EasyChair Preprint

№ 8155

Preliminary Findings of a Multimodal Sensor System for Measuring Surgeon Cognitive Workload

Ravi Naik, Kaizhe Jin, Alexandros Kogkas, Hutan Ashrafian,
Ara Darzi and George Mylonas

May 31, 2022

Preliminary findings of a multimodal sensor system for measuring surgeon cognitive workload

Ravi Naik[1, 2]*, Kaizhe Jin*[1,2] *Alexandros Kogkas*[1,2]*, Hutan Ashrafian,*[1] *Ara Darzi*[1, 2]*, George Mylonas*[1,2]

[1]*Department of Surgery and Cancer, St Mary's Hospital, Imperial College London, UK,*

[2]*Hamlyn Centre for Robotic Surgery, Institute of Global Health Innovation, Imperial College London, UK*
ravi.naik15@imperial.ac.uk

INTRODUCTION

The operating room represents a high-risk environment centred around the safe and efficient delivery of patient care. It is a complex ecosystem that encompasses many factors including communication within the multidisciplinary surgical team often led by the operating surgeon as well as execution of precise technical surgical skill. These factors are associated with the mental or cognitive workload (CWL) of the surgeon.

CWL, also described as the mental effort exerted while undertaking a task, is a construct derived from the cognitive load theory first described in the eighties during problem solving exercises [1]. There has been a growing emphasis on the measurement of CWL since then on individuals working in high-stake environments such as aviation [2]. Measurement of CWL in surgery is moving from solitary traditional subjective measures such as the Surgical Task Load Index (SURG-TLX) to objective measurement of physiological parameters secondary to changes in the CWL of the surgeon which are less exposed to subjective bias [3]. These have included heart rate variability (HRV), pupil metrics, electromyography (EMG), electroencephalography (EEG), skin conductance and functional near-infrared spectroscopy (fNIRS). More recently, there is increasing evidence to demonstrate the use of multiple sensors, or a multimodal sensor system designed to measure CWL with greater accuracy [4].

The aim of this paper is to demonstrate the use of a pilot synchronised system of multiple sensors to measure the real-time cognitive workload of surgeons in a simulated setting to demonstrate a proof of concept and to discuss the early findings.

MATERIALS AND METHODS

Ethics: Ethical approval was obtained from the research governance and integrity team at Imperial College (No: 20IC6361).

Participants: Five surgical residents from the academic surgical department at Imperial College London were recruited following study explanation and obtaining informed consent.

Surgical task: Laparoscopic peg transfer, a component of the Fundamentals of Laparoscopic surgery curriculum, was performed by participants twice per task condition. This was undertaken with standard laparoscopic equipment using the Olympus Elite system with a fixed 2D zero-degree 10mm laparoscope.

Task conditions: Participants were asked to perform the skill under four separate conditions in randomised order to minimise order effects bias and included 1. A control with no distractions (CS), 2. Mental arithmetic during the task (MA) which comprised of subtracting serial sevens from one thousand, 3. Noise from a recurring hospital bleep (ND) and 4. Dual distractions comprising of conditions 2 and 3 simultaneously (DD).

Subjective measures: Participants were asked to complete the SURG-TLX questionnaire tool after each task which consists of a multidimensional scale and pair-wise comparisons of cognitive domains including mental demands, physical demands, temporal demands, task complexity, situational stress and distractions.

Objective measures: Time to complete each task, number of pegs dropped, peg handling errors and miscalculations were recorded under each task condition and logged in real-time using a time stamp.

Physiological measures: The following physiological sensors were used: fNIRS (Artinis Brite24 2 x 11 system configured to the prefrontal cortex), EEG (TMSi Mobita 32 channel EEG system), eye tracker (Pupil Labs Pupil Core), photoplethysmography (PPG), galvanic skin response (GSR), electrocardiogram (ECG), electromyography (EMG) and skin temperature (Consensys Shimmer3 units). The arrangement of the wearable sensors is demonstrated in Figure 1A. Synchronisation of sensor data was undertaken using Lab Streaming Layer (LSL) and a bespoke graphical user interface was developed to facilitate recording.

Experimental protocol: Participants were consented and familiarised with the system. Participants were fitted with all sensors which were calibrated. An instructional video was shown, and participants were given five minutes to practice using the setup. Baseline recordings were then taken before each task condition. Participants then completed each task in a randomised order and completed the SURG-TLX questionnaire following each task. Preliminary analysis of SURG-TLX scores, pupil data, HRV and EEG is being undertaken with the aim to analyse other physiological measures in due course.

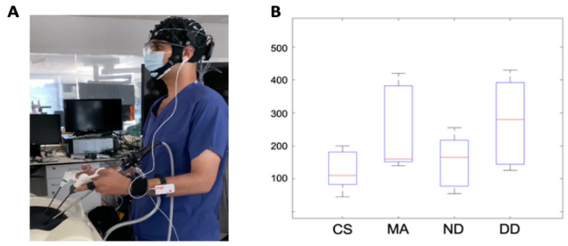

Figure 1A. Demonstration of the multimodal sensor network in use during a simulated laparoscopic peg transfer task. fNIRS, EEG, PPG, GSR sensors can be seen in use together with a wearable eye tracker. **Figure 1B.** Box plot to demonstrate distribution of overall adjusted SURG-TLX scores for each task for participants as a collective.

RESULTS

Table 1 and Figure 1B summarise the technical performance and subjective workload feedback results respectively. Participants subjectively recorded that task conditions MA and DD were significantly ($p < 0.05$) more difficult when compared to the control (CS) based on overall SURG-TLX scores ($p = 0.0071$ and $p = 0.0284$ respectively) and took longer to complete.

Figure 2 demonstrates the mean blink rate, mean single channel EEG recording and mean HRV (RMSSD) of participants during the CS and MA task. Although there was an increase in mean blink rate, single channel EEG and HRV, these were not statistically significant ($p = 0.067$, 0.254 and 0.149 respectively). There was a weak correlation between overall SURG-TLX score and time to complete task (Pearsons correlation coefficient = 0.347, $p = 0.325$). There was a strong positive correlation between raw mental demand scores and blink rate for tasks CS and MA (Pearsons correlation coefficient = 0.855, $p = 0.007$).

	Task 1 (CS)		Task 2 (MA)			Task 3 (ND)		Task 4 (DD)		
Participant	Duration (secs)	Pegs dropped	Duration (secs)	Pegs dropped	Errors in subtraction	Duration (secs)	Pegs dropped	Duration (secs)	Pegs dropped	Errors in subtraction
1	163.63	0	327.81	1	6	142.47	1	226.29	1	3
2	284.09	2	289.39	3	2	222.51	1	195	0	1
3	197.62	0	268.52	1	5	344.47	4	383.09	5	12
4	232.16	3	286.47	3	3	176.39	0	198.68	2	1
5	186.95	2	200.42	1	4	180.22	3	202.96	0	1
Mean	*212.89*	*1.4*	*274.522*	*1.8*	*4*	*213.212*	*1.8*	*241.204*	*1.6*	*3.6*

Table 1. Summary of task technical performance.

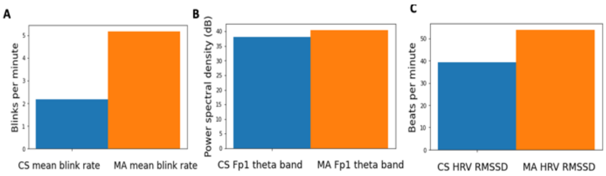

Figure 2. A) Mean blink rate of participants over task CS and MA. B) Mean HRV between task CS and MA. C) Mean change in EEG (Theta band) between task CS and MA.

DISCUSSION

The surgical task conditions designed in this pilot study successfully increased CWL when measured subjectively. There was successful synchronisation of sensors although signal loss of eye tracking and EEG data for one participant. This study demonstrates the integration and synchronisation of the physiological data required to measure CWL using LSL. Initial findings from this study support the use of heart rate variability and blink rate as measures of CWL.

Preliminary results have demonstrated a positive relationship between blink rate and subjective task difficulty (CS vs MA) in keeping with current literature [5]. A single EEG channel (Fp1) was selected for preliminary analysis as this was thought to best represent activity of the prefrontal cortex, and was also found to be higher in the MA task.

The current sample size is recognised as a limitation and a larger number of participants of the same skill level is required to establish significant correlations between task performance, subjective and objective measures of CWL. Additional exploration is also required to ensure improved signal to noise ratio when considering fusion of multimodal data to ensure benefit over single sensors.

Future work will entail quantitative analysis of the data to demonstrate causation and correlation of increased CWL to rationalise and justify components of the multimodal sensor system however, preliminary results so far support the use of a multimodal sensor system to improve reliabilty and overcome individual variability when compared to single sensors used to meaure CWL.

There is scope for further development to utilise machine learning techniques to recognise changes in CWL in an automated fashion and thus enable future integration into a smart operating environment in real-time.

Funding: This project is funded by the EPSRC Transformative Healthcare Technologies grant EP/W004755/1 and Imperial College Healthcare NHS Trust BRC.

REFERENCES

1. Sweller J. Cognitive load during problem solving: Effects on learning. *Cogn Sci*. 1988;12(2):257-285. doi:10.1016/0364-0213(88)90023-7
2. Wanyan X, Zhuang D, Zhang H. Improving pilot mental workload evaluation with combined measures. *Biomed Mater Eng*. 2014;24(6):2283-2290. doi:10.3233/BME-141041
3. Wilson MR, Poolton JM, Malhotra N, Ngo K, Bright E, Masters RSW. Development and validation of a surgical workload measure: The surgery task load index (SURG-TLX). *World J Surg*. 2011;35(9):1961-1969. doi:10.1007/s00268-011-1141-4
4. Zakeri Z, Mansfield N, Sunderland C, Omurtag A. Physiological correlates of cognitive load in laparoscopic surgery. *Sci Rep*. 2020;10(1):1-13. doi:10.1038/s41598-020-69553-3
5. Gao J, Liu S, Feng Q, et al. Subjective and objective quantification of the effect of distraction on physician's workload and performance during simulated laparoscopic surgery. *Med Sci Monit*. 2019;25:3127-3132. doi:10.12659/MSM.914635