' pupil diameter was significantly greater at the end of shifts than at the start and middle of shifts. We did not anticipate this trend as pupil diameter is positively correlated with the mental workload. This metric is modulated by the autonomic nervous system and it contracts or dilates under different conditions including cognitive workload, emotional arousal, stress, sleepiness, changes in the level of illumination or caffeine intake (Abokyi et al., 2017; Skaramagkas et al., 2021). Sleepiness, which is negatively correlated with pupil diameter (Daguet et al., 2019), gradually increases during working shifts (Geiger-Brown et al., 2012) and nurses deploy coping strategies to maintain their alertness during shifts. Some activities, such as drinking coffee (Pélissier et al., 2020), could increase pupil dilation for up to 90 min (Abokyi et al., 2017). A prior study reported that pupil diameter does not reflect overall workload comparison between periods with different levels of mental workload (Schulz et al., 2011). In the current study, the computed eye movement metrics of 12-hour shifts were analyzed when day and night shifts workload were examined. For handoff periods, an hour of eye movements at the start shifts was compared with an hour of mid-shifts and end shifts. The significant difference between NNI, number of fixations, number of saccades, and duration of saccades in handoff periods could be indicative of the potential of fixational and saccadic metrics in naturalistic workload assessment.

Correlation between Stress and ICU Nurses' Eye Movements (Hypothesis #3)

Our findings revealed that the number of fixations and gaze entropy are positively associated with higher stress levels. According to attentional control theory, anxiety adversely affects a person's ability to focus and tune out distractions, which results in reduced efficiency and performance (Caviola et al., 2017). In other words, fixation instability increased the tendency to explore the surrounding environment and to look at different stimuli which increased the number of fixations as well as visual entropy (randomness in visual search). Our results on the number of fixations and entropy are consistent with the anticipated effect of stress on nurses' eye movements. NNI is similar to entropy and increases when the visual search becomes more random, but we found a positive non-significant correlation between stress and NNI. This could be associated with the metric's degree of sensitivity to acute stress; further investigation is needed to explore the underlying reasons.

We did not expect pupil diameter to decrease with higher levels of stress; this could be associated with specific stressful activities that were performed in rooms with a lower level of lights such as checking patients in dark rooms in night shifts. Furthermore, results show that stress reduced the number and duration of saccades which could be linked with more transitions from fixation to another fixation. Work is in progress to analyze the eye tracking videos to shed light on activities that were linked with high levels of stress in night shifts.

Limitations

While this study addresses a gap in investigating workload and stress using eye movement, several confounding factors might have affected physiological responses of nurses in the uncontrolled naturalistic ICU environment, including individual differences, variability in clinical expertise, or the sequence of nursing activities. First, nurses might apply different coping strategies to deal with stress and cognitive workload (Hancock & Matthews, 2019). This may change the intensity and sensitivity of physiological responses. For example, experts are better at managing cognitive load and have a lower workload compared to novices (Castner et al., 2020; Szulewski et al., 2015; Harris et al., 1982) which leads to different visual search patterns. Second, the hysteretic effect on the cognitive workload of nurses: it is known that performing back-to-back tasks with different levels of difficulties could affect the perception of the cognitive work of participants (Hancock & Matthews, 2019); experienced nurses may follow specific orders of tasks as a coping strategy to manage stress. Different types of cognitive workload such as time pressure and dual-tasking could induce different visual search behavior (Kataoka et al., 2011); night shift and day shift could be different in this regard. Also, nurse roles vary (e.g., floor nurse, float nurse, and nurse manager) with different levels of activities, experience, workloads, and stressors that should be included in the workload assessment. Third, given the lack of previous similar studies, a power calculation was not possible. Despite the importance of this limitation, conducting naturalistic studies with eye tracking equipment that is somewhat intrusive to cover the entire shift proposes serious challenges (e.g., changing batteries may introduce interruptions to nursing tasks, wearability issues for long-duration studies) and reduces willingness to participate. Fourth, we did not measure the amount of lighting during different shifts. To filter out the effect of light as an extraneous variable on the pupil diameter, metrics such as the Index of Cognitive Activity (ICA) could be employed in future research (Matthews et al., 2015). Finally, our study does not report consensus between eye movement measures for workload or stress with qualitative or self-reported instruments due to the lack of those outcomes. Studies of convergent validity of such metrics compared to self-reported (or perceived) measures of workload and stress are warranted to contribute to setting ocular metrics threshold for workload and stress measurements in clinicians.

Future studies may employ a quasi-experimental design to control for some of these potentially confounding variables. Furthermore, a combination of eye movement-based workload assessment and qualitative approach would help researchers to identify specific nursing activities, ICU events, or ICU technologies that elicit a high level of cognitive load; to do that, one or more nurse researchers should annotate recorded videos using special software (e.g., ELAN) based on a validated system such as nursing activities score (NAS) to (1) pinpoint nursing activities bases on one of 23 items of NAS, (2) tag employed ICU technologies, and (3) mark ICU events such as end-of-life events or code blue; then, that rich dataset could be utilized for further quantitative and qualitative analyses. Using these approaches, the correlation of NAS classification of activities such as medication administration, monitoring and titration, and employed technologies could be assessed with nurses' cognitive load. The outcome will be useful to modify ICU nursing workflow or to detect usability problems of ICU technologies. Interviewing nurses as a complementary method will be beneficial to receive feedback from nurses. The explained approach will address one of the limitations of the current study that did not explore the impact of specific nursing activities or ICU events on nurses' cognitive load.

CONCLUSION

Eye tracking research has shown promise in estimating workload in simulated settings but the efficacy of this method in complex and dynamic clinical work environments (e.g., ICU) is a general research gap. A naturalistic study was conducted to evaluate ICU nurses' workload using ocular metrics to investigate correlations with stress. Our findings suggest that while workload did not differ between day and night shifts, the start of shifts resulted in significantly higher mental workload measures compared to middle and end of the shift. In addition, several eye movement metrics showed strong correlations with stress (measured in accordance with Baevsky's stress index) which may provide preliminary evidence supporting the usage of eye tracking to measure stress.

This study offers data-based insights into the variability of workload during nursing shifts and contributes to equivalence in aggregated workloads when comparing day shifts and night shifts. These findings can help nurse leaders quantify the complexity of care demands to estimate staffing needs, contribute to patient safety, and support the nurse workforce. In the future, temporal trends in workload could be detected by eye-tracking devices and communicated in real-time to nurse leaders to address complex care delivery needs for critically-ill patients. However, more work is warranted to evaluate the accuracy and validity of eye movement measures in measuring workload and stress.

ACKNOWLEDGEMENTS

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was partially funded by a Dyer Fellowship awarded to Dr. Farzan Sasangohar at Houston Methodist Hospital. The authors also thank Jacob M. Kolman, MA, ISMPP CMPP™, senior scientific writer of the Houston Methodist Center for Outcomes Research and research associate of the Department of Industrial and Systems Engineering at Texas A&M University, for language and formatting feedback.

KEY POINTS

- The cumulative mental workload of critical care nurses using ocular metrics is roughly equivalent between day and night shifts; however, stress index was higher during the night shift.
- Nurses experience a higher level of mental workload during the start of shift handoff period compared with mid-shift and the end of shifts.
- Number of fixations and entropy positively correlate with the stress index while duration of saccades and pupil diameter negatively correlate with the stress index.

ORCID iDs

Farzan Sasangohar https://orcid.org/0000-0001-9962-5470

Denny Yu https://orcid.org/0000-0002-5083-1713

Valerie Danesh https://orcid.org/0000-0002-2078-2578

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

Abokyi, S., Owusu-Mensah, J., & Osei, K. A. (2017). Caffeine intake is associated with pupil dilation and enhanced accommodation. *Eye*, *31*(4), 615–619. https://doi.org/10.1038/eye.2016.288

Acerbi, G., Rovini, E., Betti, S., Tirri, A., Rónai, J. F., Sirianni, A., Agrimi, J., Eusebi, L., & Cavallo, F. (2017). A wearable system for stress detection through physiological data analysis. In F. Cavallo, V. Marletta, A. Monteriù, & P. Siciliano (Eds.), *Lecture Notes in Electrical Engineering* (426, pp. 31-50). Springer, Cham. https://doi.org/10.1007/978-3-319-54283-6_3.

Ahmadi, N., Sasangohar, F., Nisar, T., Danesh, V., Larsen, E., Sultana, I., & Bosetti, R. (2022). Quantifying occupational stress in intensive care unit nurses: An applied naturalistic study of correlations among stress, heart rate, electrodermal activity, and skin temperature. *Human Factors*, *64*(1), 159-172. https://doi.org/10.1177/00187208211040889.

Aiken, L. H., Clarke, S. P., & Sloane, D. M. (2002). Hospital staffing, organization, and quality of care: Cross-national findings. *International Journal for Quality in Health Care*, *14*(1), 5–13. https://doi.org/10.1093/intqhc/14.1.5

Amaravadi, R. K., Dimick, J. B., Pronovost, P. J., & Lipsett, P. A. (2000). ICU nurse-to-patient ratio is associated with complications and resource use after esophagectomy. *Intensive Care Medicine*, *26*(12), 1857–1862. https://doi.org/10.1007/s001340000720

Appel, T., Scharinger, C., Gerjets, P., & Kasneci, E. (2018). Cross-subject workload classification using pupil-related measures. *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, *June 2018*, 1-8, Article No. 4. https://doi.org/10.1145/3204493.3204531.

Appel, T., Sevcenko, N., Wortha, F., Tsarava, K., Moeller, K., Ninaus, M., Kasneci, E., & Gerjets, P. (2019). Predicting cognitive load in an emergency simulation based on behavioral and physiological measures. *2019 International Conference on Multimodal Interaction*, *October 2019*, 154-163. https://doi.org/10.1145/3340555.3353735.

Armitage, G. (2009). Human error theory: Relevance to nurse management. *Journal of Nursing Management*, *17*(2), 193–202. https://doi.org/10.1111/j.1365-2834.2009.00970.x

Armstrong, E., de Waard, M. C., Heymans, M. W.H.-J. S., Girbes, A. R. J., Spijkstra, J. J., Girbes, A. R. J., & Spijkstra, J. J (2015). Using nursing activities score to assess nursing workload on a medium care unit. *Anesthesia & Analgesia*, *121*(5), 1274–1280. https://doi.org/10.1213/ANE.0000000000000968

Baevsky, R. M., & Berseneva, A. P (2008). Methodical recommendations use kardivar system for determination of the stress level and estimation of the body adaptability standards of measurements and physiological interpretation. https://pdfs.semanticscholar.org/74a2/92bfafca4fdf1149d557348800fcc1b0f33b.pdf

Brunyé, T. T., Drew, T., Weaver, D. L., & Elmore, J. G. (2019). A review of eye tracking for understanding and improving diagnostic interpretation. *Cognitive Research: Principles and Implications*, *4*, 7. https://doi.org/10.1186/s41235-019-0159-2.

Byrne, A. J., Oliver, M., Bodger, O., Barnett, W. A., Williams, D., Jones, H., & Murphy, A. (2010). Novel method of measuring the mental workload of anaesthetists during clinical practice. *British Journal of Anaesthesia*, *105*(6), 767–771. https://doi.org/10.1093/bja/aeq240

Camilli, M., Nacchia, R., Terenzi, M., & Di Nocera, F. (2008). ASTEF: A simple tool for examining fixations. *Behavior Research Methods*, *40*(2), 373–382. https://doi.org/10.3758/brm.40.2.373

Carayon, P., & Gürses, A. P. (2005). A human factors engineering conceptual framework of nursing workload and patient safety in intensive care units. *Intensive and Critical Care Nursing*, *21*(5), 284–301. https://doi.org/10.1016/j.iccn.2004.12.003

Castner, N., Appel, T., Eder, T., Richter, J., Scheiter, K., Keutel, C., Hüttig, F., Duchowski, A., & Kasneci, E. (2020). Pupil diameter differentiates expertise in dental radiography visual search. *Plos One*, *15*(5), e0223941. https://doi.org/10.1371/journal.pone.0223941

Caviola, S., Carey, E., Mammarella, I. C., & Szucs, D. (2017). Stress, time pressure, strategy selection and math anxiety in mathematics: A review of the literature. *Frontiers in Psychology*, *8*, 1488. https://doi.org/10.3389/fpsyg.2017.01488.

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: danger and experience. *Perception*, *27*(8), 951–964. https://doi.org/10.1068/p270951

Cho, S.-H., Ketefian, S., Barkauskas, V. H., & Smith, D. G. (2003). The effects of nurse staffing on adverse events, morbidity, mortality, and medical costs. *Nursing Research*, *52*(2), 71–79. https://doi.org/10.1097/00006199-200303000-00003

Daguet, I., Bouhassira, D., & Gronfier, C. (2019). Baseline pupil diameter is not a reliable biomarker of subjective sleepiness. *Frontiers in Neurology*, *10*, 108. https://doi.org/10.3389/fneur.2019.00108.

Davidson, J. E., Aslakson, R. A., Long, A. C., Puntillo, K. A., Kross, E. K., Hart, J., Cox, C. E., Wunsch, H., Wickline, M. A., Nunnally, M. E., Netzer, G., Kentish-Barnes, N., Sprung, C. L., Hartog, C. S., Coombs, M., Gerritsen, R. T., Hopkins, R. O., Franck, L. S., Skrobik, Y., Kon, A. A., Scruth, E. A., Harvey, M. A., Lewis-Newby, M., White, D. B., Swoboda, S. M., Cooke, C. R., Levy, M. M., Azoulay, E., & Curtis, J. R. (2017). Guidelines for family-centered care in the neonatal, pediatric, and adult ICU. *Critical Care Medicine*, *45*(1), 103–128. https://doi.org/10.1097/CCM.0000000000002169

Debergh, D. P., Myny, D., Van Herzeele, I., Van Maele, G., Miranda, D. R., & Colardyn, F. (2012). Measuring the nursing workload per shift in the ICU. *Intensive Care Medicine*, *38*(9), 1438–1444. https://doi.org/10.1007/s00134-012-2648-3

de Greef, T., Lafeber, H., van Oostendorp, H., & Lindenberg, J. (2009). Eye movement as indicators of mental workload to trigger adaptive automation. In D. D. Schmorrow, I. V. Estabrooke, & M. Grootjen (eds), *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*. Springer (pp. 219–228). https://doi.org/10.1007/978-3-642-02812-0_26

De Rivecourt, M., Kuperus, M. N., Post, W. J., & Mulder, L. J. M. (2008). Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight. *Ergonomics*, *51*(9), 1295–1319.

Diaz-Piedra, C., Rieiro, H., Cherino, A., Fuentes, L. J., Catena, A., & Di Stasi, L. L. (2019). The effects of flight complexity on gaze entropy: An experimental study with fighter pilots. *Applied Ergonomics*, *77*(May 2019), 92-99. https://doi.org/10.1016/j.apergo.2019.01.012.

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A random glance at the flight deck: Pilots' scanning strategies and the real-time assessment of mental workload. *Journal of Cognitive Engineering and Decision Making*, *1*(3), 271–285. https://doi.org/10.1518/155534307X255627.

Di Nocera, F., Terenzi, M., & Camilli, M. (2006). Another look at scanpath: Distance to nearest neighbour as a measure of mental workload. In *Developments in Human Factors in Transportation, Design, and Evaluation*. Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co.,, 295–303.

Di Stasi, L. L., Rieiro, H., Carrión, J. M. S., Berrido, M. M., Olivares, G., Catena, A., & Catena, A. (2016). Gaze entropy reflects surgical task load. *Surgical Endoscopy*, *30*(11), 5034–5043. https://doi.org/10.1007/s00464-016-4851-8

Donchin, Y., & Seagull, F. J. (2002). The hostile environment of the intensive care unit. *Current Opinion in Critical Care*, *8*(4), 316–320. https://doi.org/10.1097/00075198-200208000-00008

Ehmke, C., & Wilson, S. (2007). Identifying web usability problems from eyetracking data. 119–128. https://doi.org/10.14236/ewic/hci2007.12

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 65–84. https://doi.org/10.1518/001872095779049499

Foy, H. J., & Chapman, P. (2018). Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation. *Applied Ergonomics*, *73*(November 2018), 90-99. https://doi.org/10.1016/j.apergo.2018.06.006.

Friesen, M. A., White, S. V., & Byers, J. F. (2008). Handoffs: Implications for nurses. In R. G. Hughes (Ed.), *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*. Agency for Healthcare Research and Quality (US). http://www.ncbi.nlm.nih.gov/books/NBK2649/

Geiger-Brown, J., Rogers, V. E., Trinkoff, A. M., Kane, R. L., Bausell, R. B., & Scharf, S. M. (2012). Sleep, sleepiness, fatigue, and performance of 12-hour-shift nurses. *Chronobiology International*, *29*(2), 211–219. https://doi.org/10.3109/07420528.2011.645752

Grundgeiger, T., Sanderson, P., MacDougall, H. G., & Venkatesh, B. (2010). Interruption management in the intensive care unit: Predicting resumption times and assessing distributed support. *Journal of Experimental Psychology. Applied*, *16*(4), 317–334. https://doi.org/10.1037/a0021912

Hancock, P. A., & Matthews, G. (2019). Workload and performance: Associations, insensitivities, and dissociations. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(3), 374–392. https://doi.org/10.1177/0018720818809590

Harezlak, K., & Kasprowski, P. (2018). Application of eye tracking in medicine: A survey, research issues and challenges. *Computerized Medical Imaging and Graphics*, *65*(April 2018), 176-190. https://doi.org/10.1016/j.compmedimag.2017.04.006.

Harris, R. L., Glover, B. L., & Spady, A. A. (1986). *Analytic techniques of pilot scanning behavior and their application (Technical Paper No. 2525)*. Hampton, VA: NASA Langley Research Center.

Harris, R. L., Tole, J. R., Stephens, A. T., Ephrath, A. R., & Young, L. R. (1982). Visual scanning behavior and pilot workload. *Aviation, Space, and Environmental Medicine*, *53*(11), 1067–1072.

Henneman, E. A., Roche, J. P., Fisher, D. L., Cunningham, H., Reilly, C. A., Nathanson, B. H., & Henneman, P. L. (2010). Error identification and recovery by student nurses using human patient simulation: Opportunity to improve patient safety. *Applied Nursing Research*, *23*(1), 11–21. https://doi.org/10.1016/j.apnr.2008.02.004

Herrick, H., Weinberg, D., Cecarelli, C., Fishman, C. E., Newman, H., den Boer, M. C., Martherus, T., Katz, T. A., Nadkarni, V., Te Pas, A. B., & Foglia, E. E. (2020). Provider visual attention on a respiratory function monitor during neonatal resuscitation. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, *105*(6), 666–668. https://doi.org/10.1136/archdischild-2020-319291

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190–1192. https://doi.org/10.1126/science.143.3611.1190

Higgs, J., Jones, M. A., Loftus, S., & Christensen, N. (2008). *Clinical reasoning in the health professions E-book*. Elsevier Health Sciences.

Hofmaenner, D. A., Herling, A., Klinzing, S., Wegner, S., Lohmeyer, Q., Schuepbach, R. A., & Buehler, P. K. (2021). Use of eye tracking in analyzing distribution of visual attention among critical care nurses in daily professional life: An observational study. *Journal of Clinical Monitoring and Computing*, *35*, 1511-1518. https://doi.org/10.1007/s10877-020-00628-2.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de, J. W.. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Hoonakker, P., Carayon, P., Gurses, A. P., Brown, R., Khunlertkit, A., McGuire, K., & Walker, J. M. (2011). Measuring workload of ICU nurses with a questionnaire survey: The NASA task load index (TLX). *IIE Transactions on Healthcare Systems Engineering*, *1*(2), 131–143. https://doi.org/10.1080/19488300.2011.609524

Hyönä, J., Tommola, J., & Alaja, A.-M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology Section A*, *48*(3), 598–612. https://doi.org/10.1080/14640749508401407

Jeffri, N. F. S., & Awang Rambli, D. R. (2021). A review of augmented reality systems and their effects on mental workload and task performance. *Heliyon*, *7*(3), e06277. https://doi.org/10.1016/j.heliyon.2021.e06277

Kataoka, J., Sasaki, M., & Kanda, K. (2011). Effects of mental workload on nurses' visual behaviors during infusion pump operation. *Japan Journal of Nursing Science*, *8*(1), 47–56. https://doi.org/10.1111/j.1742-7924.2010.00158.x

Kiekkas, P., Sakellaropoulos, G. C., Brokalaki, H., Manolis, E., Samios, A., Skartsani, C., & Baltopoulos, G. I. (2008). Association between nursing workload and mortality of intensive care unit patients. *Journal of Nursing Scholarship*, *40*(4), 385–390. https://doi.org/10.1111/j.1547-5069.2008.00254.x

Kim, S.-H., Pinker, E., Rimar, J., & Bradley, E. (2017). Refining workload measure in hospital units: From census to acuity-adjusted census in intensive care units. In *USC Marshall School of Business Working Paper*.

Koch, S. H., Weir, C., Haar, M., Staggers, N., Agutter, J., Görges, M., & Westenskow, D. (2012). Intensive care unit nurses' information needs and recommendations for integrated displays to improve nurses' situation awareness. *Journal of the American Medical Informatics Association*, *19*(4), 583–590. https://doi.org/10.1136/amiajnl-2011-000678

Koh, R. Y. I., Park, T., Wickens, C. D., Ong, L. T., & Chia, S. N. (2011). Differences in attentional strategies by novice and experienced operating theatre scrub nurses. *Journal of Experimental Psychology. Applied*, *17*(3), 233–246. https://doi.org/10.1037/a0025171

Kwiecień-Jaguś, K., Mędrzycka-Dąbrowska, W., Czyż-Szypenbeil, K., Lewandowska, K., & Ozga, D. (2019). The use of a pedometer to measure the physical activity during 12-hour shift of ICU and nurse anaesthetists in Poland. *Intensive & Critical Care Nursing*, *55*(102750), 102750. https://doi.org/10.1016/j.iccn.2019.07.009

Larsen, E. P., Kolman, J. M., Masud, F. N., & Sasangohar, F. (2020). Ethical considerations when using a mobile eye tracker in a patient-facing area: lessons from an intensive care unit observational protocol. *Ethics & Human Research*, *42*(6), 2–13. https://doi.org/10.1002/eahr.500068

Law, B. H. Y., Cheung, P.-Y., van Os, S., Fray, C., & Schmölzer, G. M. (2020). Effect of monitor positioning on visual attention and situation awareness during neonatal resuscitation: A randomised simulation study. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, *105*(3), 285–291. https://doi.org/10.1136/archdischild-2019-316992

Lee, A., Cheung, Y. S. L., Joynt, G. M., Leung, C. C. H., Wong, W. T., & Gomersall, C. D. (2017). Are high nurse workload/staffing ratios associated with decreased survival in critically ill patients? A cohort study. *Annals of Intensive Care*, *746*(1), 46. https://doi.org/10.1186/s13613-017-0269-2

Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Frontiers in Human Neuroscience*, *13*, 57. https://doi.org/10.3389/fnhum.2019.00057.

Malacrida, R., Bomio, D., Matathia, R., Suter, P. M., & Perrez, M. (1991). Computer-aided self-observation psychological stressors in an ICU. *International Journal of Clinical Monitoring and Computing*, *8*(3), 201–205. https://doi.org/10.1007/BF01738893

Mandrick, K., Peysakhovich, V., Rémy, F., Lepron, E., & Causse, M. (2016). Neural and psychophysiological correlates of human performance under stress and high mental workload. *Biological*

Psychology, *121, Part A*(December 2016), 62-73. https://doi.org/10.1016/j.biopsycho.2016.10.002.

Marandi, R. Z., Madeleine, P., Omland, O., Vuillerme, N., & Samani, A. (2018). Reliability of oculometrics during a mentally demanding task in young and old adults. *IEEE Access*, *6*, 17500-17517. https://doi.org/10.1109/ACCESS.2018.2819211.

Marquard, J. L., Henneman, P. L., He, Z., Jo, J., Fisher, D. L., & Henneman, E. A. (2011). Nurses' behaviors and visual scanning patterns may reduce patient identification errors. *Journal of Experimental Psychology. Applied*, *17*(3), 247–256. https://doi.org/10.1037/a0025261

Marquart, G., Cabrall, C., & de Winter, J. (2015). Review of eye-related measures of drivers' mental workload. *Procedia Manufacturing*, *3*(2015), 2854-2861. https://doi.org/10.1016/j.promfg.2015.07.783.

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., & Abich, J. J. (2015). The psychometrics of mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(1), 125–143. https://doi.org/10.1177/0018720814539505

Menekse Dalveren, G. G., & Cagiltay, N. E. (2018). Insights from surgeons' eye-movement data in a virtual simulation surgical training environment: effect of experience level and hand conditions. *Behaviour & Information Technology*, *37*(5), 517-537. https://doi.org/10.1080/0144929X.2018.1460399.

Miranda, D. R., Nap, R., de Rijk, A., Schaufeli, W., & Iapichino, G.Group, the members of the T (2003). Nursing activities score. *Critical Care Medicine*, *31*(2), 374–382. https://doi.org/10.1097/01.CCM.0000045567.78801.CC

Moghadam, K. N., Chehrzad, M. M., Masouleh, S. R., Mardani, A., Maleki, M., Akhlaghi, E., & Harding, C (2021). Nursing workload in intensive care units and the influence of patient and nurse characteristics. *Nursing in Critical Care*, *26*(6), 425-431. https://doi.org/10.1111/nicc.12548.

Nakayama, M., Takahashi, K., & Shimizu, Y. (2002). The act of task difficulty and eye-movement frequency for the 'Oculo-motor indices'. *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, *March 2002*, 37-42. https://doi.org/10.1145/507072.507080.

Neill, D. (2011). Nursing workload and the changing health care environment: A review of the literature. *Administrative Issues Journal*, *1*(2), 13. https://doi.org/10.5929/2011.1.2.11

Nogueirade, L. D. S.S., Domingues, C. D. A., Poggetti, R. S., Sousa, R. M. C. D., & Sousade, R. M. C. (2014). Nursing workload in intensive care unit trauma patients: analysis of associated factors. *Plos One*, *9*(11), e112125. https://doi.org/10.1371/journal.pone.0112125

Oates, P., & Oates, R. (1996). Stress and work relationships in the neonatal intensive care unit: Are they worse than in the wards? *Journal of Paediatrics and Child Health*, *32*(1), 57–59. https://doi.org/10.1111/j.1440-1754.1996.tb01543.x

Padilha, K. G., de Sousa, R. M. C., Queijo, A. F., Mendes, A. M., & Miranda, D. R. (2008). Nursing Activities Score in the intensive care unit: Analysis of the related factors. *Intensive and Critical Care Nursing*, *24*(3), 197–204. https://doi.org/10.1016/j.iccn.2007.09.004

Pélissier, C., Cavelier, C., Vercherin, P., Roche, F., Patural, H., & Fontana, L. (2020). Vigilance and sleepiness in nurses working 12-hr shifts and their coping strategies. *Journal of Nursing Management*, Advance Online Publication. https://doi.org/10.1111/jonm.13233

Pickup, L., Wilson, J. R., Sharpies, S., Norris, B., Clarke, T., & Young, M. S. (2005). Fundamental examination of mental workload in the rail industry. *Theoretical Issues in Ergonomics Science*, *6*(6), 463–482. https://doi.org/10.1080/14639220500078021

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*, 257–266. https://doi.org/10.1109/tsmc.1983.6313160

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology. Applied*, *6*(1), 31–43. https://doi.org/10.1037//1076-898x.6.1.31

Reis Miranda, D., de Rijk, A., & Schaufeli, W. (1996). Simplified therapeutic intervention scoring system. *Critical Care Medicine*, *24*(1), 64–73. https://doi.org/10.1097/00003246-199601000-00012

Romano, J. L., Garcia, P. C., Silva, D. V., Moura, B. R. S., & Nogueirade, L. S. (2019). Type of admission and nursing workload of critical patients: A cross-sectional study. *Nursing in Critical Care*, *24*(6), 387–391. https://doi.org/10.1111/nicc.12408.

Schulz, C. M., Schneider, E., Fritz, L., Vockeroth, J., Hapfelmeier, A., Wasmaier, M., Kochs, E. F., & Schneider, G. (2011). Eye tracking for assessment of workload: A pilot study in an anaesthesia simulator environment. *British Journal of Anaesthesia*, *106*(1), 44–50. https://doi.org/10.1093/bja/aeq307

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Skaramagkas, V., Giannakakis, G., Ktistakis, E., Manousos, D., Karatzanis, I., Tachos, N., Tripoliti, E. E., Marias, K., Fotiadis, D. I., & Tsiknakis, M. (2021). Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Reviews in Biomedical Engineering*, *PP*(99), 1–1. https://doi.org/10.1109/RBME.2021.3066072

Stafseth, S. K., Solms, D., & Bredal, I. S. (2011). The characterisation of workloads and nursing staff allocation in intensive care units: A descriptive study using the nursing activities score for the first time in Norway. *Intensive and Critical Care Nursing*, *27*(5), 290–294. https://doi.org/10.1016/j.iccn.2011.07.003

Stubbings, L., Chaboyer, W., & McMurray, A. (2012). Nurses' use of situation awareness in decision-making: an integrative review. *Journal of Advanced Nursing*, *68*(7), 1443–1453. https://doi.org/10.1111/j.1365-2648.2012.05989.x

Szewczyk, G., Spinelli, R., Magagnotti, N., Tylek, P., Sowa, J., Rudy, P., & Gaj-Gielarowiec, D. (2020). The mental workload of harvester operators working in steep terrain conditions. *Silva Fennica*, *54*(3), 10355. https://doi.org/10.14214/sf.10355.

Szulewski, A., Roth, N., & Howes, D. (2015). The use of task-evoked pupillary response as an objective measure of cognitive load in novices and trained physicians. *Academic Medicine*, *90*(7), 981–987. https://doi.org/10.1097/ACM.0000000000000677

Tarnow-Mordi, W., Hau, C., Warden, A., & Shearer, A. (2000). Hospital mortality in relation to staff workload: A 4-year study in an adult intensive-care unit. *Lancet*, *356*(9225), 185–189. https://doi.org/10.1016/S0140-6736(00)02478-8

Tien, T., Pucher, P. H., Sodergren, M. H., Sriskandarajah, K., Yang, G.-Z., & Darzi, A. (2014). Eye tracking for skills assessment and training: A systematic review. *Journal of Surgical Research*, *191*(1), 169–178. https://doi.org/10.1016/j.jss.2014.04.032

Weinberg, D. D., Newman, H., Fishman, C. E., Katz, T. A., Nadkarni, V., Herrick, H. M., & Foglia, E. E. (2020). Visual attention patterns of team leaders during delivery room resuscitation. *Resuscitation*, *147*, 21-25. https://doi.org/10.1016/j.resuscitation.2019.12.008.

Wilbanks, B. A., & McMullan, S. P. (2018). A review of measuring the cognitive workload of electronic health records. *CIN: Computers, Informatics, Nursing*, *36*(12), 579–588. https://doi.org/10.1097/CIN.0000000000000469

Wu, C., Cha, J., Sulek, J., Sundaram, C. P., Wachs, J., Proctor, R. W., & Yu, D. (2021). Sensor-based indicators of performance changes between sessions during robotic surgery training. *Applied Ergonomics*, *90*(103251), 103251. https://doi.org/10.1016/j.apergo.2020.103251

Wu, C., Cha, J., Sulek, J., Zhou, T., Sundaram, C. P., Wachs, J., & Yu, D. (2020). Eye-tracking metrics predict perceived workload in robotic surgical skills training. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *62*(8), 1365–1386. https://doi.org/10.1177/0018720819874544

Yang, Y., McDonald, M., & Zheng, P. (2012). Can drivers' eye movements be used to monitor their performance? A case study. *IET Intelligent Transport Systems*, *6*(4), 444–452. https://doi.org/10.1049/iet-its.2012.0008

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics*, *58*(1), 1–17. https://doi.org/10.1080/00140139.2014.956151

Yu, D., Lowndes, B., Thiels, C., Bingener, J., Abdelrahman, A., Lyons, R., & Hallbeck, S. (2016). Quantifying intraoperative workloads across the surgical team roles: Room for better balance? *World Journal of Surgery*, *40*(7), 1565–1574. https://doi.org/10.1007/s00268-016-3449-6

Yu, F., Narayanan, A., Mackay, L., Ward, K., King, A., & Smith, M. (2020). Describing objectively measured intensive care nurses' physical work activity behavioural patterns during a 12-hr shift. *Journal of Clinical Nursing*, *29*(21–22), 4331–4342. https://doi.org/10.1111/jocn.15470

Zehnder, E. C., Schmölzer, G. M., van Manen, M., & Law, B. H. Y. (2021). Using eye-tracking augmented cognitive task analysis to explore healthcare professionals' cognition during neonatal resuscitation. *Resuscitation*, *6*(100119), 100119. https://doi.org/10.1016/j.resplu.2021.100119

Zhang, J. Y., Liu, S. L., Feng, Q. M., Gao, J. Q., & Zhang, Q. (2017). Correlative evaluation of mental and physical workload of laparoscopic surgeons based on surface electromyography and eye-tracking signals. *Scientific Reports*, *7*(1), 11095–11097. https://doi.org/10.1038/s41598-017-11584-4

Nima Ahmadi is a postdoctoral fellow in the Center of Outcomes Research, Houston Methodist. He earned his PhD in Engineering Management (2020) from Western New England University, MA, USA.

Farzan Sasangohar is an assistant professor in the Department of Industrial and Systems Engineering at Texas A&M University as well as the Center of Outcomes Research and Department of Surgery at Houston Methodist. He received his PhD in Industrial Engineering (2015) from the University of Toronto.

Jing Yang is a Ph.D. student at Purdue University. She has a MS in industrial and operations engineering from the University of Michigan.

Denny Yu is an assistant professor in the School of Industrial Engineering at Purdue University. He received his PhD in Industrial and Operations Engineering from the University of Michigan in 2014.

Valerie Danesh is a research scientist in the Center for Applied Health Research at Baylor Scott & White Research Institute and a faculty affiliate of The University of Texas at Austin School of Nursing. She received her PhD in Nursing Research from the University of Central Florida in 2015.

Steven Klahn is the Virtual ICU Nursing Director at Houston Methodist Hospital. He received his BSN in nursing from University of Texas Health Science Center at San Antonio in 1999.

Faisal Masud is the Mary A. and M. Samuel Daffin, Sr. Centennial Chair in Anesthesia and Critical Care and Medical Director of Center for Critical Care at Houston Methodist Hospital. He received his M.B. from Rawalpindi Medical College in 1988.

Date received: July 22, 2021
Date accepted: February 8, 2022