MENTAL WORKLOAD MONITORING WITH MULTIMODAL APPROACH DURING LAPAROSCOPIC SURGERY TRAINING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
ANKARA UNIVERSITY

by

Kubranur KARA BASTURK

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN BIOMEDICAL ENGINEERING

ANKARA
2023

All rights reserved

ABSTRACT

Masters Thesis

MENTAL WORKLOAD MONITORING WITH MULTIMODAL APPROACH DURING LAPAROSCOPIC SURGERY TRAINING

Kubranur KARA BASTURK

Ankara University
Graduate School of Natural and Applied Sciences
Department of Biomedical Engineering

Supervisor: Assist. Prof. Hasan Onur KELES

The surgeon in laparoscopic surgery faces several challenges in terms of cognitive load and performance. As the mental workload of surgeons increases, the rates of surgical errors increase. Surgeons and surgical residents can be trained with laparoscopic surgery simulators using standard tasks. In this study, our aim is to investigate mental workload related biomarkers in the brain and human physiology with a multimodal approach. Due to the intricate nature of surgical skills and mental workload, it is essential to adopt multimodal approaches to investigate objective measures.

In this thesis, behavioral and performance data such as NASA-TLX, completion time, and error rate were collected in addition to physiological methods such as ECG and EEG in surgeons and inexperienced subjects during laparoscopic tasks. Since surgical skills are mainly derived from brain activity, peripheral physiological parameters were also evaluated and investigated. The innovative aspect of this thesis is that these multimodal data will be used for assessment and training of surgeons.

August 2023, 52 pages

Key Words: Laparoscopic surgery, Mental workload, EEG, ECG, Multimodal approach

ÖZET

Yüksek Lisans Tezi

LAPAROSKOPİK CERRAHİ EĞİTİMİ SIRASINDA MULTİMODAL YAKLAŞIM İLE ZİHİNSEL İŞ YÜKÜNÜN İZLENMESİ

Kübranur KARA BAŞTÜRK

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Biyomedikal Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Hasan Onur KELEŞ

Laparoskopik cerrahi de hem bilişsel yük hem de performans açısından cerrah için muhtemel birçok zorluk bulunmaktadır. Cerrahların zihinsel iş yükü arttıkça, cerrahi hataların oranları artmaktadır. Laparoskopik cerrahi simülatörleri ve standart görevler ile cerrahların ve cerrahi asistanların eğitimleri yapılabilmektedir. Bu çalışmada, laparoskopi cerrahi simülatör kullanımı sırasında çoklu (multimodal) yaklaşım ile cerrahi uzmanlığın ve zihinsel iş yükünün, beyinde/insan fizyolojisinde olan belirteçlerinin araştırılması hedeflendi. Cerrahi becerinin ve zihinsel iş yükünün nörobilişsel karmaşıklığı göz önüne alındığında, objektif ölçümleri araştırmak için çoklu yaklaşımlar gereklidir.

Bu tez kapsamında, EEG ile EKG ölçümleri eş zamanlı olarak yapılmasının yanında NASA-TLX, bitirme süresi, hata oranı gibi davranışsal veriler deneyimli cerrahlar ve deneyimsiz katılımcılardan toplanmıştır. Cerrahi beceriler ağırlıklı olarak beyin aktivitesinden elde edilmesine rağmen, çalışma kapsamında sinir sistemi aktivitesinin değerlendirilmesine ek olarak periferik fizyolojik parametreler de değerlendirildi. Projenin yenilikçi yönü, bu çoklu verilerin cerrahi performans tahmini ve eğitim rehberliği sağlamaya çalışacak olmasıdır.

Ağustos 2023, 52 sayfa

Anahtar Kelimeler: Laparoskopik cerrahi, Zihinsel iş yükü, EEG, EKG, Multimodal yaklaşım

FOREWORD AND ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Assist. Prof. Dr. Hasan Onur Keles, for his invaluable guidance and unwavering support during my academic journey. His patience, motivation, enthusiasm, and extensive knowledge have been extremely valuable throughout my research and the completion of this thesis.

I also want to express my gratitude to my beloved family and my dear husband, Tarık Baştürk. They have consistently provided both financial and moral support during this challenging thesis journey and throughout my entire educational experience. For all the people who contributed in many other ways I am forever grateful.

May God bless all of you.

Kubranur KARA BASTURK
Ankara, August 2023

TABLE OF CONTENTS

THESIS APPROVAL
ETHIC ..... **i**
ABSTRACT ..... **ii**
ÖZET ..... **iii**
FOREWORD AND ACKNOWLEDGEMENTS ..... **iv**
ABBREVIATIONS ..... **vi**
LIST OF FIGURES ..... **vii**
LIST OF TABLES ..... **viii**
1. INTRODUCTION **1**
2. LITERATURE REVIEW **3**
2.1 Mental Workload **3**
2.2 Mental Workload Measurement Methods **5**
2.2.1 Performance measures **5**
2.2.2 Physiological measures **6**
2.2.3 Subjective measures **9**
2.3 Mental Workload Measurement in Surgery **10**
3. MATERIALS AND METHOD **14**
3.1 Participants **14**
3.2 Experimental Design **14**
3.3 Data Collection **17**
3.4 Data Analysis **18**
3.4.1 Behaviour and performance measures **18**
3.4.2 ECG measures **18**
3.4.3 EEG Measures **19**
3.4.3.1 Preprocessing **19**
3.4.3.2 PLV analysis **19**
3.5 Statistical Analysis **20**
4. RESULTS **21**
5. DISCUSSION **37**
5.1 Limitation **42**
6. CONCLUSION **43**
REFERENCES **44**
CURRICULUM VITAE **52**

ABBREVIATIONS

US	United States
EEG	Electroencephalography
ECG	Electrocardiogram
EOG	Electrooculogram
EMG	Electromyogram
GSR/EDR	Electrodermal response measurement
SDNN	The measure of variability in the time intervals between successive NN intervals
RR	Inter-beat Intervals
Pnn50	Percentage of successive RR intervals differing by more than 50 ms
HRV	Heart Rate Variable
HR	Heart Rate
RMSSD	Root-mean-square of successive differences of NN intervals
SDSD	Standard Deviation of Successive Differences
fMRI	Functional magnetic resonance imaging
fNIRS	Functional Near-Infrared Spectroscopy
PLV	Phase-Locking Value
ECN	Executive Control Network
SN	Salience Network
DMN	Default Mode Network
PFC	Prefrontal Cortex

LIST OF FIGURES

Figure 3.1 Peg Transfer Task 16
Figure 3.2 String Pass Task 17
Figure 3.3 PLV analysis 19
Figure 4.1 Completion time of participants 21
Figure 4.2 Reaction time of participants 21
Figure 4.3 Total NASA-TLX score of participants 22
Figure 4.4 Sub NASA-TLX score for Task 1 23
Figure 4.5 Sub NASA-TLX score for Task 2 3
Figure 4.6 RMSSD value of participants 24
Figure 4.7 Pnn50 value of participants 24
Figure 4.8 SDSD value of participants 25
Figure 4.9 Reaction Time v RMSSD for surgeons during Task 1 25
Figure 4.10 Completion Time v Total NASA-TLX for surgeons during Task 1 26
Figure 4.11 Physical Demand v Completion Time for surgeons during Task 1 26
Figure 4.12 Frustration v Completion Time for surgeons during Task 1 27
Figure 4.13 Effort v Completion Time for surgeons during Task 1 27
Figure 4.14 Completion Time v Total NASA-TLX for surgeons during Task 2 28
Figure 4.15 Reaction Time v Lap. Experience for surgeons during Task 2 28
Figure 4.16 Completion Time v Frustration for surgeons during Task 2 29
Figure 4.17 Reaction Time v Frustration for surgeons during Task 2 29
Figure 4.18 Completion Time v RMSSD for the inexperienced during Task 1 30
Figure 4.19 Pnn50 v Performance for the inexperienced during Task 1 30
Figure 4.20 Performance v Completion Time for the inexperienced during Task 1 31
Figure 4.21 Frustration v Completion Time for the inexperienced during Task 1 31
Figure 4.22 Effort v Completion Time for the inexperienced during Task 1 31
Figure 4.23 React. Time v Total NASA-TLX for the inexperienced during Task 2 32
Figure 4.24 RMSSD v Performance for the inexperienced during Task 2 33
Figure 4.25 Pnn50 v Performance for the inexperienced during Task 2 33
Figure 4.26 Completion Time v Effort for the inexperienced during Task 2 33
Figure 4.27 The PLV connectivity network among EEG channel pairs of inexperienced and surgeons during Task 1 35
Figure 4.28 The PLV connectivity network among EEG channel pairs of inexperienced and surgeons during Task 2 35

LIST OF TABLES

Table 3.1 Demographics of participants 14
Table 3.2 Experimental design..... 14
Table 4.1 Significant PLV-Pairs in Connectivity in Task 1..... 36
Table 4.2 Significant PLV-Pairs in Connectivity in Task 2..... 36

1. INTRODUCTION

In the United States, medical errors are a major cause of death, ranking third, and resulting in a yearly cost of $20 billion. Among these errors, surgical errors make up the largest portion. The World Health Organization (WHO) states that one out of every 10 patients experience medical errors (World Health Organization, 2004). Surgical errors are the most frequently reported type, contributing to about half of all avoidable mistakes in healthcare (Mülayimoğlu & Ayaz, 2012). Surgical skill isn't just a result of natural talent; it can be cultivated through proper training and technique (Rosser et al., 1997). Research shows that around 50% of surgical complications can be prevented, especially when handled by skilled and proficient teams (Mazzocco et al., 2009). A lack of adequate training contributes to increased error rates. Surgical residency programs aim to educate professionals with the necessary technical, personal, and cognitive skills demanded by society (Aggarwal et al., 2004). However, current surgical training follows an apprenticeship model, which might fall short in producing competent surgeons. Few opportunities exist for medical students to practice surgical skills in a controlled environment separate from actual patients (Porte et al., 2007). Given the large number of students and limited training resources, new training methods need to be developed for this group of trainees (Porte et al., 2007). This study focuses on laparoscopic surgery training for several reasons: (1) Laparoscopic surgery involves tasks that are well-defined and standardized, suitable for individuals at different experience levels. (2) It allows us to assess mental workload, which can be measured using techniques like neuroimaging and heart rate variability. (3) The mental demands on surgeons are substantial during laparoscopic procedures. (4) Laparoscopy is becoming more common in surgery due to technological advancements. (5) It offers an excellent opportunity for collaboration between engineers and clinicians. Many studies demonstrate that neuroimaging can effectively measure various aspects of mental state, such as cognitive workload, attention, stress, and memory. These factors are closely related to how well human operators perform (Van Merriënboer & Sweller, 2010). In surgical training, there's a need for objective methods to assess mental workload.

This thesis aims to explore mental workload patterns during laparoscopic training for both

experienced surgeons and inexperienced participants. We used multimodal noninvasive mobile monitoring to analyze brain activity and physiological changes. Our approach involves combining data from Electroencephalography (EEG) and Electrocardiography (ECG), along with evaluating behavior and performance. We looked into differences in brain activity, physiology, and behavior based on task difficulty (easy/difficult), comparing experienced surgeons with inexperienced participants. Due to the extended and uninterrupted work hours of surgeons, particularly surgical assistants, mental workload levels are significantly elevated within the operating room. Objective techniques for assessing mental workload are still under investigation in the research communities. Traditionally, mental workload has been assessed through methods like performance evaluations. It can be hypothesized that performing laparoscopic surgery places a greater mental workload on surgeons compared to other types of surgeries. Being able to recognize subtle levels of expertise will boost the confidence in identifying skilled artists and leaders, particularly in times of medical emergencies. Traditionally, mental workload has been evaluated using limited methods like performance and subjective assessments, which can sometimes yield unsatisfactory outcomes. However, technological advancements have led to the creation of user-friendly sensors, and their growing acceptance within the global scientific community has enhanced the utilization of physiological data for skill assessment. Recent advancements in sensor technology bring significant opportunities by allowing real-time processing of neurophysiological data in various fields to create innovative applications that enhance human-machine interaction. Through modern wearable, wireless, lightweight, and non-invasive interfaces, brain functions can now be monitored outside of labs, using methods like optical brain imaging, electroencephalography, and physiological markers such as eye movements, heart rate, breathing rate, and skin conductivity. These technologies enable us to understand the cognitive states of professionals like surgeons during training and actual tasks, as seen in laparoscopic surgery. By tracking brain activity and analyzing other neurophysiological indicators, we can gain insights into surgeons' mental states. This knowledge helps us detect and prevent risky conditions like unconsciousness, distraction, and fatigue at an early stage. Furthermore, such technology can lead to more personalized training for surgeons, pinpoint areas for improvement, and offer tailored feedback to enhance surgical skills. As surgical skills primarily stem from brain activity, this study also assessed peripheral physiological factors, with a specific emphasis on evaluating the activity of the nervous system.

2. LITERATURE REVIEW

Most surgeons agree that laparoscopic surgery, which is often used for minimally invasive procedures, requires more focus than open surgery. The constraints of instruments and visual performance make laparoscopic surgery more demanding and stressful (Berguer et al., 2001). Stress levels are higher, and technical performance is reduced in laparoscopic procedures versus open and robotic surgery (Sevdalis, Arora, Nestel, et al., 2010). Unlike open surgery, mental workload gradually increases in laparoscopic tasks due to heightened focus and stress (Berguer et al., 2001). Experienced users reporting lower concentration for open surgery tasks reflects their surgical expertise and confidence (Berguer et al., 2001). Laparoscopic surgery, with its limited field of view and need for assistance, presents technical challenges that necessitate specific training. Training programs should differ for experts and assistants, as instructional approaches vary based on learner experience (Van Merriënboer & Sweller, 2010). Many studies focus on cognitive processes impacting surgical performance, with mental workload being a fundamental one. Other cognitive aspects like situational awareness, divided attention, mental exhaustion, alertness, and fatigue are also part of this field (Schneider & Shiffrin, 1977; Wickens, 2014).

2.1 Mental Workload

Cognitive workload is influenced by a cognitive system that involves a working memory with limits on capacity and duration (G. A. Miller, 1956; Peterson & Peterson, 1959), along with a large long-term memory that can store and process new information with almost unlimited capacity (Antonenko, Paas, Grabner, & Van Gog, 2010). Mental workload refers to the amount of brain activity occurring within a given timeframe.

The concept of mental workload is based on the idea that individuals have a relatively consistent cognitive capacity (Carswell et al., 2005). Laparoscopic surgery presents challenges not only physically but also mentally. The cognitive strain on surgeons can greatly impact patient safety and healthcare expenses (Zakeri et al., 2020). Real-time automatic tracking of a surgeon's workload during an operation can provide valuable data

for various systems (Zander et al., 2017). This data can enable the system to offer automatic assistance, allowing communication with the surgeon during critical or stressful moments. It also facilitates communication with others while the surgeon is focused on complex maneuvers (Zander et al., 2017). Novice surgeons are particularly sensitive to stress (Anton et al., 2016). As expertise grows, cognitive load lessens. Using physiological data for robust assessment could lead to more effective training and certification methods in this field (Zakeri et al., 2020). Leveraging already learned information stored in long-term memory as cognitive schemas reduces the load on working memory. A schema acts as a single information unit in working memory (Antonenko, Paas, Grabner, & Van Gog, 2010). Various training simulation modules can assist in this process. Standardized training programs, including computer-based virtual reality simulations or training boxes, have gained traction for minimally invasive surgery. However, these tools haven't gained widespread acceptance due to their limited realism and lack of suitable performance evaluation methods (Stylopoulos et al., 2004). This highlights the significant gap between the demand for better educational methods and the current education systems in place (Stylopoulos et al., 2004). Surgical training can incorporate physical laparoscopy training simulators. A comprehensive mental skills curriculum, tailored to surgery and built upon simulators, has been developed and proven effective in improving both mental skills and surgical performance (Stefanidis et al., 2017). Using these insights, a simulator-focused curriculum could help surgeons manage stress, enhance their performance, and ultimately positively impact patient outcomes (Stefanidis et al., 2017). Higher workload is linked to reduced task performance and an increased likelihood of errors (Yurko et al., 2010). Employing a physical training simulator for multimodal mental workload monitoring can play a significant role in tracking the cognitive state of surgeons and residents throughout their training.

Laparoscopic surgery is an area where the cognitive state and training of surgeons and residents hold significant importance. During complex surgical procedures, laparoscopic surgeons must maintain high concentration and attention. The cognitive state and performance of surgeons play a direct role in patient safety and surgical results. Neurophysiological data could provide valuable chances in laparoscopic surgery. Keeping track of and studying the cognitive states of surgeons could enhance the safety

of surgeries, their performance, and the final results.

2.2 Mental Workload Measurement Methods

Methods to evaluate mental workload can be divided into three main groups: Performance measures, physiological measures, and subjective measures. These measurement methods will be discussed further in the following section. Subjective measures involve questionnaires that let individuals assess their own mental workload by rating factors like mental effort, fatigue, and frustration. Various physiological measurements are linked to cognitive performance and can therefore be used to measure mental workload. Performance criteria, such as how long it takes to complete tasks, reaction times, and error rates during experiments, can also indicate mental workload.

2.2.1 Performance measures

Mental workload performance metrics are utilized to assess performance levels associated with workload. These metrics can involve factors like task accuracy, time taken to complete tasks, response time, and error rate. For instance, the time it takes to finish a task or how accurately it's done might vary as mental workload increases. These metrics help gauge workload levels, enhance task efficiency, and boost overall performance. These measurements are taken during various cognitive tasks, and the results are analyzed to understand the connection between workload and performance. These metrics also reveal the difficulty of tasks and how their difficulty changes over time. In this approach, performance measures involve recording the time it takes an individual to complete a task, how accurately they do it, and other performance factors. Then, the data is studied to quantify the mental workload associated with a specific task. These metrics are divided into two categories: primary task and secondary task. Primary task metrics gauge participants' task completion time as a performance indicator. For the secondary task, participants' reaction time is assessed as a performance measure, where they manually respond to an auditory cue. The performance measurements in this study cover participants' time to complete the experiment and their reaction times.

2.2.2 Physiological measures

Physiological measures are based on the concept that increased mental demands can trigger stronger physical reactions in the body (S. Miller, 2001). Miller (S. Miller, 2001) suggested 5 ways to measure physiological responses to understand mental workload: brain activity, heart rate, eye movement, breathing, and speech. Mental workload can be assessed using various physiological methods, although most studies have traditionally focused on a single method. With advancements in technology and hardware, mobile and portable wireless systems are now being used to combine multiple methods outside of the lab. Neurophysiological parameters like heart rate, blink frequency, pupil dilation, eye focus, muscle movements, and skin conductivity, measured through techniques such as electrocardiogram (ECG), electrooculogram (EOG), electromyogram (EMG), and electrodermal response (GSR/EDR), can help track the impact of the nervous system on the body. Ongoing research aims to understand how these parameters relate to the cognitive state of surgeons.

ECG signals are a sensitive source for measuring mental workload. Numerous studies have established a link between ECG (Electrocardiographic) measurements and mental workload (Heine et al., 2017; Loeches De La Fuente et al., 2019; Marquart et al., 2015; Shao et al., 2020). Many studies have found a strong connection between heart rate variability (HRV) and mental workload. HRV, a widely-used feature derived from ECG signals, indicates fluctuations in heartbeat and serves as a physiological marker. Common parameters used to evaluate HR and mental workload include SDNN, pNN50 (the percentage of consecutive RR intervals differing by more than 50 ms), HR, and RMSSD (Root Mean Square of Successive RR Interval Differences). Metrics like NN mean, PNN50, total spectrum power, and low frequency are also crucial in identifying mental fatigue, often associated with mental workload. Among these, RMSSD shows a positive link with mental fatigue, while PNN50 and NN average show a negative connection (Huang et al., 2018). A study (Fan et al., 2020) indicated that Mean RR and RMSSD values from ECG parameters decreased significantly with increasing task difficulty. Research demonstrates that as mental workload rises, heart rate (HR) tends to increase (Lenneman & Backs, 2007). Conversely, heart rate variability (HRV) responds more

rapidly to changes in instant workload than heart rate (HR). Particularly, HRV values tend to decrease as mental workload rises (Chen et al., 2015). In another study, it was revealed that as task complexity increases, both mental workload and HRV exhibit noticeable variations while performing tasks of different difficulties (ÖZSEVER & TAVACIOĞLU, 2021). A study (Mukherje et al., 2011) highlighted the precision and reliability of frequency domain metrics, HRV time domain, and nonlinear indices when evaluating mental effort tasks in healthy older individuals.

In neuroimaging, limited research has explored the link between performance and the experience levels of surgical assistants and expert surgeons. Methods like fNIRS and EEG are commonly used in literature to comprehend cognitive processes crucial in surgical training and procedures. Imaging techniques such as fNIRS, EEG, and fMRI allow us to observe the functional activity of the human brain (Bunge & Kahn, 2009; Modi et al., 2017)).

EEG is the longest-standing imaging method that non-invasively records the electrical activity of neural processes using electrodes on the scalp. It excels in temporal resolution but has limited spatial accuracy. The rise of portable wireless EEG devices has opened up applications in areas like neuroergonomics and surgery (Keleş, 2022). On the other hand, fNIRS is a functional imaging technique that measures changes in blood flow within the cerebral cortex to indirectly assess brain activity (Keleş, 2022). It employs silicon photodetectors with LEDs or lasers to emit infrared light on the scalp. By using two light sources of distinct wavelengths, the brain tissue is illuminated, and the reflected photons are detected along a specific path. Differences in the sent and received light help gauge activation levels in the examined brain region.

fMRI, on the other hand, is the most widely used functional imaging method that depicts tissue activity through vascular response, measuring changes in the ratio of oxyhemoglobin to deoxyhemoglobin in the brain. However, due to its weight and complexity, fMRI is limited in its application in neuroergonomics, surgical training, and mental workload studies (Keleş, 2022). fMRI has the highest spatial resolution among functional neuroimaging methods and stands apart from fNIRS and EEG in its ability to

gather data from different brain layers. Lately, combined approaches like Multiple Functional Imaging (fNIRS+fMRI, EEG+fNIRS, etc.) have been used for a more comprehensive understanding of the neural basis of behavior.

Communication among neurons occurs through phase synchronization between pre-synaptic and post-synaptic neurons, both within a specific brain region and across distant regions. The phase lock value (PLV) is a metric used to measure the level of neural phase synchronization and EEG coupling between two groups of neurons in a given frequency range (Jian et al., 2017). This method assumes that the instantaneous phase difference of EEG signals between two connected brain regions remains relatively constant. The phase lock value (PLV) ranges from 0 to 1, where 0 signifies no synchronization between the signals and 1 represents consistent phase alignment between them. Some studies utilize PLV to differentiate synchronization levels between brain regions (Bertrand et al., 2016).

In a recent study (di Biase et al., 2023), PLV analysis was employed to investigate EEG brain connectivity in Parkinson's disease. Among healthy participants, greater PLV connectivity in the delta frequency range was evident during the motor task in comparison to individuals with Parkinson's disease. However, no distinction was noted in the resting state. Another research explored a surface tactile device capable of providing haptic feedback during an interactive touch task. The study used PLV to assess functional connectivity. The main goal was to collect PLV data from the alpha, beta, and gamma frequency ranges. The study aimed to compare outcomes under two conditions: one without haptic feedback and the other with haptic feedback through a touchscreen device. The research demonstrated that interhemispheric regions exhibited distinct functional beta connectivity (Park et al., 2019). Research also suggests that heightened theta band activity and reduced alpha band activity in EEGs indicate increased mental workload (Brookings et al., 1996; Sterman & Mann, 1995). Changes in workload are also reflected in the magnitude of evoked potential components linked to stimuli, regardless of the specific tasks being performed (Kramer et al., 1995). For example, studies conducted in simulator environments have shown that when distraction affects performance, power distribution in the theta band decreases, and the distribution in the alpha band changes. Conversely, tasks demanding focused attention exhibit increased power in the theta band,

notably from electrodes in anterior-middle and peak-middle regions (Borghini et al., 2011, 2014; Dussault et al., 2005).

In a study by Zakeri et al. (2020), fNIRS and EEG data were collected simultaneously from 38 healthy adults. Using this data, heart rate (HR), heart rate variability (HRV), and blink rate (BR) were analyzed. The aim was to explore how physiological variables can indicate the effort level associated with training tasks. The experiment revealed connections between physiological data and expended effort. Another study (Keles et al., 2016) investigated correlations among multimodal signals, neural activity across various brain frequencies, and localized changes in hemoglobin levels in 18 healthy adults. The study confirmed delayed hemodynamic modulation in posterior areas, particularly in the alpha frequency range (8-16Hz). Robust beta (16-32Hz) modulation of hemodynamics was also observed, with beta modulation likely resulting from alpha-beta coupling in EEG.

During an fMRI investigation (Causse et al., 2022), 20 participants performed a complex cognitive task under safe and challenging conditions. The study proposed that assessing activation balance in fMRI signals might be a method for objectively gauging mental status. It was suggested that task-related regions could exhibit varying ratios of prefrontal activities, a key indicator of increased mental effort.

2.2.3 Subjective measures

The most common approach for assessing mental workload is the questionnaire method. This involves using questions or surveys to gather individuals' feedback on their experiences. This method can include inquiries about the time taken to complete a specific task, the perceived difficulty level of the task, its repetitiveness, and other task-related aspects. The endeavor to comprehend mental state descriptions and their impact on individual and team performance encounters a significant challenge due to the fact that these descriptions can't be directly observed. To gauge mental workload during laparoscopic surgery, it becomes crucial to gather both behavioral and physiological data from surgeons. Traditionally, behavioral cues such as secondary task performance or

survey outcomes like the NASA-TLX survey are referred to for measuring these aspects. For example, a study showcased that mental workload evaluated using a standard self-report tool correlates with the error rate and quality of suturing in laparoscopic surgery training. Essentially, increased workload is linked to reduced task performance and a higher probability of errors (Yurko et al., 2010). The NASA-TLX covers mental demand, temporal demand, physical demand, frustration, performance, and effort level. The NASA-TLX is a comprehensive and multi-dimensional scale utilized to assess workload estimations of one or more operators during or after mission completion (Hart, 2006).

2.3 Mental Workload Measurement in Surgery

Assessing mental workload in surgery involves evaluating the cognitive demands that emerge during surgical procedures. The objective of this measurement is to understand the cognitive load, attention requirements, and mental resource utilization that surgeons encounter while performing operations. This approach is utilized to gauge surgeons' workload, improve their performance, and have a positive impact on surgical results. Numerous studies have been conducted to evaluate mental workload during lengthy and intricate laparoscopic surgical procedures. HRV and HR serve as indicators of autonomic responses, allowing for the evaluation of cognitive load in the surgical setting (Rieger et al., 2014). Changes in heart rate during the intraoperative period tend to be more prominent than those attributed to physical and musculoskeletal demands (Arora, Sevdalis, Aggarwal, et al., 2010). Hence, heart rate has been utilized as a marker to gauge acute stress levels experienced by surgeons (Alobid et al., 2011; Payne & Rick, 1986).

Heart rate variability (HRV) can detect autonomic changes even when heart rate measurements remain steady. This implies that even with a constant heart rate, HRV measurements can unveil effects on the autonomic nervous system. A decrease in HRV has been observed during surgical decision-making processes (Czyewska et al., 1983). This could signify the cognitive load surgeons undergo while making decisions and performing procedures. It has also been discovered that HRV can be linked to cognitive load (Song et al., 2009; Böhm et al., 2001). HRV can function as a measurement tool in assessing the mental demands placed on surgeons. Moreover, HRV has been found to

assist in identifying specific surgical segments that evoke varying stress levels among surgeons specializing in coronary bypass surgery and resident surgeons (Song et al., 2009; Böhm et al., 2001). This implies that surgeons can use HRV measurements to recognize stressful situations and tailor intervention strategies accordingly.

A study demonstrated that working in a hospital environment induces more stress compared to other daily activities, particularly for those with limited surgical experience, and residents experience more stress than experienced surgeons (Weenk et al., 2018). Another study (Stefanidis et al., 2017) has shown the effectiveness of improving mental skills and surgical performance through simulator-centered training. The results indicate that training with simulators can help reduce the mental workload of surgeons in surgical settings, leading to enhanced performance and positive impacts on patient outcomes. In a study involving surgeons with varying levels of laparoscopic experience (Moorthy et al., 2003), three different stressors were introduced, resulting in compromised dexterity and increased error rates across all stress conditions. A study by Rosser et al. (1997) investigated 150 trainees and assessed three standard laparoscopic drills. The findings revealed that the cognitive acquisition of skills was linked to improved performance in complex laparoscopic tasks, particularly in intracorporeal suturing, which is known to be challenging and intricate. Learning studies have highlighted the intricate relationship between cognitive load and skill mastery (Byrne et al., 2010). While direct observation and assessment by instructors may suggest satisfactory trainee performance, they may not accurately predict skill retention or real-world performance outcomes. Proficiency in complex tasks tends to coincide with a gradual decrease in cognitive load (Schneider & Shiffrin, 1977), and minimally invasive surgeries generally involve higher cognitive workload compared to open procedures (Byrne et al., 2010). Interestingly, a trainee who achieves an expert level in a surgical simulator might not perform as well in an actual operating room setting, and the extent of this drop in performance correlates with the mental workload experienced during training (Byrne et al., 2010). Trainees with similar initial performance levels can show varied subsequent performance outcomes, and these differences may be attributed to variations in their cognitive load.

In a study (Alan Gevins et .al., 1998), working memory load was investigated while

participants used computers, utilizing neural network pattern recognition on EEG spectral features. As the load increased, there was a noticeable rise in frontal theta EEG activity, coupled with a decrease in alpha activity. The researchers hypothesized that these changes likely reflect the heightened mental effort required by the task's difficulty and the allocation of more cortical resources to task performance. In other studies (Antonenko, Paas, Grabner, & Van Gog, 2010; Maimon et al., 2022), standard and robotic surgical methods were compared using physiological measurements (Hubert et al., 2013). In studies that monitored surgeons using EEG, a common finding was an increase in power in the theta (4-7 Hz) band of the frontal electrodes, indicative of high mental workload and challenging task conditions. The study involved 13 surgeons with varying levels of laparoscopic experience and 27 inexperienced participants who performed standard laparoscopic training tasks. Researchers recorded and analyzed their performance, subjective cognitive load, and prefrontal cortical activities. Experienced surgeons exhibited significantly faster task performance, experienced lower cognitive load compared to inexperienced individuals, and displayed less impact on left prefrontal activation (H. O. Keles & Omurtag, 2021). In a study by Crewther et al. (2016) using fNIRS, it was reported that after 3 weeks of laparoscopic training for medical students, a decrease in mental workload and symmetrical hemodynamic activation in the bilateral prefrontal region were observed. Another study by Modi et al. (2018) involving groups of residents with varying levels of experience found a decrease in prefrontal activation and performance during time-pressure tasks. The reduction in prefrontal activation was less pronounced in senior residents, suggesting a slower increase in mental workload. In a specific study (H. Keles et al., 2021), functional neuroimaging data were collected from sixteen participants, including eleven surgeons, five assistant surgeons, and seventeen students. Two motor tasks were completed, and participants experiencing higher mental workload showed elevated activations in the prefrontal cortex. Another study (H. O. Keles et al., 2021), involved 16 surgeons and 17 students, utilizing an advanced high-density NIRS (near-infrared spectroscopy) device. Surgeons exhibited higher activations compared to subjects with lower and higher task loads. Students displayed a more prominent response in the left prefrontal cortex, particularly in the dorso- and ventrolateral areas.

In a study comparing stress levels between daily situations and surgical procedures, it was found that stress during surgery was higher (Weenk et al., 2018). This elevated surgical stress poses a potential risk to patient safety, and efforts should be directed towards reducing this stress. It has been demonstrated that experienced surgeons tend to handle the procedure with more ease compared to less experienced ones. Notably, the experienced surgeons displayed significantly lower HRV levels (Böhm et al., 2001). Research findings have consistently indicated that experienced surgeons exhibit lower mental workload and superior technical performance compared to resident surgeons (Arora, Sevdalis, Nestel, et al., 2010). Consequently, instructional approaches need to be carefully tailored for novice learners as opposed to more experienced individuals, taking into account their varying skill levels and stress responses (Van Merriënboer & Sweller, 2010). Research has revealed that elevated stress levels can have adverse impacts on surgeons, potentially compromising their fine motor coordination and decision-making skills (Antonenko, Paas, Grabner, & Van Gog, 2010). In a study comparing robotic techniques with conventional laparoscopic methods, it was observed that complex task performance remained similar, but thumb muscle strength decreased and mental stress potentially reduced with the use of robotic techniques (Berguer & Smith, 2006).

3. MATERIALS AND METHOD

3.1 Participants

The study included two participant groups: 22 healthy adult volunteers without prior laparoscopic surgery experience, and 14 specialized surgeons experienced in laparoscopic procedures. Additional participant details are presented in Table 1. Before commencing the study, participants provided their informed consent. The research protocol was approved by the Ankara University Ethics Committee (AUTFKAEK No: 2021/435), and all procedures were conducted in compliance with applicable guidelines and ethical standards.

Table 3.1 Demographics of participants

	Female	**Male**	**Age**	**Lap. Experience (Year)**
Surgeons	7	7	41±13	18±12
Inexperienced	15	7	23±4	0

3.2 Experimental Design

Table 3.2 Experimental design

EEG+ECG record				**EEG+ECG record**	
Restingrecording	**Peg transfer (Task 1) Easy**	**NasaTLX (Task 1)**	**Restingrecording**	**String pass (Task 2) Difficult**	**Nasa- TLX (Task 2)**
2 min.	Max. 7 min		2 min.	Max. 7 min.	

This experiment was conducted at the Ankara University Biomedical Engineering laboratory, utilizing a custom-designed laparoscopy simulation box and a single monitor. Each session lasted approximately one hour, covering both task performance and setup

time for the system and devices. At the start of each session, participants were given a 10-minute trial to acquaint themselves with the surgical equipment and an overview of the experiments. A camera was installed inside the laparoscopy training simulator, with a training table positioned below the box. The training table was centered within the camera's field of view. The video feed from the laparoscopic training simulator was directly displayed on a monitor placed within the participant's line of sight.

Participants in the study performed two essential surgical tasks commonly used in laparoscopic surgery: The Peg Transfer (referred to as Task 1) and the String Pass (referred to as Task 2). The Peg Transfer task holds great significance in laparoscopic surgery as it hones the technique and skill necessary for the procedure and provides valuable self-assessment and improvement opportunities. For the Peg Transfer Task, participants engaged with five rings, using graspers to lift each ring, move it between rods, and position it in a designated area. The rods were labeled A, B, C, and D at the lower left, upper left, upper right, and lower right corners, respectively. Participants began with five rings on rod A, transferring them to rod B using their left hand. Then, they picked up each ring with their left hand, lifted it, and placed it on rod C using their right hand. Rings on rod C were moved individually with the left hand and placed on rod D. If any rings were dropped, participants could place them on the tray and then back on the rod, continuing the experiment. This task required precise coordination of the surgical trainee's hands, sharp depth perception, and strong visuospatial awareness within the simulated environment. The aim was to enhance spatial and depth perception while refining dexterity using a simulated instrument (holder). The mission equipment included a simulator laparoscopy box that displayed images on the screen, along with two attached holders. The peg board and pegs were located within the laparoscopy box. The mission concluded upon completion of the peg transfers. Dropped pegs were considered errors. Performance criteria included recording the completion time and error rate. The string pass task involves passing a string through loops positioned at various angles and heights. This task focuses on bimanual coordination and skill development, falling within the upper-intermediate difficulty range. Its purpose is to enhance the technique and skill necessary for laparoscopic surgery. Participants have the freedom to use both surgical instruments without a specific hand restriction. The task starts with threading the string

through the hole marked as "1". Using laparoscopic holders, participants transfer the string between their hands in a clockwise direction, passing it through 6 rings arranged in the shape of the letter "S". The numbers on the board indicate the sequence for looping the string. Maneuvering the laparoscopic instruments at different angles is essential to navigate the rings. The task is completed once the string has passed through all 6 rings.

To mimic real-world distractions or interruptions, an auditory task was introduced during the experiment. This secondary task served as a stimulus, providing an additional metric to assess participants' cognitive load by analyzing their reaction time and responsiveness. Participants were tasked with pressing a footswitch in response to randomly generated beeps, aiming for a swift response. Before beginning the experiment, participants were instructed to keep their heads still. The time taken by participants to finish the laparoscopic tasks was recorded, with a maximum time limit of 7 minutes from the moment they received instructions to start the experiment.

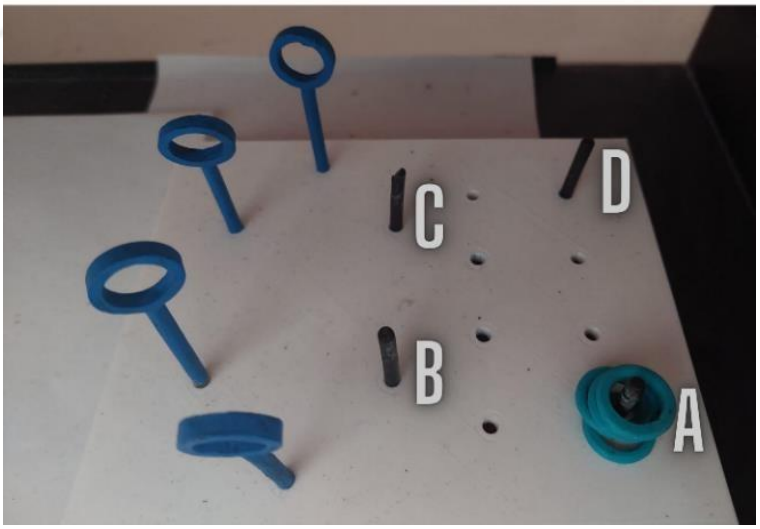


Figure 3.1 Peg Transfer Task

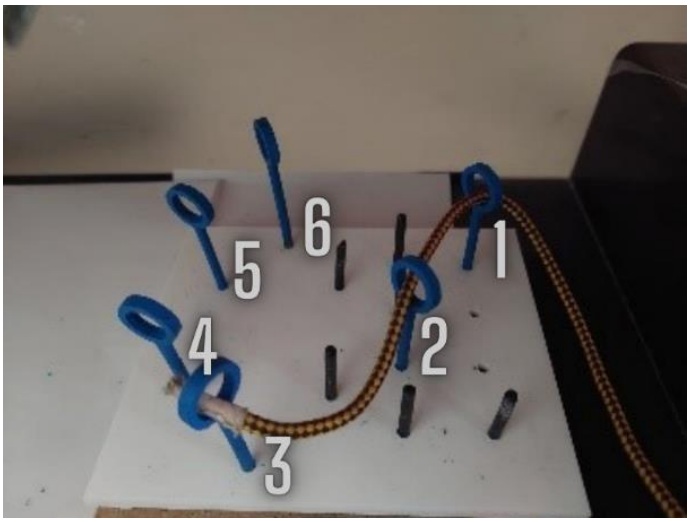


Figure 3.2 String Pass Task

3.3 Data Collection

Data were collected from surgeons and inexperienced groups. To collect data from the Surgeons, they went to Istanbul with the laparoscopy training simulator box and other equipment. Data were collected from surgeons working in the private and government hospitals. For the inexperienced group, data were collected from inexperienced participants at Ankara University in the Biomedical Engineering Neuroimaging laboratory. During the study, EEG caps were attached to the participants and Mentalab Explore device was integrated. We used the Mentalab Explore (EEG+ECG) device, a wireless data collection tool, to simultaneously record participants' brain electrical activities and heart rate variability. This device features 8 channels and a sampling rate of 250 Hz. The electrodes were positioned following the 10-20 electrode system. Specifically, EEG data were collected from 7 channels (FZ, F4, C3, C4, P4, F3, P3), and ECG data were collected from 1 channel. As behavioral data, we recorded task completion times and reaction times of the participants. After each experiment, participants were asked to complete the Nasa-TLX questionnaire. This questionnaire assesses mental workload using a multidimensional rating approach with six subscales, providing insight into participants' subjective experience and mental effort levels.

3.4 Data Analysis

3.4.1 Behaviour and performance measures

During the study, researchers recorded the time it took for participants to complete the three laparoscopic tasks and their reaction times. To assess mental workload, participants filled out a NASA-TLX questionnaire where each criterion was rated on a scale from 1 to 20. The overall mental workload score was calculated by adding up the scores for each criterion. The data were analyzed using linear regression to explore relationships between participants' task completion times, reaction times, NASA-TLX scores, HRV analysis, and other relevant parameters.

3.4.2 ECG measures

To reduce high-frequency noise, a low-pass filter with a cut-off frequency of 30 Hz was applied to each signal. To address the low-frequency wandering artifact in participants' ECG signals, a high-pass filter with a cut-off frequency of 0.5 Hz was used. Removing this wandering artifact is crucial for clear ECG interpretation. In the preprocessing stage, a band-pass filter ranging from 0.5 Hz to 30 Hz was applied in the Python interface to eliminate noise and correct the navigation artifact. For ECG analysis, participants' RR interval series were used. This series formed the basis for HRV analysis. To ensure accuracy, incorrectly detected artifacts that could cause inaccuracies in RR intervals have been removed. HRV analysis was performed using the HRV-Analysis package, a Python tool that supports various time, frequency, and nonlinear HRV analysis methods within the Python statistical computing environment. In the analysis of Heart Rate Variability (HRV) in the time domain, three categories of parameters were considered: statistical, geometric, and nonlinear. For this study, Python HRV analysis calculated key statistical parameters including RMSSD, SDSD, and PNN50. SDSD measures the average RR interval and the standard deviation of normal RR intervals. RMSSD represents the root mean square difference between consecutive RR intervals. PNN50 indicates the percentage of RR intervals where the difference between successive NN intervals exceeds 50 ms. The study aimed to explore the relationships between these ECG parameters and individuals' stress levels.

3.4.3 EEG Measures

3.4.3.1 Preprocessing

The raw and processed signals were visually examined to check for muscle and motion artifacts, eye movements, and other disturbances. The EEG signal was filtered to include frequencies between 0.5 Hz and 30 Hz. Each epoch had a window size of 5 seconds. High-amplitude signals were eliminated using threshold methods.

3.4.3.2 PLV analysis

EEG-based features were derived from the phase locking value (PLV), which is a measure of synchronization in electrical brain activities and biosignals. It quantifies synchronization between frequencies without a specific direction, indicating how different brain areas work together. PLV also shows how the phase difference between two signals changes over time (Aydore et al., 2013; Lachaux et al., 1999). PLV was calculated for four frequency bands, each covering a range of 1 Hz. These bands are named delta (4–5 Hz), theta (7–8 Hz), alpha (10–11 Hz), and beta (14-15 Hz). We also used labels (FZ, F4, C3, C4, P4, F3, P3) to refer to specific brain areas within these frequency bands. PLV was calculated for all possible combinations of brain areas, and separate analyses were conducted for EEG recordings at rest and EEG recordings during motor tasks.

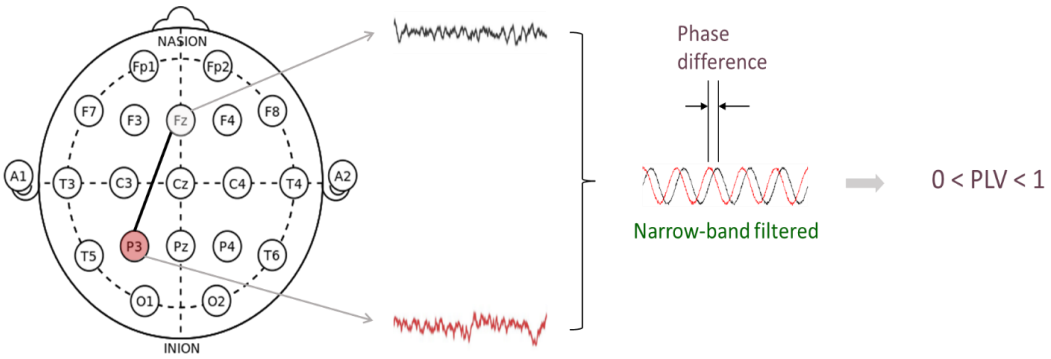


Figure 3.3 PLV analysis

3.5 Statistical Analysis

To perform the calculations mentioned above, we utilized Matlab functions including regress, ftlm, cascadeforwardnet, autofndpeaks, cvpartition, corr, kstest2, and signrank (Matlab R2017b, MathWorks, Inc., Natick, Massachusetts, United States). Additionally, we utilized GraphPad 10 for statistical analysis. The statistical methods used in this thesis included the Kolmogorov-Smirnov test, Wilcoxon signed-rank test and linear regression analysis. The Kolmogorov-Smirnov test is a method that assesses how closely a given distribution matches a specified distribution, like the normal distribution. It involves plotting the cumulative distribution function of observed data against the theoretical distribution function for comparison. A test statistic is then calculated to measure the difference between the two distributions. This test statistic is compared to a critical value. If the test statistic is lower than the critical value, it suggests that the parameters follow a normal distribution. Conversely, if the test statistic exceeds the critical value, it indicates that the parameters have a distribution different from a normal distribution. In this study, the critical value used was 0.05. The descriptive results comparing two groups(Surgeons vs Inexperienced), based on completion time; reaction time, NASA TLX, ECG data contained non-paired data. In this thesis, linear regression analysis was used to explore the relationship between various factors, including performance, subjective perceptions, and physiological parameters. This type of analysis helps us understand how changes in a dependent variable are associated with changes in one or more independent variables. In our analysis of group differences of EEG analysis, we employed the Wilcoxon signed-rank test when dealing with paired subjects within groups (such as participants who completed both the Peg Transfer and String Pass tasks). These tests were chosen for their conservative nature and lack of assumptions about normal distribution. We avoided using a null hypothesis that would have necessitated multiple comparisons.

4. RESULTS

We collected subjective and physiological data from participants who completed tasks on the laparoscopic surgery training simulator. The study involved 36 participants, including 14 surgeons and 22 individuals with inexperienced. Demographic and experience information of participants are shown in Table 3.1.

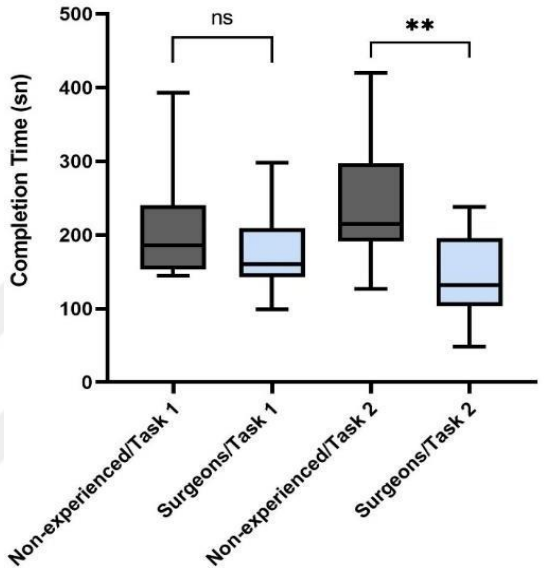


Figure 4.1 Completion time of participants

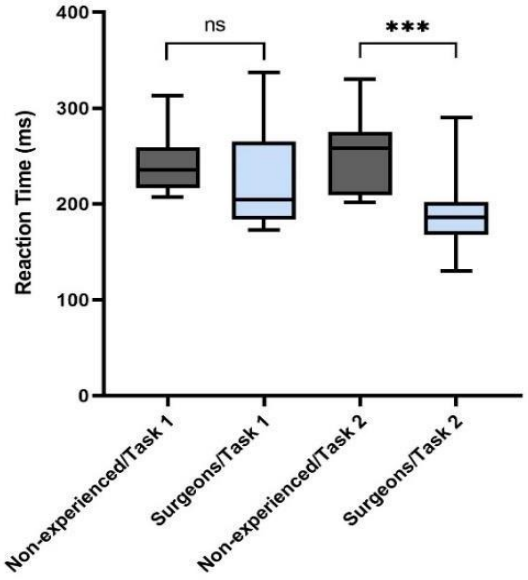


Figure 4.2 Reaction time of participants

In Figure 4.1, the comparison of task completion times between inexperienced participants and surgeons indicates that surgeons complete tasks more quickly in both groups. Notably, there is a significant difference in completion time for Task 2 between the surgeons and inexperienced group ($P < 0.05$), clearly distinguishing the two groups. The difference is even more pronounced in Task 2. Similar trends were observed in Task 1, although the group differences did not reach statistical significance. This might be because Task 2 is more challenging than Task 1. In Figure 4.2, it can be seen that for Task 1, there was no statistically significant difference in reaction times between inexperienced participants and surgeons when responding to the sound stimulus. However, it was statistically proven that surgeons reacted faster than the inexperienced group in Task 2 ($P<0.05$).

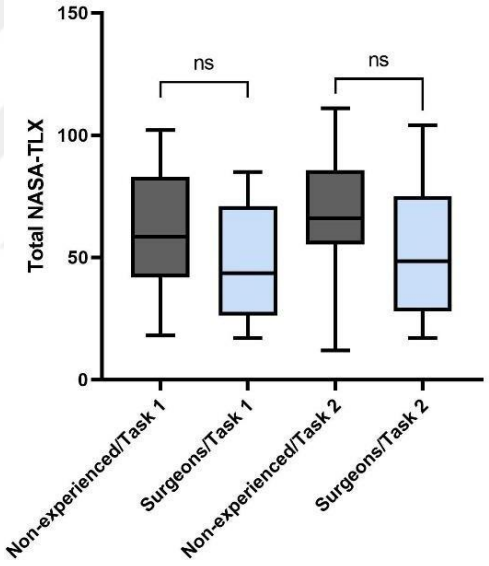


Figure 4.3 Total Nasa-TLX score of participants

Figure 4.3 shows the Nasa-TLX scores for both the inexperienced and surgeon groups in Task 1 and Task 2.

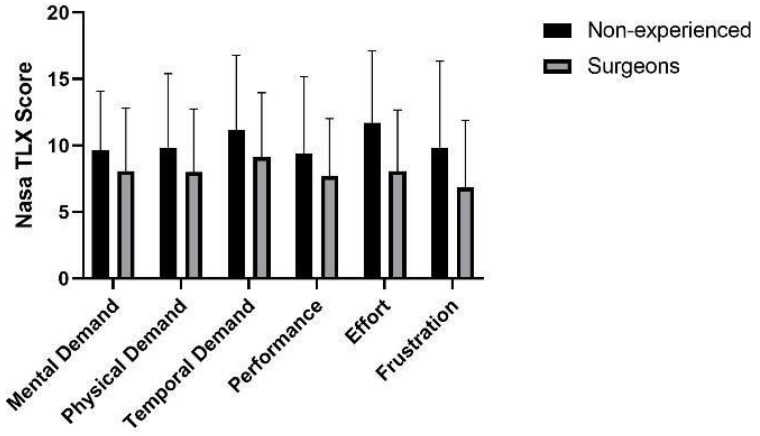


Figure 4.4 Sub Nasa-TLX score for Task 1

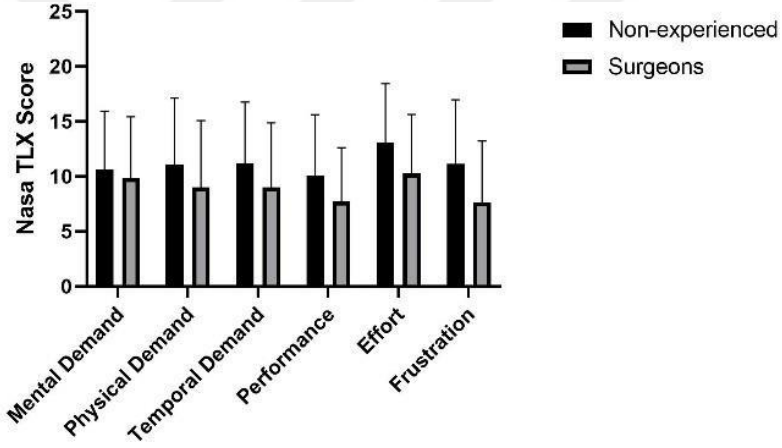


Figure 4.5 Sub Nasa-TLX score for Task 2

Figure 4.4 shows the Sub Nasa-TLX scores for Task 1 of both the inexperienced and surgeon groups. It's evident that the inexperienced group scored higher than the surgeon group in all Sub Nasa-TLX scores. Figure 4.5 illustrates the Sub Nasa-TLX scores for Task 2 among the inexperienced and surgeon groups. Once again, it's noticeable that the inexperienced group achieved higher scores than the surgeon group in all Sub Nasa-TLX scores.

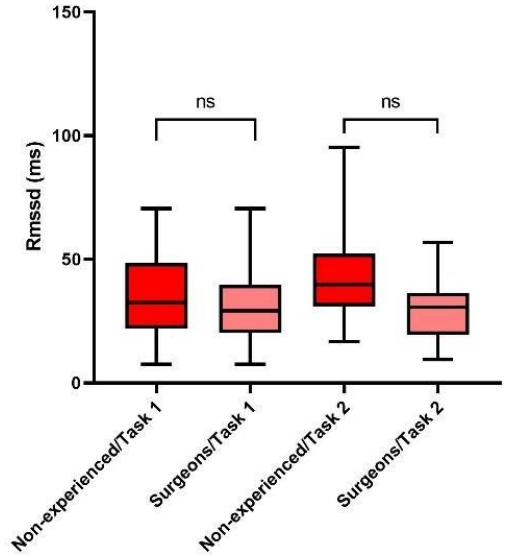


Figure 4.6 RMSSD value of participants

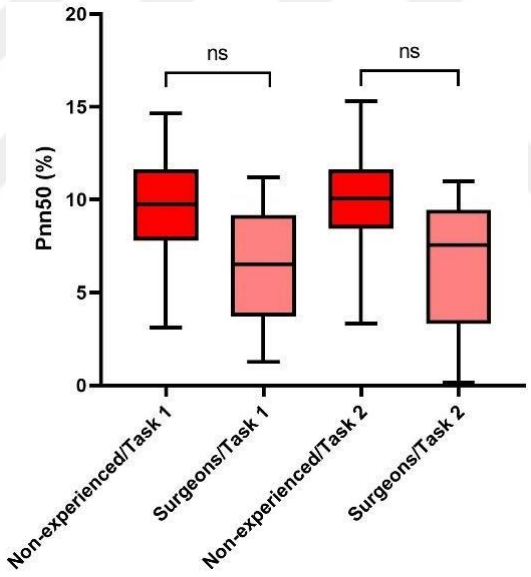


Figure 4.7 Pnn50 value of participants

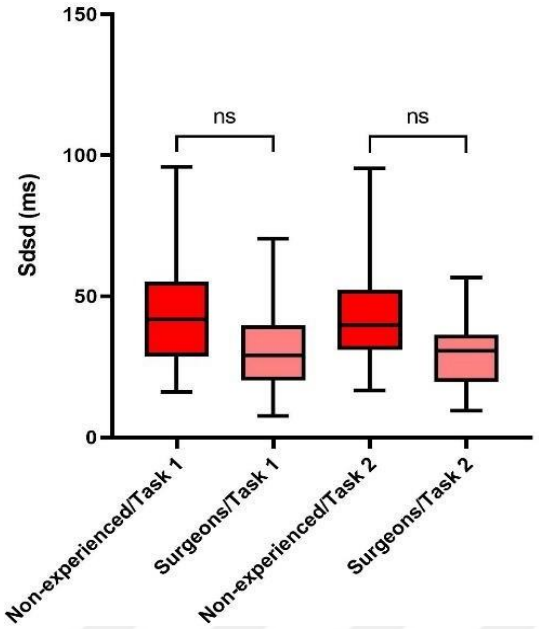


Figure 4.8 SDSD value of all participants for Task 1 and Task 2

Figure 4.6 shows the RMSSD values for both Task 1 and Task 2 in the inexperienced and surgeon groups. Notably, the RMSSD values were observed to be higher in the inexperienced group. Figure 4.7 illustrates the Pnn50 values for Task 1 and Task 2 among the inexperienced and surgeon groups. Similarly, the Pnn50 values were noted to be higher in the inexperienced group. Figure 4.8 shows the SDSD values for Task 1 and Task 2 in both the inexperienced and surgeon groups. Consistently, the SDSD values were higher in the inexperienced group.

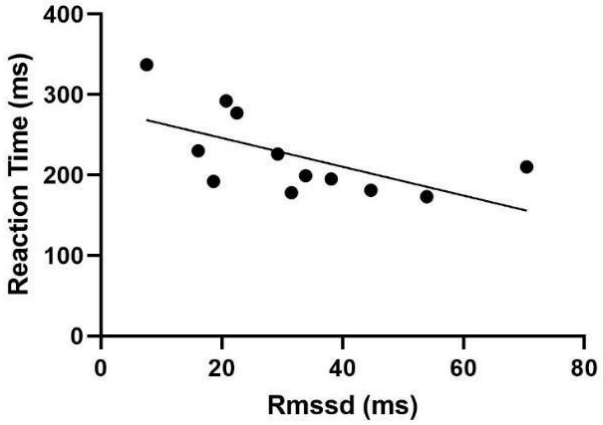


Figure 4.9 Reaction Time v RMSSD for surgeons during Task 1. The solid black line indicates the linear best fit (R2 = 0.37, p= 0.03)

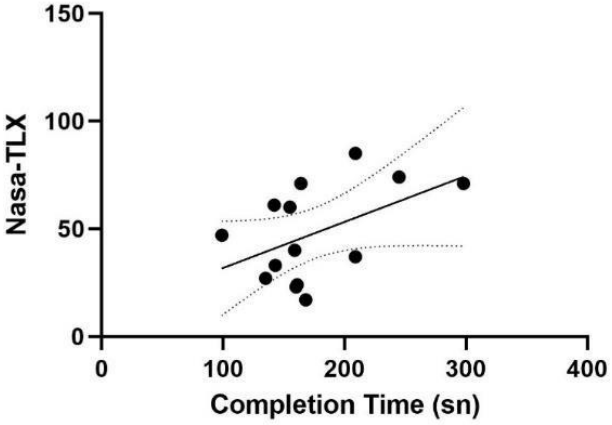


Figure 4.10 Completion Time v Total Nasa-TLX for surgeons during Task 1. The solid black line indicates the linear best fit (R2 = 0.23, p= 0.07)

As depicted in Figure 4.9, a notable statistical differences was evident between the reaction times of the surgeons in Task 1 and the RMSSD value (P<0.05). The study observed an inverse relationship, where the RMSSD value increased as the reaction time of the surgeons shortened during the task. Figure 4.10 illustrates a statistically significant correlation (P<0.05) between the completion time of Task 1 and the Nasa-TLX scores of the surgeons. These finding highlights that as the completion time increased, the Nasa-TLX scores also exhibited a corresponding increase. This can be attributed to heightened stress during the task and the delayed completion of the task.

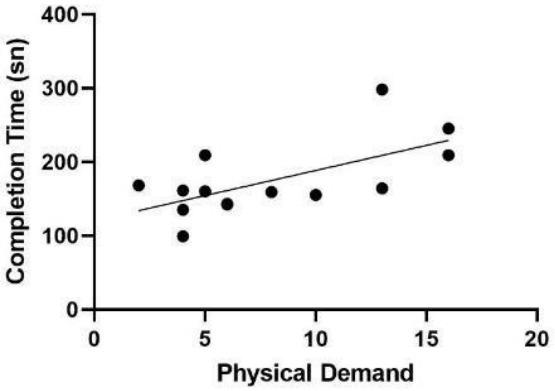


Figure 4.11 Physical Demand v Completion Time for surgeons during task 1. The solid black line indicates the linear best fit (R2 = 0.41, p= 0.01)

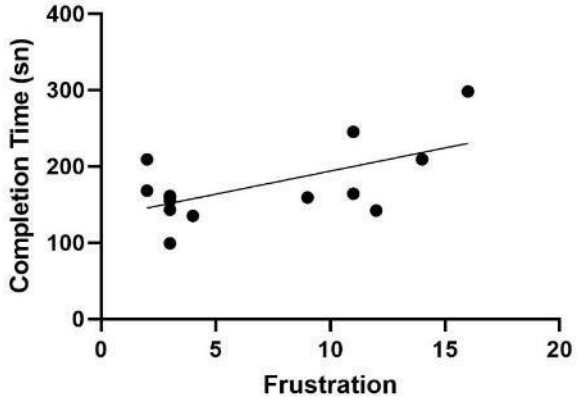


Figure 4.12 Frustration v Completion Time for surgeons during task 1. The solid black line indicates the linear best fit (R2 = 0.36, p= 0.02)

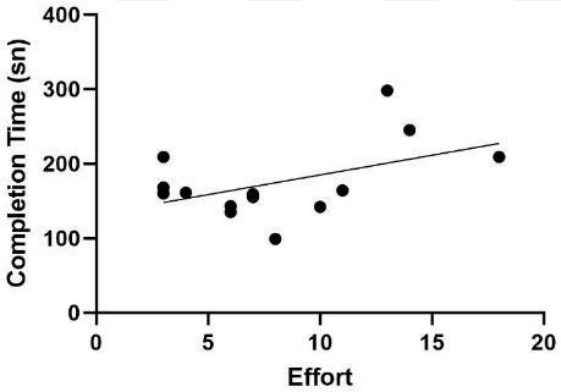


Figure 4.13 Effort vs Completion Time for surgeons during Task 1. The solid black line indicates the linear best fit (R2 = 0.23, p= 0.08)

Figure 4.11 shows the connection between the time surgeons took to complete Task 1 and the level of physical demand. Through statistical analysis, a correlation was identified between these factors ($P<0.05$). This finding indicates that as the completion time increases, the physical demand also rises. In Figure 4.12, a noteworthy pattern emerges, revealing that surgeons' frustration scores escalate with longer Task 1 completion times ($P<0.05$). The prolonged completion times are associated with heightened frustration scores among surgeons. This trend suggests a link between extended completion times and elevated levels of anxiety and stress. Figure 4.13 unveils a significant statistical distinction between the surgeons' Task 1 completion times and their reported effort levels ($P<0.05$). As surgeons take more time to complete the task, their reported effort scores also increase.

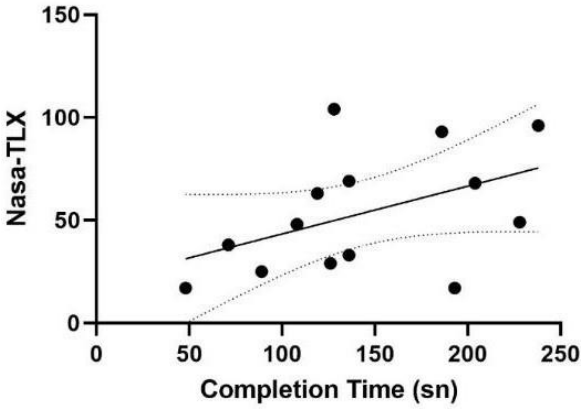


Figure 4.14 Completion Time v Total Nasa-TLX for surgeons during Task 2. The solid black line indicates the linear best fit. (R2 = 0.21, p= 0.05)

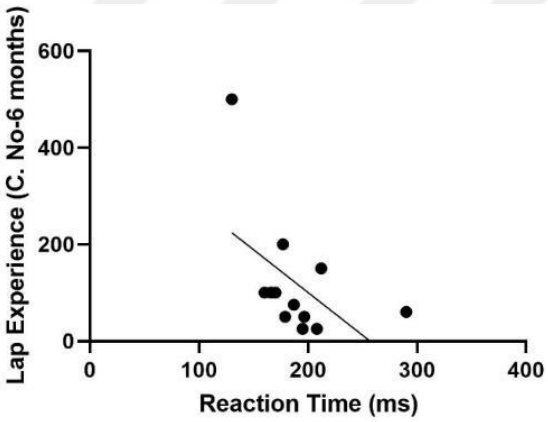


Figure 4.15 Reaction Time v Lap. Experience for surgeons during Task 2. The solid black line indicates the linear best fit. (R2 = 0.25, p= 0.05)

Figure 4.14 shows a significant correlation between the time surgeons required to finish Task 2 and their Nasa-TLX scores (P<0.05). Notably, the Nasa-TLX scores exhibited an increase in tandem with the extended completion time of Task 2 for surgeons. In Figure 4.15, a clear relationship emerges, illustrating the connection between surgeons' reaction times in Task 2 and their level of experience in laparoscopic surgery. With an increase in laparoscopic surgery experience, surgeons' reaction times were consistently reduced.

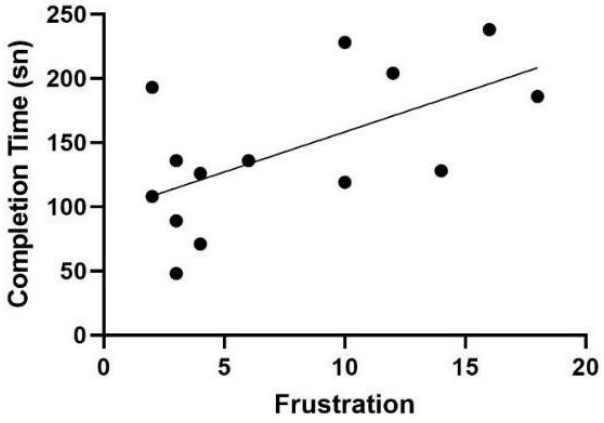


Figure 4.16 Completion Time v Frustration for surgeons during Task 2. The solid black line indicates the linear best fit. (R2 = 0.36, p= 0.02)

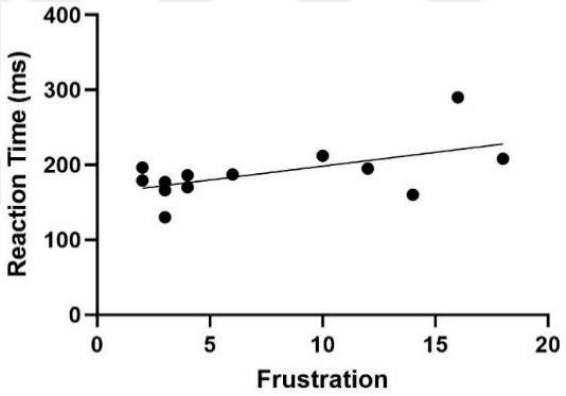


Figure 4.17 Reaction Time v Frustration for surgeons during Task 2. The solid black line indicates the linear best fit. (R2 = 0.33, p= 0.04)

Figure 4.16 illustrates the link between the time surgeons took to finish Task 2 and their frustration scores ($P<0.05$). This connection reveals that as the completion time for Task 2 increases among surgeons, their frustration scores also rise. In Figure 4.17, the graph demonstrates the relationship between surgeons' reaction times in Task 2 and their frustration scores. A clear pattern emerges: as surgeons' reaction times lengthen in Task 2, their frustration scores also increase.

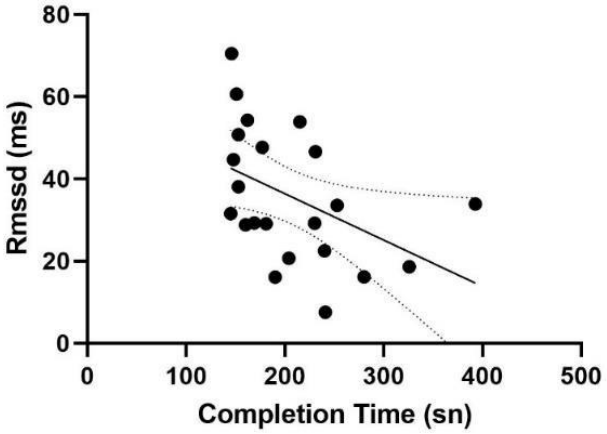


Figure 4.18 Completion Time v RMSSD for the inexperienced during Task 1. The solid black line indicates the linear best fit. (R2 = 0.20, p= 0.03)

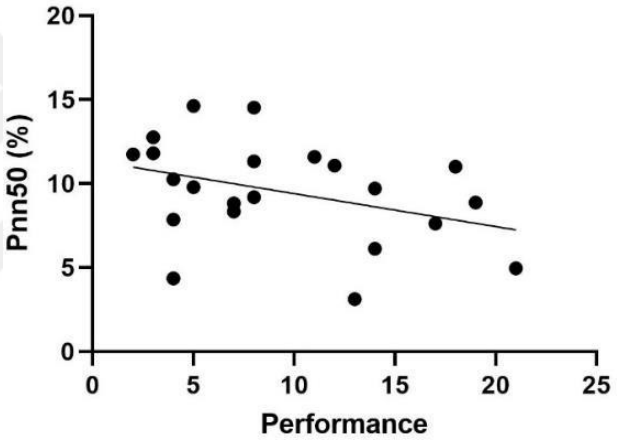


Figure 4.19 Pnn50 v Performance for the inexperienced during Task 1. The solid black line indicates the linear best fit. (R2 = 0.14, p= 0.05)

Figure 4.18 shows the connection between the time needed to finish task 1 and the RMSSD values of the inexperienced group. This connection was statistically significant (P<0.05). Notably, as completion times increased, the RMSSD values decreased. In Figure 4.19, the graph showcases the relationship between the performance score and the Pnn50 value in Task 1 for the inexperienced group (P<0.05). A clear trend is observed: as the Pnn50 value decreases, the performance score of the inexperienced group increase

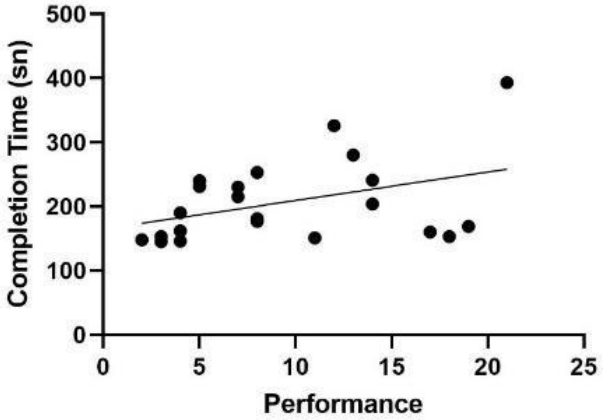


Figure 4.20 Performance v Completion Time for the inexperienced during Task 1. The solid black line indicates the linear best fit. (R2 = 0.16, p= 0.05)

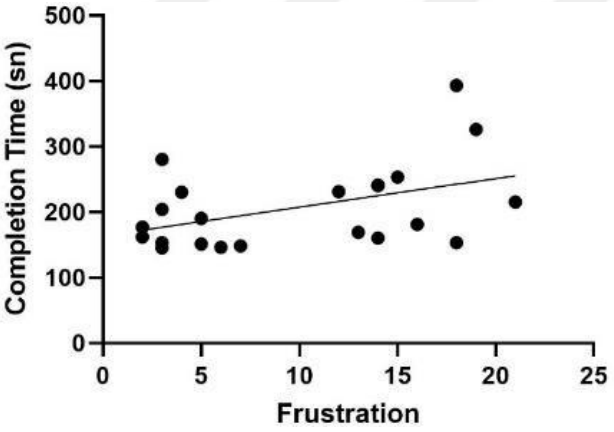


Figure 4.21 Frustration v Completion Time for the inexperienced during Task 1. The solid black line indicates the linear best fit. (R2 = 0.19, p= 0.04)

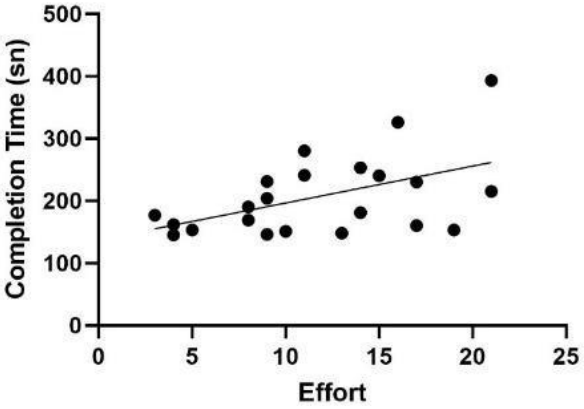


Figure 4.22 Effort v Completion Time for the inexperienced during Task 1. The solid black line indicates the linear best fit. (R2 = 0.25, p= 0.02)

Figure 4.20 illustrates a significant correlation between the completion time of Task 1 and the performance score of the inexperienced group ($P<0.05$). As the completion time increases, the performance score also increases for the inexperienced group. Figure 4.21 displays a noteworthy relationship between the completion time of Task 1 for the inexperienced group and their frustration score ($P<0.05$). It's evident that as the completion time increases, so does the frustration score. Figure 4.22 demonstrates a significant connection between the completion time of Task 1 for the inexperienced group and their effort score ($P<0.05$). As the completion time increases, the effort score also increases for the inexperienced group.

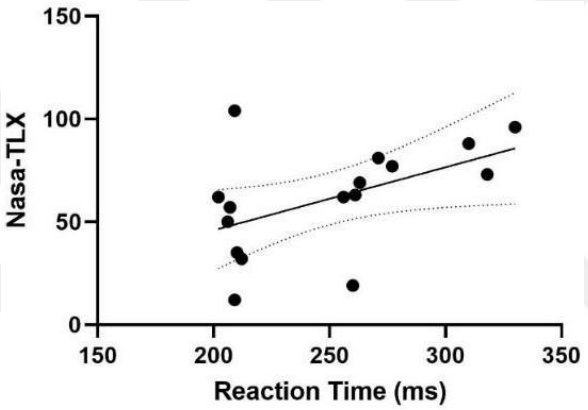


Figure 4.23 Reaction Time v Total Nasa-TLX for the inexperienced during Task 2. The solid black line indicates the linear best fit. ($R2 = 0.26$, $p= 0.04$)

Figure 4.23 illustrates the connection between the reaction time in Task 2 of the inexperienced group and the Nasa-TLX scores ($P<0.05$). It's evident that as the reaction time in Task 2 of the inexperienced group increases, the Nasa-TLX scores also increase. Figure 4.24 depicts the relationship between the performance scores and the RMSSD values in Task 2 of the inexperienced group, which was found to be statistically significant ($P<0.05$). As the RMSSD value in task 2 of the inexperienced group increases, a decrease is observed in the performance scores.

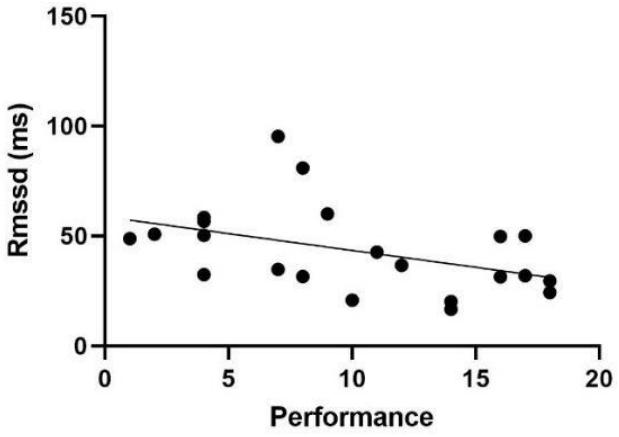


Figure 4.24 RMSSD v Performance for the inexperienced during Task 2. The solid black line indicates the linear best fit. (R2 = 0.19, p= 0.04)

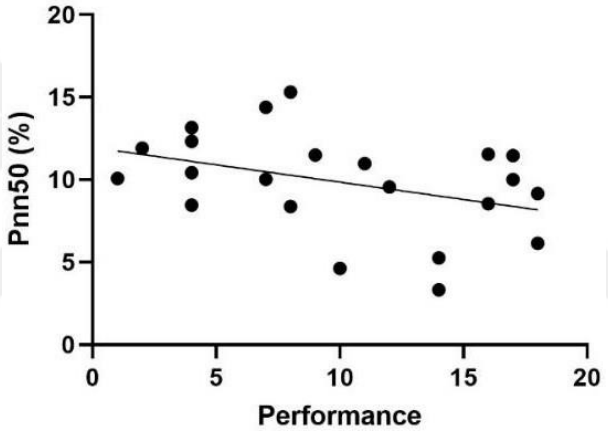


Figure 4.25 Pnn50 v Performance for the inexperienced during Task 2. The solid black line indicates the linear best fit. (R2 = 0.15, p= 0.05)

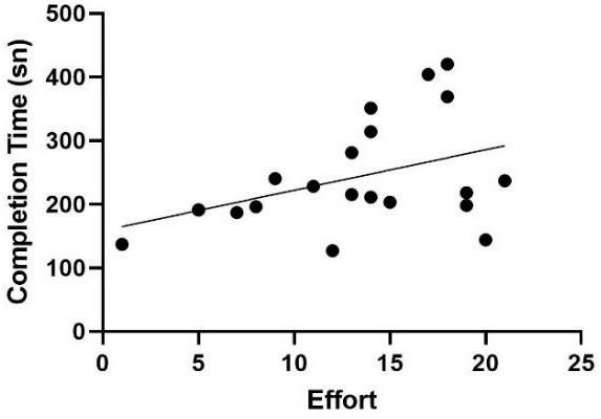


Figure 4.26 Completion Time v Effort for the inexperienced during Task 2. The solid black line indicates the linear best fit. (R2 = 0.17, p= 0.05)

Figure 4.25 illustrates the connection between the Pnn50 value and the performance score during Task 2 of the inexperienced group. It's evident that as the Pnn50 value of the inexperienced group increased during Task 2, there was a decrease in the performance score. Figure 4.26 depicts the relationship between task completion time and effort score during Task 2 of the inexperienced group. The study observed that as the time taken to complete Task 2 increased for the inexperienced group, the effort score also increased.

Figure 4.27 and Figure 4.28 indicates the functional connectivity network features of PLV, that show the changes (decrease/increase) in the connectivity network between two pairs of channels. The PLV was extracted from four frequency bands of both tasks between surgeons and inexperienced. The significant channels denoted *. in the connectivity network measurements between surgeons and inexperienced in Table 3 and in Table 4. In the result, we observed the patters differences between inexperienced and surgeons during Task1 and Task2. Figure 4.27 shows that significant functional connectivity networks in delta, theta, alpha and beta from the inexperienced to surgeons during Task 1. Our results showed that there are significant pairs in the different brain region. The statistically significant pairs are the below: **Delta:** Fz-F4, F4-C4, F4-F3, F4-P3, P4-P3. **Theta:** Fz-C3, C3-C4, C3-P4 C3-F3, C3-P3, P4-P3, F3-P3**. Alpha:** Fz-C3, F4-C4, C3-C4, F4-P3, C4-P4, C3-P3 **, P4-P3. **Beta:** F4-C4, C3-P3. Figure 11.2 shows that significant functional connectivity networks in delta, theta, alpha and beta from the inexperienced to surgeons during Task 2. Our results showed that there are significant pairs in the different brain region. The statistically significant pairs are the below: **Delta:** C3-P4, C4-P4, C3-P3, P4-P3, F3-P3. **Theta:** F3-P3, C4-P3**, C3-P3**, P4-P3**, C3-F3, FZ-C3, C3-C4, C3-P4**, C4-P4**. Alpha:** FZ-C3, F4-C4, C3-P4, C3-F3, C3-P3, P4-P3, F3-P3. **Beta**: C3-P4, C4-P4, FZ-F3, F4-F3, C3-P3, P4-P3. All bands showed decreases in the connectivity network between different brain regions The significant discrimination connectivity features between surgeons and inexperienced conditions were selected using a Wilcoxon test.

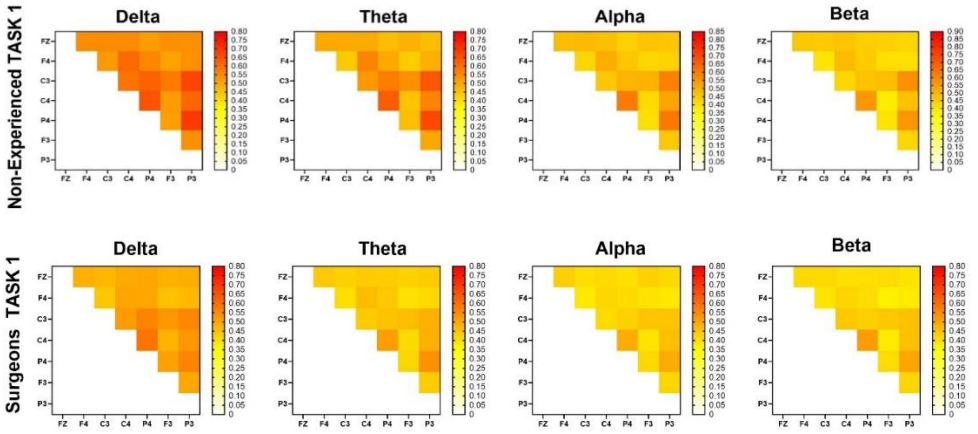


Figure 4.27 The PLV connectivity network among EEG channel pairs of inexperienced and surgeons during Task 1

Figure 4.27 shows the PLV connectivity network among EEG channel pairs of inexperienced during Task 1.

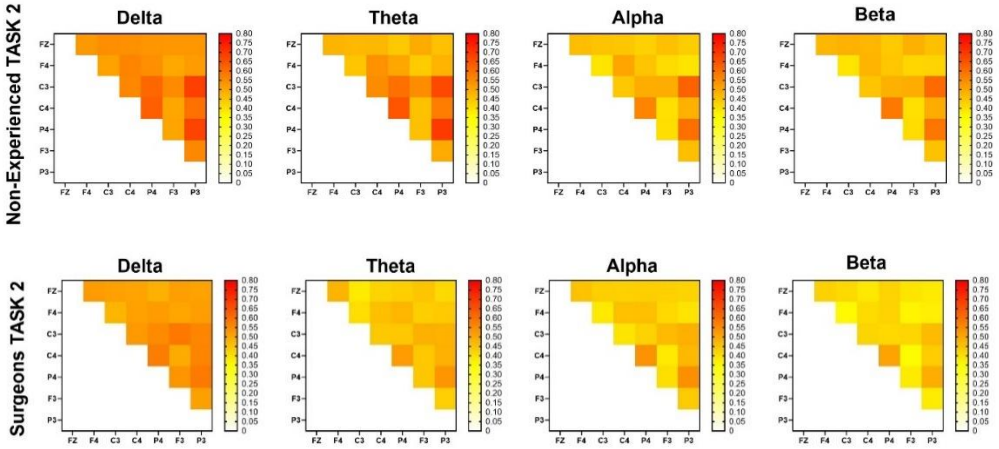


Figure 4.28 The PLV connectivity network among EEG channel pairs of inexperienced and surgeons during Task 2

Figure 4.28 shows the PLV connectivity network among EEG channel pairs of surgeons during Task 2.

Table 4.1 Significant PLV-Pairs in Connectivity in Task 1

Task 1-Delta			Task 1-Theta			Task 1-Alpha			Task 1-Beta		
Pairs	P Value		Pairs	P Value		Pairs	P Value		Pairs	P Value	
Fz-F4	<0.05	*	Fz-F4	>0.05		Fz-F4	>0.05		Fz-F4	>0.05	
Fz-C3	>0.05		Fz-C3	<0.05	*	Fz-C3	<0.05	*	Fz-C3	>0.05	
F4-C3	>0.05		F4-C3	>0.05		F4-C3	>0.05		F4-C3	>0.05	
Fz-C4	>0.05		Fz-C4	>0.05		Fz-C4	>0.05		Fz-C4	>0.05	
F4-C4	<0.05	*	F4-C4	>0.05		F4-C4	<0.05	*	F4-C4	<0.05	*
C3-C4	>0.05		C3-C4	<0.05	*	C3-C4	<0.05	*	C3-C4	>0.05	
Fz-P4	>0.05		Fz-P4	>0.05		Fz-P4	>0.05		Fz-P4	>0.05	
F4-P4	>0.05		F4-P4	>0.05		F4-P4	>0.05		F4-P4	>0.05	
C3-P4	>0.05		C3-P4	<0.05	*	C3-P4	>0.05		C3-P4	>0.05	
C4-P4	>0.05		C4-P4	>0.05		C4-P4	<0.05	*	C4-P4	>0.05	
Fz-F3	>0.05		Fz-F3	>0.05		Fz-F3	>0.05		Fz-F3	>0.05	
F4-F3	<0.05	*	F4-F3	>0.05		F4-F3	>0.05		F4-F3	>0.05	
C3-F3	>0.05		C3-F3	<0.05	*	C3-F3	>0.05		C3-F3	>0.05	
C4-F3	>0.05		C4-F3	>0.05		C4-F3	>0.05		C4-F3	>0.05	
P4-F3	>0.05		P4-F3	>0.05		P4-F3	>0.05		P4-F3	>0.05	
Fz-P3	>0.05		Fz-P3	>0.05		Fz-P3	>0.05		Fz-P3	>0.05	
F4-P3	<0.05	*	F4-P3	>0.05		F4-P3	<0.05	*	F4-P3	>0.05	
C3-P3	>0.05		C3-P3	<0.05	*	C3-P3	<0.01	**	C3-P3	<0.05	*
C4-P3	>0.05		C4-P3	>0.05		C4-P3	>0.05		C4-P3	>0.05	
P4-P3	<0.05	*	P4-P3	<0.05	*	P4-P3	<0.05	*	P4-P3	>0.05	
F3-P3	>0.05		F3-P3	<0.05	*	F3-P3	>0.05		F3-P3	>0.05	

Table 4.2 Significant PLV-Pairs in Connectivity in Task 2

Task 2-Delta			Task 2-Theta			Task 2-Alpha			Task 2-Beta		
Pairs	P Value		Pairs	P Value		Pairs	P Value		Pairs	P Value	
Fz-F4	>0.05		Fz-F4	>0.05		Fz-F4	>0.05		Fz-F4	>0.05	
Fz-C3	>0.05		Fz-C3	<0.05	*	Fz-C3	<0.05		Fz-C3	>0.05	
F4-C3	>0.05		F4-C3	>0.05		F4-C3	>0.05		F4-C3	>0.05	
Fz-C4	>0.05		Fz-C4	>0.05		Fz-C4	>0.05		Fz-C4	>0.05	
F4-C4	>0.05		F4-C4	>0.05		F4-C4	<0.05	*	F4-C4	>0.05	
C3-C4	>0.05		C3-C4	<0.05	*	C3-C4	>0.05		C3-C4	>0.05	
Fz-P4	>0.05		Fz-P4	>0.05		Fz-P4	>0.05		Fz-P4	>0.05	
F4-P4	>0.05		F4-P4	>0.05		F4-P4	>0.05		F4-P4	>0.05	
C3-P4	<0.05	*	C3-P4	<0.01	**	C3-P4	<0.05	*	C3-P4	<0.05	*
C4-P4	<0.05	*	C4-P4	<0.05	*	C4-P4	>0.05		C4-P4	<0.05	*
Fz-F3	>0.05		Fz-F3	>0.05		Fz-F3	>0.05		Fz-F3	<0.05	*
F4-F3	>0.05		F4-F3	>0.05		F4-F3	>0.05		F4-F3	<0.05	*
C3-F3	>0.05		C3-F3	<0.05	*	C3-F3	<0.05		C3-F3	>0.05	
C4-F3	>0.05		C4-F3	>0.05		C4-F3	>0.05		C4-F3	>0.05	
P4-F3	>0.05		P4-F3	>0.05		P4-F3	>0.05		P4-F3	>0.05	
Fz-P3	>0.05		Fz-P3	>0.05		Fz-P3	>0.05		Fz-P3	>0.05	
F4-P3	>0.05		F4-P3	>0.05		F4-P3	>0.05		F4-P3	>0.05	
C3-P3	<0.05	*	C3-P3	<0.01	**	C3-P3	<0.05	*	C3-P3	<0.01	**
C4-P3	>0.05		C4-P3	<0.01	**	C4-P3	>0.05		C4-P3	>0.05	
P4-P3	<0.05	*	P4-P3	<0.01	**	P4-P3	<0.01	**	P4-P3	<0.01	**
F3-P3	<0.05	*	F3-P3	<0.05	*	F3-P3	<0.05	*	F3-P3	>0.05	

5. DISCUSSION

The main purpose of this thesis is to investigate mental workload patterns during laparoscopic training for both experienced surgeons and inexperienced participants. We looked into differences in brain activity, physiology, and behavior based on task difficulty (easy/difficult), comparing experienced surgeons with inexperienced participants.

In the analysis of task completion times among participants, it was evident that the surgeons demonstrated superior speed compared to the inexperienced group. This observed disparity can be attributed to the prior experience possessed by the surgeons. Statistical tests conducted confirmed a significant difference in completion times for Task 2 between the surgeon and inexperienced groups. Upon comparing task 1 and Task 2, it became apparent that the inexperienced group took longer to complete Task 2. This could be attributed to the elevated strain experienced during laparoscopic tasks with varying workloads. It is plausible to assert that the surgeons' swifter completion of Task 2 can be attributed to their experience or difficulty task. In a study, it was demonstrated that as the task difficulty increased with a continuous auditory stimulus, there was a difference in task completion. This finding aligns with the literature. It was observed that as the task difficulty increased, the completion time also increased by 14%. (Siu et al., 2010)

Examining the reaction times for the secondary auditory stimulus sent during the main task, we found that there was no significant difference in participants' reaction times during Task 1. However, a statistically significant difference in reaction times emerged during Task 2. Notably, the reaction times of the surgeons during Task 2 were faster compared to the inexperienced group. This divergence is likely due to the increased difficulty and higher workload of Task 2, as reflected in the completion time. In one study, it was observed that experienced surgeons were less affected by distracting elements while performing a surgical skill.(Hsu et al., 2008) It is believed that the experience of performing a specific technical skill leads to automatization, requiring less conscious effort or cognitive capacity, thereby reducing the impact of distracting events. (Schneider & Shiffrin, 1977)

After completing the tasks, participants filled out the Nasa-TLX questionnaire, a widely used standardized metric for assessing workload. This multidimensional scale enabled participants to rate their perceived mental demand, physical demand, temporal demand, effort, performance, and frustration levels. The outcomes gathered from this scale provided valuable insights into participants' experiences of workload across different dimensions. NASA-TLX is currently the most widely used tool for assessing mental workload(Noyes & Bruneau, 2007; Rutledge et al., 2009) This method allows for the multi-dimensional rating of tasks, hence it has been proven to be useful due to its diagnostic capacity regarding potential sources of workload(EVA et al., 2010) In this thesis, correlations have been found among Nasa-TLX scores. Differences in Nasa-TLX workload scores were anticipated between high and low workload conditions. As expected, the results indicated variations both between tasks and among participant groups. In terms of Nasa-TLX scores, we anticipated lower scores for task 2, which had a high workload, and higher scores for Task 1, which had a low workload. Upon reviewing the graphs generated for Task 1 and Task 2, we found that the results aligned with our expectations. Nasa-TLX scores for both surgeons and the inexperienced group were lower for Task 2 compared to Task 1, reflecting the higher workload of Task 2. In the overall Nasa-TLX assessment, the scores for surgeons were lower than those of the inexperienced group. This discrepancy can be attributed to the surgeons' perception of reduced stress due to their experience in task performance. We interpret the higher scores among the inexperienced group compared to the surgeons as an indication that they consider themselves less successful in task completion. Upon analyzing the Sub Nasa-TLX graphs, it's evident that surgeons consistently scored lower than the inexperienced group for both Task 2 and Task 1 across all subscales. Notably, the 'effort' subscale of the Nasa-TLX showed a significant difference between surgeons and inexperienced participants. Surgeons rated their effort lower during tasks compared to the inexperienced group, likely due to their experience, resulting in a perceived lower level of effort, while the inexperienced group might exert more effort due to their lack of experience.

Physiological measures are often considered indirect indicators of workload, as they reflect the connection between cognitive activity and performance. In this study, EEG and ECG were utilized as primary physiological measures to assess and infer mental

workload. Heart rate variability (HRV), obtained from heart rate or electrocardiogram (ECG), is widely used and has a long history in the assessment of mental workload. (Jorna, 1992; Roscoe, 1992) It has been shown that HRV is a more reliable and temporally correlated measurement for stress compared to other indicators. (Moorthy et al., 2003)

We observed that the HRV values of the surgeons were lower than the inexperienced group in. When the mental workload and heart rate variability were examined in the literature, it was seen that the surgeons has high HRV parameters (RMSSD, SDSD and pnn50). This is opposite of our finding. We explain this differences sınce surgeons take seriously to complete task. This situation changes the HRV values. We didn't see much differences HRV parameters based on task difficulty level. As the task difficulty increases, heart rate decreases along with significant changes in various linear and nonlinear HRV parameters; this is attributed to mental stress induced by the influence of the autonomic nervous system during mental arithmetic. (Singh et al., 2019)It has been observed that HRV parameters decrease during more challenging tasks. (Luque-Casado et al., 2016; VELTMAN & GAILLARD, 1998)

In this study, we utilized linear regression analysis to explore the connections between variables. The graphs depicting the relationships between participants' subjective, Nasa-TLX, and physiological values are presented in the results section. When investigating the connection between completion time and other variables, we observed a positive correlation with total Nasa-TLX and Sub Nasa-TLX scores. Longer task completion times were linked to higher stress levels. This implies that as completion time increases, Nasa-TLX scores rise, suggesting heightened mental workload among participants. Examining the link between completion times and heart rate variables, we noted discrepancies in both groups. As participants' completion times increased, their heart rate variables decreased. Comparing these results with previous findings and existing literature, decreased heart rate variables have been associated with increased mental workload. Therefore, we interpret our findings as follows: 'As participants' task completion times increase, heart rate variables decrease, indicating elevated mental workload levels". When examining the relationship between the reaction times of the inexperienced group during Task 2 and their total Nasa-TLX scores, an interesting trend emerged. As reaction

times increased, Nasa-TLX scores also increased. This pattern could be attributed to heightened mental workload during the auditory task. In Task 2, the reaction times of the surgeons correlated with the Sub Nasa-TLX subscale of frustration. As reaction times increased, frustration scores also rose. Similarly, significant differences emerged between reaction times and RMSSD. As reaction times increased, RMSSD values decreased. These observations were indicative of increased mental workload.

The study also brought to light significant disparities between the number of laparoscopic operations performed by surgeons in the last 6 months and their reaction times. Remarkably, as surgeons gained more experience with surgical procedures, their reaction times increased. This intriguing finding suggests that surgeons' experience contributes to improved reaction times. While there may not be a definitive answer regarding the impact of distracting elements on varying levels of experience, experimental studies reviewed in one study indicate that inexperienced subjects have difficulty maintaining focus on their tasks when their attention is diverted. (Mentis et al., 2016). Furthermore, noteworthy differences surfaced between the Sub Nasa-TLX scale performance and Pnn50 values in both Task 1 and Task 2 of the inexperienced group. As the performance score increased, the Pnn50 value decreased. This phenomenon can be explained by the fact that the inexperienced group rated their performance higher when they experienced greater mental workload. Surgeons reported poorer perceived performance during cases with unexpectedly high difficulty ($P < 0.001$).

Our results seem to show a generalized decrease of the connectivity from inexperienced to surgeons for both Task 1(Easy) and Task 2(Difficult). We observed the less significant connections in Task 2 compared to Task 1 in Alpha band. This finding aligns with the traditional understanding of the alpha rhythm, suggesting that a reduction in alpha activity may have contributed to a generalized cortical activation. This reduced alpha activity could result in decreased inhibition of irrelevant information, potentially serving as an indirect indicator of increased workload linked to the task's complexity. These results also point in the same direction of previously published studies, concluding that tasks with increased difficulty are associated with a decrease of alpha connectivity. Previous studies have observed a disruption in alpha synchronization in the brain area related to the task

(e.g., lower alpha power during highly demanding tasks), while synchronization has been observed in brain areas unrelated to the task. (Fukuda et al., 2015; Pfurtscheller et al., 1996; Pfurtscheller & Klimesch, 1990)In one study, it was shown to decrease in the frontal area during more challenging tasks. (Lim et al., 2023a)

We observed the more significant connections in Task 2 (Difficult) compared to Task1 (Easy) in beta and theta bands. The more significant connections in beta bands may be associated with greater attention, preparation dedicated to the difficult task and high mental workload. Additionally, it is established in the literature that the beta band is closely connected to motor function, with a noted reduction in beta activity occurring during the preparation and execution of voluntary movements. Despite lacking a complete understanding, the presence of a distinct beta connectivity pattern gives rise to the hypothesis that there may be cognitive differences between individuals highly experienced in a specific task and those with less experience. Beta connectivity was also expected to increase from Task 1 to Task 2. Another research study found that when performing tasks involving hand movement, simply repeating the same motion didn't have a major impact on beta power. However, both getting ready for a mental task and paying more attention were linked to increased connectivity in the sensorimotor area within this specific frequency range.(Darch et al., 2020) In one of the surgical studies, participants exhibited an increase in beta power in response to increased workload during tasks. (Lim et al., 2023b) However, caution should be exercised when using beta band power as an indicator of workload because changes in beta band powers have also been observed with increasing cognitive task demands, and the role of beta band power remains uncertain. (Spitzer & Haegens, 2017). In this thesis, we observed that theta bands were also affected by difficulty level. There was a more significant connections in the theta from Task 1 to Task 2. Jochen et al, demonstrated that in motor tasks, the theta connectivity in frontal and central regions increased when the performance improved. Increases in theta with both increased task difficulty and with practice suggests that focusing attention required more effort. Examination of previous studies reveals evidence that theta band power increases with more challenging tasks. (Gevins & Smith, 2003; Lim et al., 2023c)

5.1 Limitation

This study comes with several limitations that need to be acknowledged. Firstly, there were also large differences in age and gender distribution among the participants. This may restrict the generalizability of the findings to a broader population. A larger and more similar demographics sample could provide a more comprehensive understanding. Secondly, the study employed a cross-sectional design, which captures data at a single point in time. This design may not allow for the examination of causal relationships between variables, and it may not capture changes over time. Thirdly, our control group consists of only inexperienced participants, we assume that different known level of expertise among the participants would be useful. Fourthly, task difficulty level between Task1 and Task 2 is low. We would assume that the high-level differences in the task difficulty provide the sharp differences in behavioral and physiological data. Finally, we have the limited number of EEG electrodes due to restriction of mobile system. Scalp EEG has a relatively low sensitivity so this makes also differences in the analysis.

6. CONCLUSION

In this thesis, we evaluated and investigated the physiological and behavioral differences between surgeons and inexperienced subjects during two different tasks which has the different difficulty level and cognitive load. This thesis showed that the higher the degree of difficulty of the task will reveal the physiological and behavioral differences between surgeons and inexperienced group. Given the neurocognitive complexity of surgery skill and mental workload, multimodal approaches are required. In the future, we plan to combine EEG+fNIRS for understanding the cognitive processing of surgeons in the operation room.

REFERENCES

Aggarwal, R., Hance, J., & Darzi, A. (2004). Surgical education and training in the new millennium. In Surgical Endoscopy and Other Interventional Techniques (Vol. 18, Issue 10,pp. 1409–1410). https://doi.org/10.1007/s00464-004-8238-x

Alan Gevins, Michael E. Smith, Harrison Leong, Linda McEvoy, Susan Whitfield, Robert Du, & Georgia Rush. (1998). Monitoring Working Memory Load during Computer-Based Tasks with EEG Pattern Recognition Methods. https://doi.org/10.1518/001872098779480578

Alobid, I., Juan De Pablo, ;, Mullol, J., Centellas, S., Parramon, G., Carrasco, J., Armario, A., & Bernal-Sprekelsen, M. (2011). Increased Cardiovascular and Anxiety Outcomes but Not Endocrine Biomarkers of Stress During Performance of Endoscopic Sinus Surgery A Pilot Study Among Novice Surgeons. In Arch Otolaryngol Head Neck Surg (Vol. 137, Issue 5). https://jamanetwork.com/

Anton, N. E., Howley, L. D., Pimentel, M., Davis, C. K., Brown, C., & Stefanidis, D. (2016).

Effectiveness of a mental skills curriculum to reduce novices' stress. Journal of Surgical Research, 206(1), 199–205. https://doi.org/10.1016/j.jss.2016.07.019

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using Electroencephalography to Measure Cognitive Load. In Educational Psychology Review (Vol. 22, Issue 4, pp. 425– 438). https://doi.org/10.1007/s10648-010-9130-y

Antonenko, P., Paas, F., Grabner, R., & Van Gog, T. (2010). Using Electroencephalography to Measure Cognitive Load. Source: Educational Psychology Review, 22(4), 425–438. https://doi.org/10.1007/s

Arora, S., Sevdalis, N., Aggarwal, R., Sirimanna, P., Darzi, A., & Kneebone, R. (2010). Stress impairs psychomotor performance in novice laparoscopic surgeons. Surgical Endoscopy, 24(10), 2588–2593. https://doi.org/10.1007/s00464-010-1013-2

Arora, S., Sevdalis, N., Nestel, D., Woloshynowych, M., Darzi, A., & Kneebone, R. (2010). The impact of stress on surgical performance: A systematic review of the literature. Surgery,147(3). https://doi.org/10.1016/j.surg.2009.10.007

Aydore, S., Pantazis, D., & Leahy, R. M. (2013). A note on the phase locking value and its properties. NeuroImage, 74, 231–244. https://doi.org/10.1016/j.neuroimage.2013.02.008

Berguer, R., & Smith, W. (2006). An Ergonomic Comparison of Robotic and Laparoscopic Technique: The Influence of Surgeon Experience and Task Complexity. Journal of Surgical Research, 134(1), 87–92. https://doi.org/10.1016/j.jss.2005.10.003

Berguer, R., Smith, W. D., & Chung, Y. H. (2001). Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery. Surgical Endoscopy,15(10), 1204–1207. https://doi.org/10.1007/s004640080030

Bertrand, J. A., McIntosh, A. R., Postuma, R. B., Kovacevic, N., Latreille, V., Panisset, M., Chouinard, S., & Gagnon, J. F. (2016). Brain Connectivity Alterations Are Associated with the Development of Dementia in Parkinson's Disease. Brain Connectivity, 6(3), 216–224. https://doi.org/10.1089/brain.2015.0390

Böhm, B., Rötting, N., Schwenk, W., Grebe, S., & Mansmann, U. (2001). A Prospective Randomized Trial on Heart Rate Variability of the Surgical Team During Laparoscopic and Conventional Sigmoid Resection. In Arch Surg (Vol. 136). http://archsurg.jamanetwork.com/

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. In Neuroscience and Biobehavioral Reviews (Vol. 44, pp. 58–75). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2012.10.003

Borghini, G., Isabella, R., Vecchiato, G., Toppi, J., Astolfi, L., Caltagirone, C., & Babiloni, F. (2011). Brainshield: HREEG study of perceived pilot mental workload. Italian Journal of Aerospace Medicine.

Brookings, J. B., Wilsonh'*, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. In Biological Psychology (Vol. 42).

Bunge, S. A., & Kahn, I. (2009). Cognition: An Overview of Neuroimaging Techniques. https://doi.org/10.1016/B978-008045046-9.00298-9

Byrne, J. V., Beltechi, R., Yarnold, J. A., Birks, J., & Kamran, M. (2010). Early experience in the treatment of intra-cranial aneurysms by endovascular flow diversion: A multicentre prospective study. PLoS ONE, 5(9), 1–8. https://doi.org/10.1371/journal.pone.0012492

Carswell, C. M., Clarke, D., & Seales, W. B. (2005). Assessing mental workload during laparoscopic surgery. Surgical Innovation, 12(1), 80–90. https://doi.org/10.1177/155335060501200112

Causse, M., Lepron, E., Mandrick, K., Peysakhovich, V., Berry, I., Callan, D., & Rémy, F. (2022).

Facing successfully high mental workload and stressors: An fMRI study. Human Brain Mapping, 43(3), 1011–1031. https://doi.org/10.1002/hbm.25703

Chen, S. W., Liaw, J. W., Chang, Y. J., Chuang, L. L., & Chien, C. T. (2015). Combined heart rate variability and dynamic measures for quantitatively characterizing the cardiac stress status during cycling exercise. Computers in Biology and Medicine, 63, 133–142. https://doi.org/10.1016/j.compbiomed.2015.05.026

Crewther, B. T., Shetty, K., Jarchi, D., Selvadurai, S., Cook, C. J., Leff, D. R., Darzi, A., & Yang, G. Z. (2016). Skill acquisition and stress adaptations following laparoscopic surgery training and detraining in novice surgeons. Surgical Endoscopy, 30(7), 2961–2968. https://doi.org/10.1007/s00464-015-4584-0

Czyewska, E., Kiczka, K., Czarnecki, A., & Pokinko, P. (1983). The Surgeon's Mental Load During Decision Making at Various Stages of Operations. In European Journal of Applied Physiology and Occupational Physiology (Vol. 51). Springer-

Verlag.

Darch, H. T., Cerminara, N. L., Gilchrist, I. D., & Apps, R. (2020). Pre-movement changes in sensorimotor beta oscillations predict motor adaptation drive. Scientific Reports, 10(1), 17946. https://doi.org/10.1038/s41598-020-74833-z

Di Biase, L., Ricci, L., Caminiti, M. L., Pecoraro, P. M., Carbone, S. P., & Di Lazzaro, V. (2023).

Quantitative High Density EEG Brain Connectivity Evaluation in Parkinson's Disease: The Phase Locking Value (PLV). Journal of Clinical Medicine, 12(4). https://doi.org/10.3390/jcm12041450

Dussault, C., Jouanin, J.-C., Philippe, M., & Guezennec, C.-Y. (2005). EEG and ECG changes during simulator operation reflect mental workload and vigilance. Aviation, Space, andEnvironmental Medicine, 76(4), 344–351.

Eva, D. R., Susana, R. V., Jesus, M. G., & Lourdes, L. M. (2010). Estudio Psicométrico del Índice de Carga Mental NASA-TLX con una Muestra de Trabajadores Españoles. Revista de Psicología Del Trabajo y de Las Organizaciones, 26(3), 191–199. https://doi.org/10.5093/tr2010v26n3a3

Fan, X., Zhao, C., Zhang, X., Luo, H., & Zhang, W. (2020). Assessment of mental workload basedon multi-physiological signals. Technology and Health Care, 28(S1), S67–S80. https://doi.org/10.3233/THC-209008

Fukuda, K., Mance, I., & Vogel, E. K. (2015). α Power Modulation and Event-Related Slow Wave Provide Dissociable Correlates of Visual Working Memory. The Journal of Neuroscience, 35(41), 14009–14016. https://doi.org/10.1523/JNEUROSCI.5003-14.2015

Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. Theoretical Issues in Ergonomics Science, 4(1–2), 113–131. https://doi.org/10.1080/14639220210159717

Halsband, U., & Lange, R. K. (2006). Motor learning in man: A review of functional and clinical studies. Journal of Physiology Paris, 99(4–6), 414–424. https://doi.org/10.1016/j.jphysparis.2006.03.007

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. Proceedings of the Human Factors and Ergonomics Society, 904–908. https://doi.org/10.1177/154193120605000909

Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association, 28(2), 194–200. https://doi.org/10.1037/a0013145

Heine, T., Lenis, G., Reichensperger, P., Beran, T., Doessel, O., & Deml, B. (2017).

Electrocardiographic features for the measurement of drivers' mental workload. Applied Ergonomics, 61, 31–43. https://doi.org/10.1016/j.apergo.2016.12.015

Hsu, K. E., Man, F.-Y., Gizicki, R. A., Feldman, L. S., & Fried, G. M. (2008). Experienced surgeons can do more than one thing at a time: effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons. Surgical Endoscopy, 22(1), 196–201. https://doi.org/10.1007/s00464-007-9452-0

Huang, S., Li, J., Zhang, P., & Zhang, W. (2018). Detection of mental fatigue state with wearable ECG devices. International Journal of Medical Informatics, 119, 39–46. https://doi.org/10.1016/j.ijmedinf.2018.08.010

Hubert, N., Gilles, M., Desbrosses, K., Meyer, J., Felblinger, J., & Hubert, J. (2013). Ergonomic assessment of the surgeon's physical workload during standard and robotic assisted laparoscopic procedures. International Journal of Medical Robotics and Computer Assisted Surgery, 9(2), 142–147. https://doi.org/10.1002/rcs.1489

Jian, W., Chen, M., & McFarland, D. J. (2017). Use of phase-locking value in sensorimotor rhythm-based brain–computer interface: zero-phase coupling and effects of spatial filters. Medical and Biological Engineering and Computing, 55(11), 1915–1926. https://doi.org/10.1007/s11517-017-1641-y

Jorna, P. G. (1992). Spectral analysis of heart rate and psychological state: a review of its validity as a workload index. Biological Psychology, 34(2–3), 237–257. https://doi.org/10.1016/0301-0511(92)90017-o

Keles, H., Cengiz, C., Demiral, I., Ozmen, M., & Omurtag, A. (2021). 155 A Machine Learning-Based Approach for Predicting Surgeons' Subjective Experience and Skill Levels: Neuroimaging Study. British Journal of Surgery, 108. https://doi.org/10.1093/bjs/znab259.595

Keleş, H. O. (2022). Cerrahların Zihinsel İş Yükü Değerlendirmesinde İşlevsel Beyin Görüntüleme Uygulamaları.

Keles, H. O., Barbour, R. L., & Omurtag, A. (2016). Hemodynamic correlates of spontaneous neural activity measured by human whole-head resting state EEG + fNIRS. NeuroImage, 138, 76–87. https://doi.org/10.1016/j.neuroimage.2016.05.058

Keles, H. O., Cengiz, C., Demiral, I., Ozmen, M. M., & Omurtag, A. (2021). High density optical neuroimaging predicts surgeons's subjective experience and skill levels. PLoS ONE, 16(2 February). https://doi.org/10.1371/journal.pone.0247117

Keles, H. O., & Omurtag, A. (2021). Video game experience affects performance, cognitive load, and brain activity in laparoscopic surgery training. Research Square. https://doi.org/10.21203/rs.3.rs-904919/v1

Kramer, A. F., Trejob, L. J., & Humphrey, D. (1995). Assessment of mental workload with task-irrelevant auditory probes. In Biological Psychology (Vol. 40).

Lachaux, J.-P., Rodriguez, E., Martinerie, J., & Varela, F. J. (1999). Measuring Phase Synchrony in Brain Signals. In Hum Brain Mapping (Vol. 8).

Lenneman, J. K., & Backs, R. W. (2007). LNAI 4562 - Diagnosticity of Cardiac Modes of Autonomic Control Elicited by Simulated Driving and Verbal Working Memory Dual-Tasks. In LNAI (Vol. 4562).

Li, L., Wu, M., Liao, Y., Ouyang, L., Du, M., Lei, D., Chen, L., Yao, L., Huang, X., & Gong, Q. (2014). Grey matter reduction associated with posttraumatic stress disorder and traumatic stress. In Neuroscience and Biobehavioral Reviews (Vol. 43, pp. 163–172). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2014.04.003

Lim, C., Barragan, J. A., Farrow, J. M., Wachs, J. P., Sundaram, C. P., & Yu, D. (2023a).

Physiological Metrics of Surgical Difficulty and Multi-Task Requirement during Robotic Surgery Skills. Sensors, 23(9), 4354. https://doi.org/10.3390/s23094354

Loeches De La Fuente, H., Berthelon, C., Fort, A., Etienne, V., De Weser, M., Ambeck, J., & Jallais, C. (2019). Electrophysiological and performance variations following driving events involving an increase in mental workload. European Transport Research Review, 11(1). https://doi.org/10.1186/s12544-019-0379-z

Luque-Casado, A., Perales, J. C., Cárdenas, D., & Sanabria, D. (2016). Heart rate variability and cognitive processing: The autonomic response to task demands. Biological Psychology, 113, 83–90. https://doi.org/10.1016/j.biopsycho.2015.11.013

Mentis, H. M., Chellali, A., Manser, K., Cao, C. G. L., & Schwaitzberg, S. D. (2016). A systematic review of the effect of distraction on surgeon performance: directions for operating room policy and surgical training. Surgical Endoscopy, 30(5), 1713–1724. https://doi.org/10.1007/s00464-015-4443-z

Moorthy, K., Munz, Y., Dosis, A., Bann, S., & Darzi, A. (2003). The effect of stress-inducing conditions on the performance of a laparoscopic task. Surgical Endoscopy, 17(9), 1481–1484. https://doi.org/10.1007/s00464-002-9224-9

Noyes, J. M., & Bruneau, D. P. J. (2007). A self-analysis of the NASA-TLX workload measure. Ergonomics, 50(4), 514–519. https://doi.org/10.1080/00140130701235232

Maimon, N. B., Bez, M., Drobot, D., Molcho, L., Intrator, N., Kakiashvilli, E., & Bickel, A. (2022). Continuous Monitoring of Mental Load During Virtual Simulator Training for Laparoscopic Surgery Reflects Laparoscopic Dexterity: A Comparative Study Using a Novel Wireless Device. Frontiers in Neuroscience, 15. https://doi.org/10.3389/fnins.2021.694010

Marquart, G., Cabrall, C., & de Winter, J. (2015). Review of Eye-related Measures of Drivers' Mental Workload. Procedia Manufacturing, 3, 2854–2861. https://doi.org/10.1016/j.promfg.2015.07.783

Mazzocco, K., Petitti, D. B., Fong, K. T., Bonacum, D., Brookey, J., Graham, S., Lasky, R. E., Sexton, J. B., & Thomas, E. J. (2009). Surgical team behaviors and patient outcomes. American Journal of Surgery, 197(5), 678–685. https://doi.org/10.1016/j.amjsurg.2008.03.002

Miller, G. A. (1956). The Psychological Review The Magical Number Seven, Plus Or Minus Two: Some Limits On Our Capacity For Processing Information 1.

Miller, S. (2001). Workload measures "National Advanced Driving Simulator." http://www.nads-sc.uiowa.edu/

Modi, H. N., Singh, H., Orihuela-Espina, F., Athanasiou, T., Fiorentino, F., Yang, G. Z., Darzi, A., & Leff, D. R. (2018). Temporal Stress in the Operating Room: Brain Engagement Promotes "coping" and Disengagement Prompts "choking." Annals of Surgery, 267(4), 683–691.https://doi.org/10.1097/SLA.0000000000002289

Modi, H. N., Singh, H., Yang, G. Z., Darzi, A., & Leff, D. R. (2017). A decade of imaging surgeons' brain function (part I): Terminology, techniques, and clinical translation. In Surgery (United States) (Vol. 162, Issue 5, pp. 1121–1130).

Mosby Inc. https://doi.org/10.1016/j.surg.2017.05.021

Moorthy, K., Munz, Y., Dosis, A., Bann, S., & Darzi, A. (2003). The effect of stress-inducing conditions on the performance of a laparoscopic task. Surgical Endoscopy and Other Interventional Techniques, 17(9), 1481–1484. https://doi.org/10.1007/s00464-002-92249

Mukherjee, S., Yadav, R., Yung, I., Zajdel, D. P., & Oken, B. S. (2011). Sensitivity to mental effort and test-retest reliability of heart rate variability measures in healthy seniors. Clinical Neurophysiology, 122(10), 2059–2066. https://doi.org/10.1016/j.clinph.2011.02.032

Mülayimoğlu, A., & Ayaz, S. (2012). Halk Eğitim Merkezine Devam Eden Bireylerin HastaGüvenliğine İlişkin Görüşleri ve Tıbbi Hatalarla İlgili Deneyimleri.

Özsever, B., & Tavacioğlu, L. (2021). Measuring Mental Workload and Heart Rate Variabilityof Officers During Different Navigation Conditions. Marine Science and Technology Bulletin, In Press, 306–312. https://doi.org/10.33714/masteb.976409

Park, W., Jamil, M. H., & Eid, M. (2019). Prefrontal and Interhemispheric Functional EEG Connectivity Associated with Tactile Stimulation. https://hal.science/hal-02013505

Payne, R. L., & Rick, J. T. (1986). Heart rate as an indicator of stress in suegeons and anaesthetists. Pergamon Journals.

Peterson, L. R., & Peterson, M. J. (1959). Short-Term Retention Of Individual Verbal Items 1. In Journal of Experimental Psychology (Vol. 58, Issue 3).

Pfurtscheller, G., & Klimesch, W. (1990). Topographical display and interpretation of event-related desynchronization during a visual-verbal task. Brain Topography, 3(1), 85–93. https://doi.org/10.1007/BF01128865

Pfurtscheller, G., Stancák, A., & Neuper, Ch. (1996). Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: A review. International Journal of Psychophysiology, 24(1–2), 39–46. https://doi.org/10.1016/S0167-8760(96)00066-9

Porte, M. C., Xeroulis, G., Reznick, R. K., & Dubrowski, A. (2007). Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learningnew surgical skills. American Journal of Surgery, 193(1), 105–110. https://doi.org/10.1016/j.amjsurg.2006.03.016

Rieger, A., Stoll, R., Kreuzfeld, S., Behrens, K., & Weippert, M. (2014). Heart rate and heart rate variability as indirect markers of surgeons' intraoperative stress. International Archives ofOccupational and Environmental Health, 87(2), 165–174. https://doi.org/10.1007/s00420-013-0847-z

Roscoe, A. H. (1992). Assessing pilot workload. Why measure heart rate, HRV and respiration? Biological Psychology, 34(2–3), 259–287. https://doi.org/10.1016/0301-0511(92)90018-p

Rosser, J. C., Rosser, L. E., & Savalgi, R. S. (1997). Skill Acquisition and Assessment for Laparoscopic Surgery Objectives: To describe a training method with objective evaluation to enhance laparoscopic surgical skills, to provide training in

laparoscopic suturing techniques, and to assess whether specific training exercises were help-ful in the attainment of intracorporeal suturing skills.

Rutledge, T., Stucky, E., Dollarhide, A., Shively, M., Jain, S., Wolfson, T., Weinger, M. B., & Dresselhaus, T. (2009). A real-time assessment of work stress in physicians and nurses.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention (Vol. 84).

Shao, S., Wang, T., Wang, Y., Su, Y., Song, C., & Yao, C. (2020). Research of hrv as a measure of mental workload in human and dual-arm robot interaction. Electronics (Switzerland), 9(12), 1–17. https://doi.org/10.3390/electronics9122174

Singh, N., Aggarwal, Y., & Sinha, R. K. (2019). Heart rate variability analysis under varied task difficulties in mental arithmetic performance. Health and Technology, 9(3), 343–353. https://doi.org/10.1007/s12553-018-0272-0

Siu, K.-C., Suh, I. H., Mukherjee, M., Oleynikov, D., & Stergiou, N. (2010). The impact of environmental noise on robot-assisted laparoscopic surgical performance. Surgery, 147(1), 107–113. https://doi.org/10.1016/j.surg.2009.08.010

Sleimen-Malkoun, R., Temprado, J. J., Thefenne, L., & Berton, E. (2011). Bimanual training in stroke: How do coupling and symmetry-breaking matter? BMC Neurology, 11. https://doi.org/10.1186/1471-2377-11-11

Song, M. H., Tokuda, Y., Nakayama, T., Sato, M., & Hattori, K. (2009). Intraoperative heart rate variability of a cardiac surgeon himself in coronary artery bypass grafting surgery. Interactive Cardiovascular and Thoracic Surgery, 8(6), 639–641.https://doi.org/10.1510/icvts.2008.195941

Spitzer, B., & Haegens, S. (2017). Beyond the Status Quo: A Role for Beta Oscillations in Endogenous Content (Re)Activation. Eneuro, 4(4), ENEURO.0170-17.2017. https://doi.org/10.1523/ENEURO.0170-17.2017

Stefanidis, D., Anton, N. E., McRary, G., Howley, L. D., Pimentel, M., Davis, C., Yurco, A. M., Sevdalis, N., & Brown, C. (2017). Implementation results of a novel comprehensive mental skills curriculum during simulator training. American Journal of Surgery, 213(2), 353–361. https://doi.org/10.1016/j.amjsurg.2016.06.027

Sterman, M., & Mann, C. (1995). Concepts and applications of EEG analysis in aviation performance evaluation. In Biological Psychology (Vol. 40).

Stylopoulos, N., Cotin, S., Maithel, S. K., Ottensmeyer, M., Jackson, P. G., Bardsley, R. S., Neumann, P. F., Rattner, D. W., & Dawson, S. L. (2004). Computer-enhanced laparoscopic training system (CELTS): Bridging the gap. Surgical Endoscopy and Other Interventional Techniques, 18(5), 782–789. https://doi.org/10.1007/s00464-003-8932-0

Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies. In Medical Education (Vol. 44, Issue 1, pp.85–93). https://doi.org/10.1111/j.1365-2923.2009.03498.x

Veltman, J. A., & Gaillard, A. W. K. (1998). Physiological workload reactions to

increasing levels of task difficulty. Ergonomics, 41(5), 656–669. https://doi.org/10.1080/001401398186829

Weenk, M., Alken, A. P. B., Engelen, L. J. L. P. G., Bredie, S. J. H., van de Belt, T. H., & van Goor, H. (2018). Stress measurement in surgeons and residents using a smart patch. American Journal of Surgery, 216(2), 361–368. https://doi.org/10.1016/j.amjsurg.2017.05.015

Wickens, C. D. (2014). Effort in human factors performance and decision making. Human Factors, 56(8), 1329–1336. https://doi.org/10.1177/0018720814558419

World Health Organization. (2004). 10 facts on patient safety. http://www.who.int

Yurko, Y. Y., Scerbo, M. W., Prabhu, A. S., Acker, C. E., & Stefanidis, D. (2010). Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. Simulation in Healthcare, 5(5), 267–271. https://doi.org/10.1097/SIH.0b013e3181e3f329

Zakeri, Z., Mansfield, N., Sunderland, C., & Omurtag, A. (2020). Physiological correlates of cognitive load in laparoscopic surgery. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-69553-3

Zander, T. O., Shetty, K., Lorenz, R., Leff, D. R., Krol, L. R., Darzi, A. W., Gramann, K., & Yang, G. Z. (2017). Automated Task Load Detection with Electroencephalography: Towards Passive Brain-Computer Interfacing in Robotic Surgery. Journal of Medical Robotics Research, 2(1). https://doi.org/10.1142/S2424905X17500039