SURGICAL MODALITY AND TASK COMPLEXITY: AN fNIRS BASED INVESTIGATION OF COGNITIVE LOAD IN MINIMALLY INVASIVE TECHNIQUES

by

Fuat Ücrak

B.S., in Bioengineering, Yıldız Technical University, 2013

M.S., in Electronics & Communication Engineering, Yıldız Technical University, 2017

Submitted to the Institute of Biomedical Engineering

in partial fulfillment of the requirements

for the degree of

Doctor

of

Philosophy

Boğaziçi University

2025

SURGICAL MODALITY AND TASK COMPLEXITY: AN fNIRS BASED INVESTIGATION OF COGNITIVE LOAD IN MINIMALLY INVASIVE TECHNIQUES

APPROVED BY:

Prof. Dr. Cengizhan Öztürk .....
(Thesis Advisor)

Assoc. Prof. Dr. Mehmet Emin Aksoy .....
(Thesis Co-advisor)

Prof. Dr. Ahmet Ademoğlu .....

Prof. Dr. Muzaffer Olcay Çizmeli .....

Assoc. Prof. Dr. Daniela Schulz .....

Assoc. Prof. Dr. Kurtuluş İzzetoğlu .....

DATE OF APPROVAL: 27 May 2025

ACKNOWLEDGMENTS

First and foremost, I extend my heartfelt gratitude to my family. I am especially thankful to my mother, Zozan Ücrak, for her unwavering love and strength, and to my father, Ferman Ücrak, for his constant encouragement and belief in my academic journey. I also appreciate my brothers and sisters for their support and practical advice during this process.

I sincerely thank my thesis advisor, Prof. Dr. Cengizhan Öztürk, for his guidance, thoughtful feedback, and patient mentorship. His support was essential in shaping and completing this research.

I am grateful to my co-advisor, Assoc. Prof. Dr. Mehmet Emin Aksoy, for his clarity in helping define the research focus and for his steady support and insightful suggestions throughout the study.

I also thank Assoc. Prof. Dr. Kurtuluş İzzetoğlu for his expertise in data analysis methods and for allowing me to work in his lab. His contributions improved the technical quality of this work.

I am thankful to Prof. Dr. Muzaffer Olcay Çizmeli, who encouraged me to begin this PhD and provided strong support at the start of this academic path.

I gratefully acknowledge Burhanettin Ataş, Mert Deniz Polat, Ümit Gür, Turan Şahin, Serhat Ilgaz Yöner, Neslihan Gökmen, and Ayça Yılmaz for their support and contributions during the experimental phases.

Finally, I thank the CASE (Center of Advanced Simulation and Education) staff at Acıbadem Mehmet Ali Aydınlar University for their help and for providing access to essential facilities that enabled the practical components of this research.

ACADEMIC ETHICS AND INTEGRITY STATEMENT

I, Fuat Ücrak, hereby certify that I am aware of the Academic Ethics and Integrity Policy issued by the Council of Higher Education (YÖK) and I fully acknowledge all the consequences due to its violation by plagiarism or any other way.

Name :

Signature:

Date:

ABSTRACT

SURGICAL MODALITY AND TASK COMPLEXITY: AN fNIRS BASED INVESTIGATION OF COGNITIVE LOAD IN MINIMALLY INVASIVE TECHNIQUES

Minimally invasive surgical techniques, such as laparoscopic and robotic surgery, have advanced practice by improving precision, reducing recovery times, and minimizing complications. However, these approaches impose different cognitive demands and require distinct skills. This study uses functional near-infrared spectroscopy (fNIRS) to assess cognitive workload differences across surgical modalities and task complexities. We examined how surgical modality (laparoscopy vs. robotic surgery) and task complexity (pick and place vs. knot tying) influence cognitive workload, and compared demands in simulation based and real world environments. Twenty six trainees and specialists in general and gynecologic surgery participated. They performed standardized laparoscopic and robotic tasks of varying complexity while prefrontal cortex activity and task times were recorded using fNIRS. Both simulation based and real settings were included. Laparoscopic surgery induced greater prefrontal cortex activation than robotic surgery, especially during complex tasks like knot tying, indicating higher cognitive workload. Task complexity was a major factor, with more intricate procedures prompting increased neural activation. Real surgical environments led to greater cognitive engagement than simulations. Robotic surgery was linked to lower cognitive load, likely due to ergonomic and motor control advantages. While simulation based training prepares surgeons, it does not fully replicate the cognitive demands of real operations. These findings underscore the importance of cognitive workload assessment in surgical education and support integrating neuroimaging tools like fNIRS into training programs to improve skill acquisition and performance.

Keywords: Laparoscopic Surgery; Robotic Surgery, Cognitive Workload, fNIRS, Surgical Training, Task Complexity, Minimally Invasive Surgery.

ÖZET

CERRAHİ YÖNTEM VE GÖREV ZORLUĞU: MİNİMAL İNVAZİV YAKLAŞIMLARDA fNIRS TEMELLİ BİLİŞSEL YÜK ANALİZİ

Laparoskopik ve robotik cerrahi gibi minimal invaziv teknikler, cerrahi uygulamalarda hassasiyeti artırmakta, iyileşme süresini kısaltmakta ve komplikasyonları azaltmaktadır. Ancak bu yöntemler, farklı bilişsel yükler oluşturur ve özgün beceriler gerektirir. Bu çalışmada, fonksiyonel yakın kızılötesi spektroskopi (fNIRS) kullanılarak cerrahi modaliteler (laparoskopi vs. robotik cerrahi) ve görev zorlukluklarının (Peg Transferi ve Düğüm Atma) bilişsel iş yüküne etkisi ile simülasyon ve gerçek cerrahi ortamlar karşılaştırılmıştır. Genel ve jinekolojik cerrahi alanlarında çalışan ya da eğitim gören 26 katılımcı, farklı zorluklardaki laparoskopik ve robotik görevleri yerine getirmiştir. Görevler sırasında prefrontal korteks aktiviteleri ve görev süreleri fNIRS ile kaydedilmiştir. Hem simülasyon hem de gerçek cerrahi ortamlar değerlendirmeye alınmıştır. Laparoskopik cerrahi, özellikle düğüm atma gibi daha zor görevlerde, robotik cerrahiye kıyasla daha yüksek prefrontal aktivasyon göstermiştir; bu da daha fazla bilişsel iş yüküne işaret etmektedir. Görev zorluğu, artan sinirsel aktivasyonla birlikte önemli bir belirleyici olmuştur. Gerçek cerrahi ortamda yapılan uygulamalar, simülasyonlara göre daha yüksek bilişsel katılım sağlamıştır. Robotik cerrahi ise ergonomik ve motor kontrol avantajları sayesinde daha düşük bilişsel yükle ilişkilendirilmiştir. Simülasyon temelli eğitim faydalı olsa da gerçek cerrahi ortamın bilişsel taleplerini tam olarak yansıtmamaktadır. Bulgular, cerrahi eğitimde bilişsel iş yükü değerlendirmesinin önemini vurgulamakta ve fNIRS gibi nörogörüntüleme tekniklerinin eğitim programlarına entegrasyonunu desteklemektedir.

Anahtar Sözcükler: Laparoskopik Cerrahi, Robotik Cerrahi, Bilişsel İşyükü, fNIRS, Cerrahi Eğitim, Görev Zorluğu, Minimal İnvaziv Cerrahi.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

G	Geometric factor in the Beer–Lambert Law
I_0	Input light intensity in fNIRS measurements
I	Detected light intensity after transmission through tissue
L	Path length of photons through tissue
t	Time
p	Probability value in statistical significance testing
d	Cohen's d, a measure of effect size
μ_a	Absorption coefficient of tissue
μ_s	Scattering coefficient of tissue
λ	Wavelength of light used in fNIRS (typically 730 nm and 850 nm)
ΔOD	Change in optical density
$\Delta CHbO_2$	Change in oxygenated hemoglobin concentration
$\Delta CHbR$	Change in deoxygenated hemoglobin concentration
HbO	Oxygenated hemoglobin concentration
HbR	Deoxygenated hemoglobin concentration
$HbDiff$	Hemoglobin difference ($HbO - HbR$)
$HbTotal$	Total hemoglobin concentration ($HbO + HbR$)
α_{HB}	Molar extinction coefficient for deoxyhemoglobin
α_{HBO_2}	Molar extinction coefficient for oxyhemoglobin

LIST OF ABBREVIATIONS

ANS	Autonomic Nervous System
AMPFC	Anterior Medial Prefrontal Cortex
CWL	Cognitive Workload
DLPFC	Dorsolateral Prefrontal Cortex
fNIRS	Functional near infrared Spectroscopy
FLS	Fundamentals of Laparoscopic Surgery
FRS	Fundamentals of Robotic Surgery
Hb	Hemoglobin
HbO2	Oxygenated hemoglobin
HbR	Deoxygenated hemoglobin
ΔHbO	Change in Oxygenated Hemoglobin
ΔHbR	Change in Deoxygenated Hemoglobin
HbDiff	Hemoglobin Difference (Oxy - Deoxy)
HbTotal	Total Hemoglobin Concentration
KT	Knot tying
LAMPFC	Left Anterior Medial Prefrontal Cortex
LDLPFC	Left Dorsolateral Prefrontal Cortex
PFC	Prefrontal Cortex
PP	Pick and place
RAMPFC	Right Anterior Medial Prefrontal Cortex
RDLPFC	Right Dorsolateral Prefrontal Cortex
ROI	Region of interest
RNI	Relative neural involvement
RNE	Relative neural efficiency

1. INTRODUCTION

Minimally invasive surgical techniques such as laparoscopy and robotic surgery have transformed patient care by offering smaller incisions and greater operative precision. Over the past three decades, laparoscopic surgery has rapidly advanced, with significant improvements in instrumentation, visualization, and surgical techniques, establishing it as a standard of care across numerous surgical specialities. Despite its more recent emergence, robotic surgery has proven equally transformative, revolutionizing the field by providing surgeons with enhanced dexterity, precision, and three dimensional visualization [1]. While robotic platforms reduce tactile feedback compared to laparoscopy, they compensate through enhanced 3D visualization and intuitive controls. Studies have shown that improved depth perception reduces reliance on spatial memory and facilitates faster decision making, thereby affecting cognitive resource allocation [2, 3].

1.1 The Historical Development of Laparoscopic and Robotic Surgery

Laparoscopic surgery, a minimally invasive surgical technique that involves the use of small incisions and specialized instruments, has undergone significant advancements since its inception. The pioneering work of Kurt Semm in the 1970s marked a turning point in the development of laparoscopy, as he conducted the first laparoscopic appendectomy and cholecystectomy [4]. Building upon these early successes, subsequent decades witnessed rapid progress in instrumentation, visualization techniques, and surgical procedures, solidifying laparoscopy's position as a standard of care in many surgical specialities [1].

The evolution of robotic surgery, while more recent, has been equally transformative. The introduction of the Da Vinci Surgical System in the early 2000s marked a groundbreaking advancement in surgery, providing enhanced dexterity, improved precision, and three dimensional visualization for surgeons [5]. Robotic assisted surgery has

gained widespread adoption in various surgical specialities, particularly urology, gynecology, and general surgery [2].

1.2 Advantages of Laparoscopic and Robotic Surgery

Laparoscopic and robotic surgery offer numerous advantages for both surgeons and patients. For surgeons, these techniques provide enhanced visualization, improved dexterity, and reduced hand tremor, leading to more precise and controlled surgical maneuvers [1,2]. Additionally, laparoscopic and robotic surgery often result in shorter hospital stays, faster recovery times, and less postoperative pain compared to traditional open surgery [1,4].

Previous research using neuroimaging approaches demonstrated that robotic surgery not only enhances technical performance under time pressure compared to conventional laparoscopy but also results in greater prefrontal cortex activation, indicative of improved attentional control and task engagement during demanding conditions [6].

From a patient's perspective, the benefits of laparoscopic and robotic surgery are substantial. Smaller incisions translate to reduced scarring, decreased blood loss, and a lower risk of complications. These advantages often lead to improved patient outcomes, enhanced quality of life, and a faster return to normal activities [2,5,7].

While both laparoscopic and robotic surgery training are designed to develop minimally invasive surgical skills, they differ markedly in technology, learning curves, and training methodologies. Laparoscopic training is more widely accessible and typically serves as the foundational step before progressing to robotic surgery. In contrast, robotic surgery training emphasizes mastery of advanced instrumentation and console based operation [1,7–9]. Although both approaches aim to achieve proficiency in minimally invasive techniques, they require distinct skill sets and training strategies, along with dedicated practice and continuous skill maintenance [2,5,7,10,11].

1.3 Educational Impact of Simulation in Laparoscopic and Robotic Skill Development

Over the last three decades, simulation has become a cornerstone of medical education, providing a risk free environment for acquiring technical skills, promoting patient safety, and mitigating medical errors, thus enhancing physician training and improving healthcare outcomes [12, 13]. Patient safety should be the paramount objective of healthcare. The prevention of medical errors is a critical strategy to achieve this goal. Simulation based training offers a promising educational approach to both promote patient safety and mitigate the occurrence of medical errors. Widely employed in fields such as aviation, military, industry, and medicine, simulation has become an integral component of medical education. The benefits of medical simulation are well established, including the minimization of ethical concerns, enhancement of the educational experience, creation of a learner centered and teacher enabled environment, provision of a patient risk free learning space, facilitation of the acquisition and practice of new techniques, and the enabling of performance assessment [14–16].

Simulation offers a standardized and safe approach to training and assessing surgeons. The emergence of laparoscopic techniques has coincided with a significant increase in the use of simulation for surgical training. Research has consistently demonstrated that skills acquired through simulation are transferable to real world clinical settings [12]. Although simulation has gained widespread acceptance as an educational training tool, with ample evidence supporting its use in health education, the effectiveness of simulation based assessments in evaluating competence and performance remains a subject of ongoing debate [14]. Given the heightened cognitive demands imposed by modern technology, recent research trends underscore the importance of quantitatively assessing mental workload to enhance performance, productivity, and safety across various working environments. Physiological indicators of mental workload offer accurate and practical indices of an operator's performance [17].

To maintain patient safety during complex laparoscopic and robotic procedures, surgeons must undergo rigorous training with advanced simulation tools that closely repli-

cate real world surgical scenarios and challenges. Essential training programs like the Fundamentals of Laparoscopic Surgery (FLS) and the Fundamentals of Robotic Surgery (FRS) play a crucial role in equipping surgeons with the skills needed for these advanced techniques [18,19]. The Fundamentals of Laparoscopic Surgery (FLS) program is a structured, comprehensive curriculum designed to teach and assess the core skills required for laparoscopic surgery [6]. It consists of online modules that focus on cognitive knowledge, hands on skills training using simulators, and a standardized assessment to evaluate proficiency. Widely adopted in surgical training, the FLS program plays a critical role in preparing surgeons, assessing their competency, and enhancing patient safety by ensuring they are equipped with the necessary skills to perform laparoscopic procedures effectively [18,20].

The Fundamentals of Robotic Surgery (FRS) protocol is an emerging standard designed to train surgeons in robotic assisted procedures. Building on the principles of the Fundamentals of Laparoscopic Surgery (FLS) program, FRS emphasizes simulation based training, often utilizing advanced virtual reality platforms to teach essential robotic skills such as console operation, instrument manipulation, and 3D navigation[5]. This structured approach ensures that surgeons develop the technical proficiency and spatial awareness required for robotic surgery, while also addressing the unique challenges of working with wristed instruments and mastering the robotic console [19].

Recent findings also suggest that robotic systems can accelerate early skill acquisition. In a controlled simulated study, Leijte et al. (2020) found that novices learning robotic suturing achieved significantly faster completion times compared to those performing laparoscopic suturing [21]. However, initial suturing quality, measured by knot integrity, was higher in the laparoscopic group. These results support the premise that while robotic systems may enhance procedural efficiency and reduce cognitive burden during early learning, laparoscopic training may promote more conservative, quality focused technique in initial stages [21].

1.4 Assessing Mental Workload in Complex Tasks: A Multidimensional Perspective

As defined in the literature, mental workload is the interplay between the cognitive demands imposed on operators by their tasks and the effort exerted to accomplish these tasks [22, 23]. Consequently, it is essential to evaluate mental workload independently of performance measures.

Comprehending how the brain allocates mental resources in response to task demands is crucial for complex and high risk operational settings such as aviation, air traffic control, and surgery. Excessive mental workload in challenging tasks can lead to performance failures with potentially devastating consequences. Assessing mental workload is complex, as it involves the interplay of environmental demands, individual characteristics, and task performance. Consequently, solely considering task characteristics is insufficient for inferring an individual's level of mental workload [24]. High immersion VR environments with realistic distractors have been shown to significantly alter cognitive workload, as evidenced by increased HbO levels in the prefrontal cortex [25]. These findings emphasize the role of environmental fidelity in simulating real life task complexity.

Mental workload, a well established concept rooted in the multiple resources model of human factors, increases substantially when users must cognitively process large volumes of information within a single modality (spatial or verbal) at the same stage of cognitive processing. For instance, individuals may struggle to retain two number sequences simultaneously but can effectively scan text for a specific keyword while rehearsing a single number sequence or processing spatial information. Mental workload can be characterized as the relationship between primary task performance and the cognitive resources demanded by the task. If a user's cognitive engagement is insufficient or excessive, task performance may decline. Similar concepts are encompassed within the cognitive load literature, and the terms "mental workload" and "cognitive load" are often used interchangeably in scholarly publications [26].

It is well established that there are three primary categories of mental workload

measures: self report measures, physiological (and neurophysiological) measures, and primary task performance measures. Each class of mental workload measure possesses distinct advantages and disadvantages [27].

Self report measures, while widely used due to their ease of administration and low computational cost, are primarily administered post task and lack real time granularity. Task performance measures, though direct indicators of human performance, offer limited diagnostic value when considered in isolation. Physiological measures, though sensitive and continuous, are often more expensive and require extensive preprocessing, limiting their real time applicability [27]. However, subjective self report measures often fail to accurately reflect dynamic cognitive states during high demand clinical scenarios [28]. Psychophysiological tools like fNIRS offer real time, objective tracking, especially in ecologically valid settings such as simulated surgery or real world driving environments, where behavioral cues may be insufficient or misleading.

Several physiological changes have been observed to correlate with mental workload, including pupil dilation, skin temperature changes, galvanic skin response, and fluctuations in cardiac activity. These measures often provide valuable insights into mental workload dynamics. Additionally, direct measurements of brain activity offer a more comprehensive approach to estimating mental workload [17, 24, 26, 27].

Recent studies have demonstrated that integrating multiple physiological signals, such as fNIRS, EEG, and eye tracking data, using multimodal deep learning models, can enhance the accuracy of cognitive workload detection during surgical tasks [29].

A recent systematic review underscores the need for objective and standardized approaches to assess CWL, highlighting that subjective methods suffer from granularity limitations and recall bias, especially in dynamic clinical settings [30]. The authors advocate the integration of physiological and neurophysiological sensors to capture real time changes in CWL, emphasizing their potential to minimize human error and enhance patient safety [30].

Neural and physiological measures present a promising method for assessing mental workload because of their direct connection to brain function, allowing evaluation before any decline in performance occurs. These measures can be linked to the mental demands of a task, offering continuous, unobtrusive monitoring of the operator without disrupting their activities, unlike secondary task performance or subjective workload assessments. [31–33].

1.5 Neurophysiological Methods for Evaluating Mental Workload

While simulation improves technical proficiency, it does not fully capture the mental workload experienced during actual procedures. Neuroimaging tools like fNIRS offer an objective, physiological approach to assess this cognitive dimension of surgical performance. Beyond traditional performance metrics, brain based measures obtained through functional neuroimaging techniques can provide deeper insights into traineeś performance [34]. Various neuroimaging techniques have been utilized to investigate the neural mechanisms underlying cognitive processes. Methods such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG), positron emission tomography (PET), and functional near infrared spectroscopy (fNIRS) provide alternative approaches for assessing cognitive workload and performance. Table 1.1 summarizes the key advantages and limitations of these neurocognitive imaging techniques [34].

Table 1.2: Comparison of neuroimaging techniques. fMRI, fNIRS, and PET are hemodynamic imaging modalities, while EEG relies on electrical activity, and MEG is based on electromagnetic signals. PET: Positron Emission Tomography; EEG: Electroencephalography; MEG: Magnetoencephalography [34].

As Chakladar and Roy (2024) categorize, physiological measures for cognitive workload can be divided into brain activity based (e.g., EEG, fNIRS, fMRI) and non brain activity based (e.g., heart rate, respiration, eye tracking) methods. Each class

Table 1.1
Comparison of Neuroimaging Techniques. This table summarizes the advantages and disadvantages of commonly used neuroimaging modalities in cognitive workload studies.

Technique	Advantages	Disadvantages
fMRI	High spatial resolution (millimeter level); whole brain coverage; provides structural and functional data; good source localization	Low temporal resolution (seconds); susceptible to motion artifacts; requires body immobilization; contraindications (e.g., pacemakers); high cost; MRI compatibility issues
fNIRS	High temporal resolution (milliseconds); allows natural body movement; low sensitivity to motion artifacts; portable and cost effective	Low spatial resolution (centimeters); limited to cortical activity; susceptible to extracerebral hemodynamic influences; affected by hair and skull properties
PET	High spatial resolution (millimeter level); whole brain coverage; provides metabolic data	Low temporal resolution (seconds); requires injection of radioactive tracer; high cost
EEG	Very high temporal resolution (milliseconds); portable and affordable	Low spatial resolution; susceptible to environmental noise; poor source localization; time consuming setup
MEG	High temporal and spatial resolution; good source localization	Sensitive to environmental noise; limited portability; high cost; contraindications (e.g., metal implants like dental crowns)

has distinct advantages: brain based tools provide high specificity and temporal spatial resolution, while non brain measures offer low cost, accessible alternatives albeit with increased susceptibility to noise [35].

Given its noninvasive nature, portability, and suitability for real world applications, fNIRS is particularly advantageous for measuring prefrontal cortex activity in field settings. As a result, fNIRS has been selected as the primary neuroimaging method for this study. Numerous studies have successfully integrated fNIRS based brain imaging with established performance assessment frameworks in simulation based medical training, demonstrating its effectiveness in evaluating cognitive workload [17, 22, 24, 34].

Compared to EEG and fMRI, fNIRS offers a balance between spatial resolution and ecological validity, making it well suited for real time, movement allowed settings like surgical training. Unlike fMRI, which restricts motion and requires high infrastructure, fNIRS enables bedside or intraoperative use, with sufficient sensitivity to detect workload changes in the prefrontal cortex [36]. fNIRS has been increasingly used to investigate cortical dynamics within contextual interference paradigms, offering valuable insights into cognitive load and learning trajectories across blocked and random practice schedules [37]. Recent reviews highlight that fNIRS has transitioned from a developmental stage into a mature methodology, particularly suited for ecological and applied neuroscience contexts. Its flexibility in mobile and real world settings opens new avenues for studying cognitive workload in operational domains like surgery, aviation, and driving [38].

The integration of neural indices alongside behavioral metrics is increasingly recognized as essential for comprehensive cognitive workload assessment in surgical training. fNIRS has been shown to capture workload differences that are not always reflected in task performance metrics alone. Aksoy et al. (2025) demonstrated that even when Robotic assisted surgery and laparoscopic tasks yield similar skill acquisition rates, significant differences in prefrontal cortex activation are detectable only through neurophysiological monitoring [39].

Recent studies have further highlighted the advantages of fNIRS over traditional neuroimaging modalities, especially in applied cognitive workload research. For example, in a dual task drone piloting study, fNIRS was successfully employed to monitor real time prefrontal activity under increasing task demands, revealing its robustness and portability in dynamic task environments [40].

1.6 fNIRS

Over the past decade, functional near-infrared spectroscopy (fNIRS) has emerged as an innovative neuroimaging technique for conducting functional brain imaging research [41]. fNIRS is an emerging neuroimaging technology that serves as an alternative to EEG

and fMRI. With its lightweight design, affordability, noninvasive nature, ease of setup, and resilience to movement artifacts, fNIRS presents a promising approach for brain imaging studies, particularly in dynamic and real world environments [17, 22, 33, 34, 41].

fNIRS is a neuroimaging technique that leverages the optical characteristics of biological tissues and hemoglobin chromophores to monitor brain activity. Utilizing wavelengths between 700 and 900 nm, where neural tissues are largely transparent. fNIRS primarily detects the absorption of light by oxygenated (HbO2) and deoxygenated hemoglobin (Hb). By applying the modified Beer Lambert Law, fNIRS quantifies changes in hemoglobin concentrations within cortical tissue, providing a direct measure of brain activity based on hemodynamic responses [22, 34, 42].

1.6.1 fNIRS Working Principles

An optical apparatus typically consists of a light source that emits near infrared light into biological tissue and a detector that captures the light after its interaction with the medium. Within the tissue, photons undergo two primary interactions: absorption, where energy is lost to the medium, and scattering. Most biological tissues exhibit relative transparency within the near infrared range (700-900 nm), commonly referred to as the "optical window." This transparency results from the low absorbance of key tissue constituents, such as water, oxyhemoglobin, and deoxyhemoglobin, allowing deeper light penetration (Figure 1.1) [43]. Among the primary absorbers, oxy and deoxyhemoglobin play a crucial role in tissue oxygenation and metabolism. Notably, within the optical window, their absorption spectra remain sufficiently distinct, enabling spectroscopic differentiation using a limited set of specific wavelengths [41].

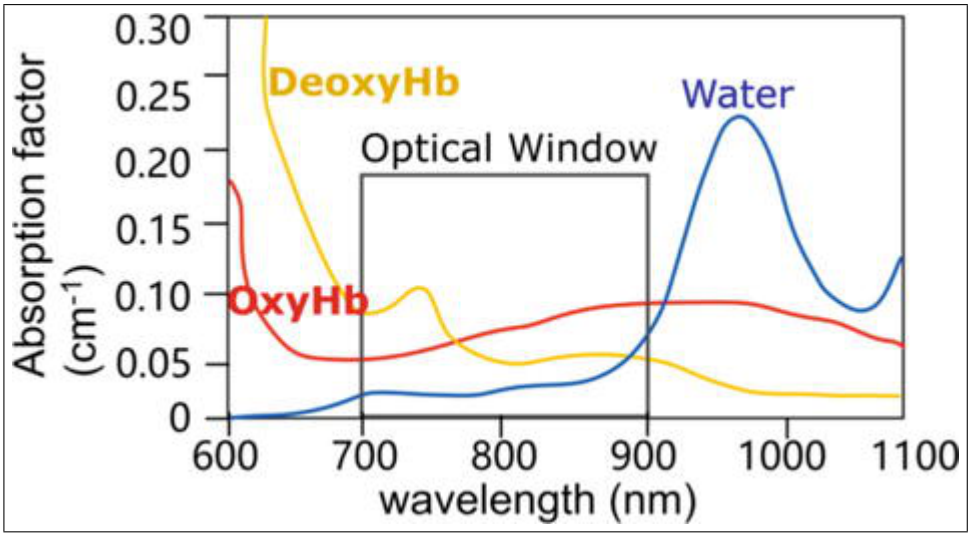

Figure 1.1 Absorption spectrum in NIR window [43].

fNIRS technology operates using specific wavelengths of light within the optical window. When introduced into the human head, photons interact with biological tissues through two primary mechanisms: scattering at extracellular and intracellular boundaries across various head layers (skin, skull, cerebrospinal fluid, brain, etc.) and absorption primarily by oxy and deoxyhemoglobin [32, 34, 41]. The Figure 1.2 illustrates the setup of a functional near-infrared spectroscopy (fNIRS) system on the human scalp, showing the arrangement of emitters and detectors on a headcap. Light emitted from the sources penetrates the scalp and skull, following a characteristic "banana-shaped" path through the cortical tissue before being detected. Two types of channels are shown: a standard channel (approximately 3 cm) and a short-separation channel (approximately 0.8 cm), the latter designed to capture superficial signals for improved signal quality. The layered anatomy includes the scalp, skull, and brain tissue [6] .

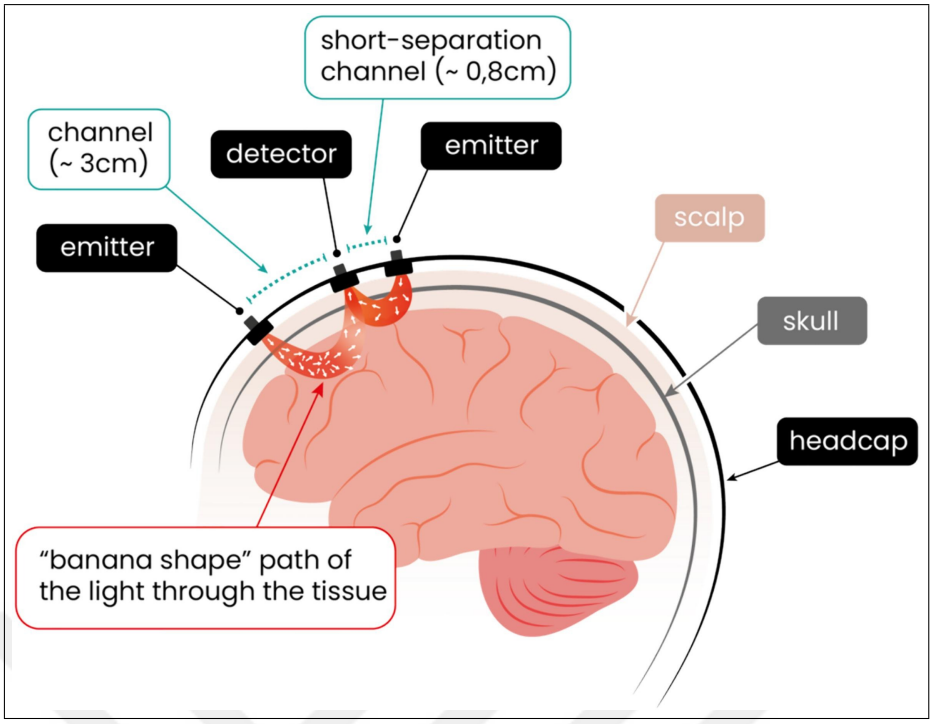

Figure 1.2 Simplified functional near-infrared spectroscopy (fNIRS) montage [6].

In functional optical brain imaging studies, light attenuation caused by scattering is typically assumed to remain constant, since the number of scatterers within various head layers does not vary with cognitive activity. Therefore, any observed changes in attenuation during cognitive tasks are attributed to variations in absorption, driven by fluctuations in oxy and deoxyhemoglobin concentrations within brain tissue. As highlighted by Pinti et al. (2020), fNIRS enables the monitoring of cortical hemodynamics by emitting near infrared light into biological tissue and measuring absorption changes associated with oxygenated and deoxygenated hemoglobin. The detected optical signal is influenced not only by cerebral activity but also by extracerebral hemodynamics such as scalp blood flow [38]. Therefore, recent advancements in signal processing, like short separation channel regression, are crucial for isolating cortical signals from superficial artifacts. These methods enhance the reliability of fNIRS in complex, real world environments where motion and physiological noise are prevalent. The integration of such techniques is increasingly recognized as essential in neuroergonomics and cognitive workload assessment domains, including surgical simulation [38].

The neurovascular coupling mechanism underlying fNIRS is critical for interpreting

hemodynamic signals as proxies for neuronal activity, especially during complex cognitive tasks that elicit regional changes in HbO and HbR levels [36]. Increases in local neuronal activity demand greater oxygen delivery, resulting in an overcompensation of regional cerebral blood flow and a detectable rise in HbO and decrease in HbR, which fNIRS captures through optical density changes. This pattern of increased HbO and decreased HbR reflects the underlying neurovascular coupling, whereby regional cerebral blood flow increases disproportionately to local oxygen consumption. As a result, HbO serves as a proxy for task evoked cortical activation, and HbR reductions signal decreased oxygen extraction at the site [44, 45]. This relationship aligns with the principle of neurovascular coupling, which links cerebral hemodynamic changes to functional brain activity. By leveraging this physiological mechanism, fNIRS technology enables the measurement of oxygen related chromophore concentration changes, facilitating functional optical brain imaging [41, 42, 46].

In functional near infrared spectroscopy (fNIRS), changes in oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (HbR) concentrations are widely recognized as indirect markers of localized neural activity due to neurovascular coupling. When a specific brain region is activated during a cognitive or motor task, there is an increase in regional cerebral blood flow (rCBF) to meet the heightened metabolic demand. This typically results in an elevation of HbO and a concurrent decrease in HbR, reflecting an oversupply of oxygenated blood relative to consumption [38, 47]. An increase in HbO is thus commonly interpreted as a sign of cortical activation, while a decrease in HbO may suggest reduced engagement or task disengagement. Conversely, elevated HbR may indicate a mismatch between oxygen delivery and demand, potentially signifying cognitive overload or fatigue, whereas a reduction in HbR is generally associated with efficient metabolic response and cognitive efficiency [45]. These patterns of HbO and HbR dynamics are critical for interpreting workload, attention, and performance in fNIRS based studies of cognitive function.

fNIRS systems are categorized into continuous wave, frequency domain, and time domain modalities. Most cognitive neuroscience studies use continuous wave systems, which measure relative changes in light intensity. In contrast, frequency domain and time

domain systems offer absolute concentration measurements and deeper spatial sensitivity but require more complex hardware [44].

According to the modified Beer-Lambert Law, the intensity of light after it has undergone absorption and scattering by biological tissue is represented as:

$$I = GI_0 e^{-(\alpha_{HB} C_{HB} + \alpha_{HBO_2} C_{HBO_2})L} \tag{1.1}$$

where:

- G accounts for the measurement geometry and is assumed to remain constant during changes in chromophore concentration.
- I_0 is the input light intensity.
- α_{HB} and α_{HBO_2} are the molar extinction coefficients for deoxyhemoglobin and oxyhemoglobin respectively.
- C_{HB} and C_{HBO_2} are the concentrations of deoxyhemoglobin and oxyhemoglobin respectively.
- L is the photon path length which is a function of absorption and scattering coefficients (μ_a and μ_s)

By measuring optical density (OD) changes at two wavelengths, the relative change of oxy and deoxyhemoglobin versus time can be obtained. If the intensity measurement at the initial time (baseline) is I_b, and at another time is I, the OD change due to variation in C_{HB} and C_{HBO_2} during that period is:

$$\Delta OD = \log\left(\frac{I_b}{I}\right) = \alpha_{HB}\Delta C_{HB} + \alpha_{HBO_2}\Delta C_{HBO_2} \tag{1.2}$$

By performing measurements at two different wavelengths, changes in deoxyhemoglobin ΔC_{HB} and oxyhemoglobin ΔC_{HBO_2} concentrations can be calculated. From these values, oxygenation and blood volume can subsequently be determined:

$$Blood \quad Volume = \Delta C_{HBO_2} + \Delta C_{HB} \tag{1.3}$$

It is important to note that fNIRS primarily captures signals from superficial cortical structures due to the limited penetration depth of near infrared light. Therefore, most studies focus on regions like the prefrontal cortex, while subcortical structures remain inaccessible [44]. Using this technique and its associated measures, researchers have assessed various brain functions, including motor and visual activation, auditory stimulation, and the execution of diverse cognitive tasks. In the studies presented in this article, we specifically utilized oxygenation data to evaluate different cognitive functions [41, 42].

1.7 The Significance of Knot Tying and Pick and Place in Laparoscopic and Robotic Surgery Training

Knot tying and pick and place are fundamental skills in laparoscopic surgery, providing a solid foundation for more complex procedures [48]. These tasks are crucial for developing the fine motor skills, hand eye coordination, and spatial awareness necessary for successful laparoscopic surgery [49]. Pick and place is another vital skill that helps trainees develop dexterity and control within the confined space of the laparoscopic field [50]. This task involves manipulating objects with forceps or other instruments, requiring precision and coordination [50].

The pick and place task is widely recognized as a fundamental assessment of fine motor skills, spatial awareness, and hand eye coordination in minimally invasive surgical training. Its structured design enables objective evaluation across both laparoscopic and robotic platforms, providing a standardized baseline for cognitive and motor performance

comparisons. Recent evidence supports its utility as a sensitive task for detecting neural workload differences between surgical modalities [39]. Pick and place exercises can improve a surgeon's ability to handle delicate tissues, suture needles, and other instruments without causing damage.

Both knot tying and pick and place are essential for developing the fundamental skills required for laparoscopic surgery [51]. They provide a solid foundation for more advanced procedures and help to reduce the risk of complications. As such, these tasks are commonly included in laparoscopic surgery training programs [51].

1.8 The Prefrontal Cortex

The prefrontal cortex (PFC), the anterior region of the frontal lobes, plays a crucial role in a range of higher order cognitive functions, including attentional control, emotional regulation, complex learning, and social cognition, and is implicated in executive functions, behavioral inhibition, and general intelligence [52–55]. The dorsolateral prefrontal cortex (DLPFC) is a key component of the prefrontal cortex, which is the brain region primarily responsible for executive function such as planning, strategy building and executive decisions [56]. The anterior medial prefrontal cortex (AMPFC) is a key region involved in a wide range of higher order cognitive functions, including task management, planning, reasoning, attention, multitasking, task set representations, and decision making [57, 58]. Neuroimaging studies consistently demonstrate that PFC exhibits increased activity during tasks with high cognitive demands, suggesting its crucial role in executive functions such as working memory and decision making [17, 59, 60].

1.9 The Role of fNIRS in Surgical Training and Cognitive Workload Assessment

Several studies have demonstrated the effectiveness of fNIRS in assessing cognitive workload and skill acquisition in surgical training. For instance, Izzetoglu et al.

(2021) found that cortical oxygenation levels decreased in robotic surgery trainees during repeated sessions, indicating improved efficiency [61]. Similarly, Aksoy et al. (2023) showed that experts exhibited lower prefrontal cortex activation compared to novices, highlighting differences in cognitive resource utilization [62]. These findings underscore the potential of neurophysiological biomarkers in evaluating surgical proficiency and informing adaptive training protocols.

In addition, Nemani et al. (2018) explored the use of fNIRS to assess bimanual motor skills in laparoscopic surgery, revealing that experts showed lower prefrontal cortex activation and higher activation in the primary motor cortex (M1) and supplementary motor area (SMA) [63]. This suggests that expert surgeons achieve a more efficient distribution of neural resources, reinforcing the role of fNIRS in distinguishing skill levels during surgical training. Furthermore, Fu et al. (2023) investigated neuroimaging changes during laparoscopic suturing tasks and found a transition in cortical activation patterns from the prefrontal cortex to sensorimotor areas, indicating progressive skill acquisition [64].

These findings are consistent with the work of Holper et al. (2014), who demonstrated that fNIRS derived signals from the dorsolateral prefrontal cortex during dynamic decision making tasks reflect both subjective valuation and individual risk attitude. Their results underscore the modality,s sensitivity to cognitive demands and its relevance in applied high stakes settings such as surgery [65].

Similarly, Ayaz et al. (2012) validated the use of fNIRS in realistic, ecologically valid settings such as air traffic control and UAV piloting. Their results suggest that fNIRS can distinguish between practice levels and mental workload, making it suitable for monitoring skill acquisition in surgical training [17].

Izzetoglu et al. (2007) further demonstrated that fNIRS can be used to monitor hemodynamic changes in brain regions linked to executive function, reinforcing its potential to inform training programs with real time neurofeedback [41].

Beyond surgery, fNIRS has been widely applied in other high stakes fields such as aviation and air traffic control to assess cognitive workload during complex decision making tasks. Kawaguchi et al. (2024) examined prefrontal cortex activation in pilots, demonstrating that experienced aviators exhibited more efficient neural processing, similar to expert surgeons [66]. These findings suggest that fNIRS derived cognitive workload measures can provide valuable insights into expertise development across domains.

The integration of fNIRS with surgical simulation training has also gained traction. Mark et al. (2022) employed neuroadaptive training protocols using fNIRS in flight simulators and found that real time workload assessments could enhance skill acquisition efficiency [67]. A similar approach could be applied to surgical education, where adaptive training based on cognitive workload measures might optimize learning outcomes and reduce skill acquisition time.

The review by Andersen et al. (2024) further supports the use of fNIRS in surgical training. Among the studies included, fNIRS was the most commonly applied modality for real time assessment of cognitive workload, especially in minimally invasive surgery settings [68]. The review also confirmed that fNIRS can differentiate between novice and expert surgeons, with novices typically exhibiting greater activation in the PFC during complex tasks. These findings validate the role of fNIRS as both a research and educational tool in simulation based training programs.

Collectively, these studies highlight the potential of fNIRS in objectively measuring cognitive demands, optimizing surgical training protocols, and bridging the gap between novice and expert performance. These findings provide a foundation for understanding the neural correlates of task performance and inform the development of adaptive training systems, which is central to the objectives of this dissertation. Future research should focus on developing standardized fNIRS based metrics for surgical proficiency assessment and integrating real time cognitive feedback into training modules. Recent studies have proposed the use of composite neurobehavioral metrics like Relative Neural Efficiency (RNE) and Relative Neural Involvement (RNI), combining behavioral scores with hemodynamic activity to reflect both effort and motivation during task performance [25].

1.10 Electrodermal Activity

Electrodermal activity (EDA), a noninvasive and portable measure of skin electrical properties, offers a promising avenue for studying autonomic nervous system activity in ambulatory settings. While its usage in neuroergonomics and mobile brain body research may be less prevalent compared to other biomedical signals, EDA's ability to track physiological arousal makes it a valuable adjunct tool. As an indicator of sweat gland activity, EDA, also known as galvanic skin response (GSR), can provide insights into emotional states, stress levels, and cognitive processes. Incorporating EDA into neuroergonomics and mobile brain body research can enhance our understanding of human technology interactions and inform the design of more effective and user friendly systems [69–73]. The characteristic shape of an SCR, with a rapid onset and gradual recovery, is illustrated in Figure 1.3 the signal waveform of a typical skin conductance response (SCR) [74]. The figure illustrates both tonic (SCL) and phasic (SCR) components of EDA. Phasic responses (SCRs) are typically elicited within 1-3 seconds of stimulus onset and recover within 2-10 seconds. These responses are closely tied to cognitive events and are a primary marker of transient sympathetic arousal [75].

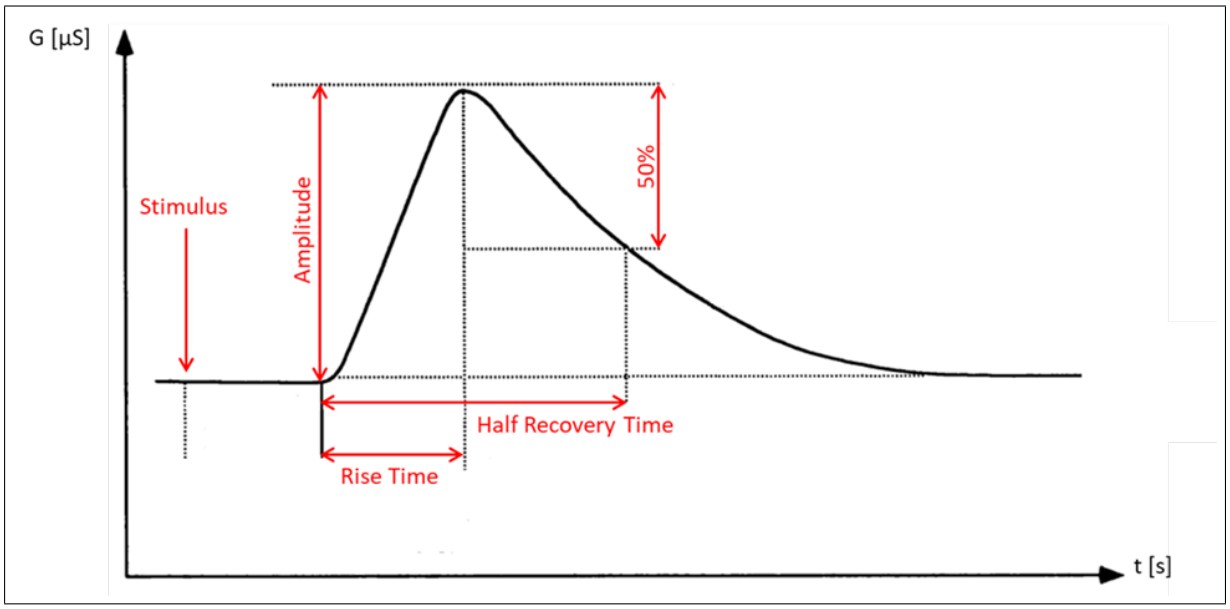

Figure 1.3 The signal waveform of a typical Skin Conductance Response (SCR), illustrating the tonic and phasic components of Electrodermal Activity (EDA) [74].

Electrodermal activity, a well established psychophysiological index, provides valuable insights into emotional arousal by measuring changes in skin electrical properties. The noninvasive application of a low constant voltage allows for the measurement of skin conductance variations, which are influenced by sweat secretion. EDA's ability to assess peripheral autonomic nervous system activity makes it a valuable tool for understanding emotional responses and cognitive processes. The time series of skin conductance can be characterized by both tonic activity (SCL) and phasic activity (SCRs), offering a comprehensive picture of emotional arousal dynamics. Research has consistently shown that EDA is more responsive to emotionally salient stimuli, such as pleasant or unpleasant acoustic stimuli, compared to neutral stimuli [69–72].

EDA comprises two primary components: tonic and phasic [76]. The tonic component, often measured as Skin Conductance Level (SCL), reflects the slower, baseline changes in skin conductance and is associated with general autonomic arousal [77]. The phasic component, manifested as Skin Conductance Responses (SCRs), represents rapid fluctuations in conductance triggered by specific stimuli or internal events [78]. While SCRs have traditionally been the focus of research, recent evidence suggests that both tonic and phasic components are crucial and may originate from distinct neural mechanisms [77]. It's important to recognize that phasic responses constitute only a small fraction of the overall EDA signal, emphasizing the significance of the tonic component in understanding the broader spectrum of autonomic nervous system activity. Figure 1.4 provides a visual representation of the relationship between tonic and phasic components of electrodermal activity in relation to emotional and cognitive states.

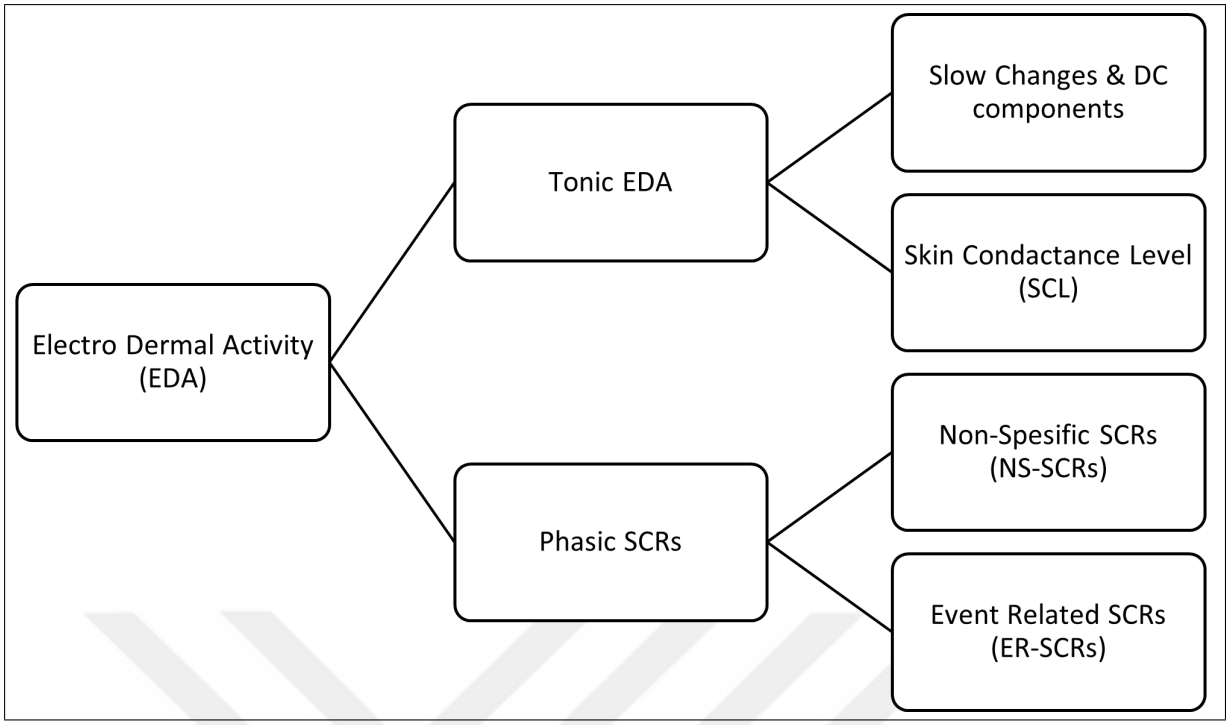

Figure 1.4 Illustrative representation of Electrodermal Activity (EDA) components and their association with cognitive and emotional states. Adapted for neuroergonomic applications.

Electrodermal activity has emerged as a valuable tool for understanding the peripheral physiological underpinnings of emotional states, including fear and disgust. Numerous studies have highlighted differences in the sympathetic nervous system activity associated with these emotions, as reflected in skin conductance changes. EDA's sensitivity to sympathetic activity, coupled with its ability to provide insights into central affective processes, makes it a promising biomarker for studying limbic and ventral prefrontal activations. Previous fNIRS research has demonstrated correlations between prefrontal cortex activation and sympathetic activity during mental tasks, further supporting the link between EDA and central nervous system function [69, 79]. The ability of EDA to differentiate between various affective processes underscores its importance as a complementary method for studying both peripheral autonomic activity and central nervous system activity in parallel [69–73].

1.10.1 Physiological Principles of EDA

Electrodermal activity measures changes in skin electrical properties, primarily influenced by sweat secretion from eccrine sweat glands. These glands play a vital role in thermoregulation and are activated by sympathetic activity within the autonomic nervous system (ANS). The ANS regulates essential bodily functions like temperature, heart rate, and blood pressure, contributing to homeostasis. The sympathetic branch of the ANS is responsible for the fight or flight response, associated with heightened arousal and emotional expressions [69–73, 80]. Studies have consistently demonstrated a strong correlation between bursts of sympathetic ANS activity and EDA signals [80]. This relationship firmly establishes EDA as a reliable indicator of emotional states and arousal [80].

In this study, EDA is employed alongside functional near infrared spectroscopy (fNIRS) to offer a multimodal understanding of psychophysiological responses during task performance. While fNIRS captures cortical activity associated with cognitive workload, EDA provides complementary insight into autonomic arousal, reflecting the activity of the sympathetic nervous system. This combined approach allows for a more comprehensive evaluation of cognitive and emotional demands, particularly within the context of surgical training, and is well aligned with contemporary neuroergonomic frameworks.

The effectiveness of this integrative strategy has been validated by several studies. For instance, Holper and Murphy (2014) utilized fNIRS and EDA concurrently during the Columbia Card Task to investigate hemodynamic and affective responses during risk based decision making. Their results highlighted the differential but complementary roles of fNIRS in capturing prefrontal cognitive activation and EDA in reflecting outcome related emotional arousal [81]. Similarly, Watson et al. (2020) examined audiovisual messaging using a two dimensional model of emotion and found that fNIRS and EDA together provided robust indicators of valence and arousal changes induced by emotionally engaging content. This study further supports the value of combining central and peripheral physiological markers in assessing affective responses under real time experimental conditions [82].

Collectively, the literature highlights the complementary strengths of fNIRS and EDA in capturing the cognitive and emotional components of task engagement. Yet, existing research rarely applies these tools together in high stakes environments like surgery, nor does it fully explore how surgical modality, task complexity, and training context influence cognitive load.

Although fNIRS is increasingly validated in ecological contexts (e.g., classrooms, driving, and surgery), there remains a lack of standardized protocols for integrating it into surgical curricula. The need to bridge simulation and real life environments is especially critical [38]. These gaps motivate the present study.

1.11 Gaps in Existing Research

Despite these advancements, key research gaps remain:

- The comparative cognitive workload associated with laparoscopic versus robotic surgery is not well understood.
- The impact of task complexity on neural efficiency requires further investigation.
- The fidelity of simulation based training compared to real world surgical environments needs validation.

1.12 Research Objectives

This study aims to address these gaps by employing fNIRS to:

1. Compare cognitive workload between laparoscopic and robotic modalities.
2. Evaluate the effects of task complexity (pick and place vs. knot tying) on prefrontal cortex activation.

3. Assess differences in cognitive engagement between simulation and real world surgical training.

1.13 Hypotheses

1. Laparoscopic surgery will elicit higher prefrontal cortex activation than robotic surgery, reflecting greater cognitive workload.
2. Increased task complexity (knot tying vs. pick and place) will lead to elevated cognitive demands and prolonged completion times.
3. Simulation based training, while beneficial, will not fully replicate the cognitive workload demands of real world surgery.

By integrating neuroimaging with surgical training, this research sought to optimize skill acquisition and enhance patient safety through data driven training improvements.

1.14 Conceptual Model

This study is grounded in a conceptual model that maps out the expected relationships between the primary study variables. The three independent variables, Surgical Modality (Laparoscopic vs. Robotic), Task Complexity (pick and place vs. knot tying), and Training Environment (Simulation vs. Real) are hypothesized to influence two dependent outcomes: Cognitive Workload measured by changes in prefrontal HbO/HbR using fNIRS, and Performance, quantified by task completion time.

Building on this, we propose an Efficiency Index, calculated as the ratio of HbO to completion time, to serve as a composite measure of mental effort per unit of performance. This metric may offer valuable insights for optimizing surgical training and understanding

skill acquisition under varying task demands and modalities. Figure 1.5 illustrates the conceptual model of the study.

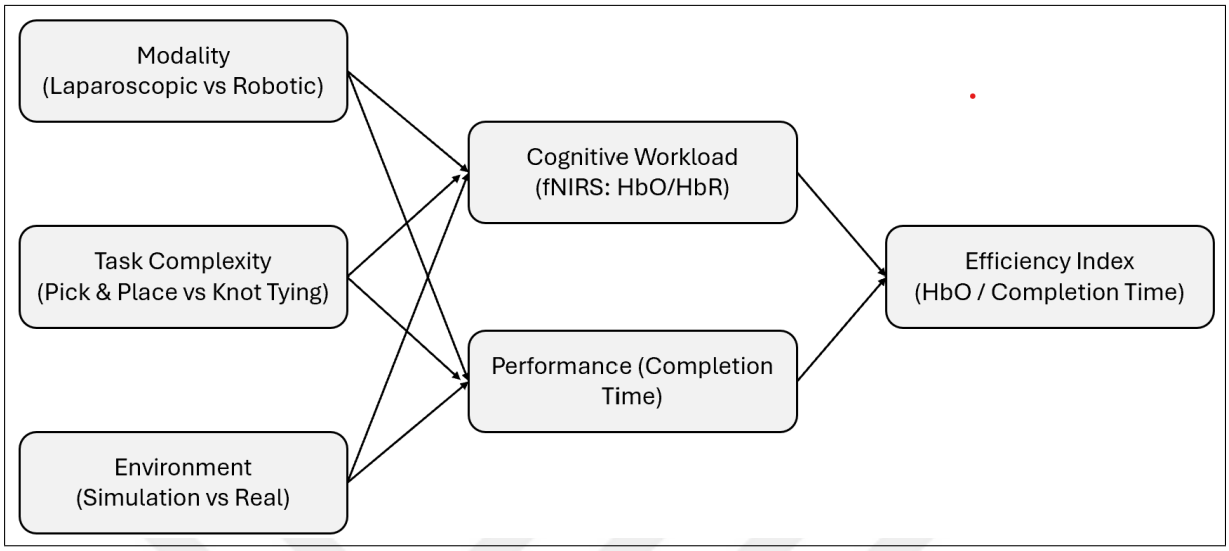

Figure 1.5 Conceptual model of the study. Surgical modality, task complexity, and environment are expected to influence cognitive workload (fNIRS) and performance (completion time), which together inform an efficiency index.

2. METHODS AND MATERIALS

2.1 Participants

A group of volunteers participated in this study, including 21 resident medical doctors from the General Surgery Department of Istanbul Haseki Training and Research Hospital and the Gynecologic Surgery Department of Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, along with 5 specialists in General Surgery and Gynecologic Surgery.

Twenty six participants (14 males, 12 females) were included in the study. The participants had an average age of 30 ± 3 years, with a median age of 29 years (range: 26 to 35 years). All participants were right handed, except for one male who was left handed. Although all participants were novices in robotic surgery, they had comparable laparoscopic surgery experience. The details of the participants are provided in Table 2.1. One participant, due to time constraints, was unable to fully participate in the study and was therefore excluded. Exclusion of this participant did not significantly affect the demographic balance or statistical power of the analysis.

Prior to their involvement in the study, all resident medical specialists from general surgery and gynecologic surgery provided informed consent, ensuring their voluntary participation and comprehension of the research objectives, procedures, and potential risks. This ethical practice is fundamental in safeguarding the rights and well being of participants in research studies. This study was approved by the Ethical Committee of Acibadem Mehmet Ali Aydinlar University (Registration number: ATADEK-2023/05/164) and conducted in accordance with the Declaration of Helsinki, with all participants providing written informed consent. By obtaining informed consent, researchers establish a trusting relationship with participants and ensure that they are fully aware of the potential benefits and drawbacks of their involvement.

Table 2.1
Subject demographics.

Subject	Age	Gender	Handedness	Specialization	Experience Level
Resident (R)					
R1	30	Male	Right	General Surgery	3.5 years
R2	28	Male	Right	General Surgery	1 year
R3	29	Female	Right	General Surgery	3 years
R4	28	Female	Right	General Surgery	2 years
R5	27	Male	Right	General Surgery	1.5 years
R6	33	Male	Right	General Surgery	2 years
R7	28	Female	Right	Gynecology Surgery	0.5 year
R8	28	Female	Right	Gynecology Surgery	0.5 year
R9	29	Female	Right	Gynecology Surgery	1 year
R10	26	Female	Right	Gynecology Surgery	0 year
R11	29	Male	Right	Gynecology Surgery	0 year
R12	35	Male	Right	Gynecology Surgery	0 year
R13	28	Female	Right	Gynecology Surgery	0.5 year
R14	33	Female	Right	Gynecology Surgery	0 year
R15	28	Female	Right	Gynecology Surgery	0 year
R16	26	Male	Right	Gynecology Surgery	0.5 year
R17	26	Male	Right	Gynecology Surgery	0.5 year
R18	28	Female	Right	Gynecology Surgery	2 years
R19	34	Male	Right	Gynecology Surgery	1 year
R20	31	Male	Right	Gynecology Surgery	2 years
R21	31	Male	Right	Gynecology Surgery	3 years
Expert (E)					
E1	35	Male	Left	Gynecology Surgery	3.5 years
E2	34	Female	Right	Gynecology Surgery	3.5 years
E3	34	Female	Right	Gynecology Surgery	4 years
E4	32	Male	Right	Gynecology Surgery	4 years
E5	33	Male	Right	Gynecology Surgery	4 years

2.2 Experimental Protocol

To objectively quantify participantś cognitive workload in real time during simulated laparoscopic and robotic surgical procedures, functional near infrared spectroscopy (fNIRS) sensors were employed. Two standardized surgical tasks were used for this study: knot tying , a complex task, and pick and place , a simpler dexterity based task. These task types are referred to as KT and PP in the rest of the manuscript. Laparoscopic and robotic surgery tasks with different difficulty levels were chosen for simulation sessions. After a baseline measurement with fNIRS, participants first completed a low difficulty pick and place task for familiarization. Task difficulty was calibrated based on complexity of motor coordination and time to completion, with knot tying designated as the complex task. Participants then completed both PP and KT tasks using laparoscopic and robotic systems. Figure 2.1 shows the experimental sensor configuration, including the placement of fNIRS optodes over the participant's forehead targeting the prefrontal cortex (PFC), and EDA electrodes attached to the medial phalanges of the index and middle fingers of the non-dominant hand, used to monitor cognitive workload and physiological arousal during surgical task performance.

Although EDA electrodes were properly positioned and raw signals were recorded throughout the experimental sessions, the study protocol did not include conditions necessary to elicit robust sympathetic nervous system activation. Specifically, no acute stressors such as time pressure, negative performance feedback, or social evaluation were introduced. These types of stimuli are typically required to induce meaningful tonic or phasic EDA responses. As a result, the recorded EDA signals lacked sufficient physiological variation to support reliable interpretation. Therefore, EDA data were excluded from all signal processing, statistical analysis, and result discussions. This decision was made to ensure analytical validity and prevent speculative conclusions based on non-responsive data. Although fNIRS offers portability and motion tolerance, its data interpretation can be affected by extracerebral blood flow, systemic physiological signals, and scalp-hemodynamic coupling. Advanced signal processing techniques, such as short channel regression and systemic physiology removal, are increasingly recommended to isolate true cortical signals [38].

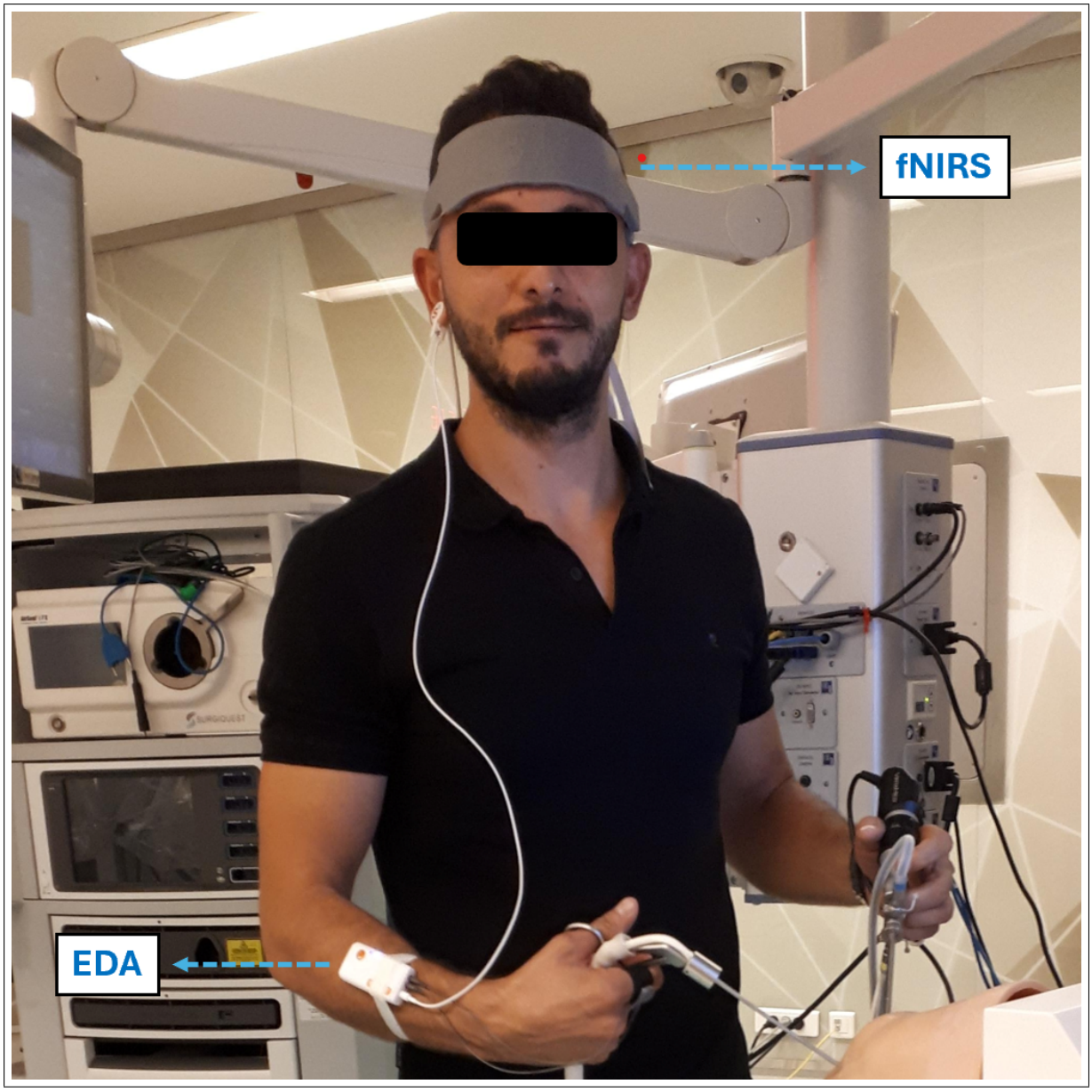

Figure 2.1 fNIRS and EDA sensor placement on a participant during task performance.

To ensure comprehensive data collection and analysis, all simulated laparoscopic and robotic surgery tasks were recorded on video and observed by trained medical specialists. The time taken to complete each task was meticulously recorded, providing valuable data for analysis. This comprehensive approach enabled a detailed analysis of performance, task complexity, and the effects of cognitive workload on surgical outcomes.

KT and PP are fundamental skills in laparoscopic surgery, providing a solid foun-

dation for more complex procedures [48]. These tasks are crucial for developing the fine motor skills, hand eye coordination, and spatial awareness necessary for successful laparoscopic surgery [49].

PP is another vital skill that helps trainees develop dexterity and control within the confined space of the laparoscopic field [35]. This task involves manipulating objects with forceps or other instruments, requiring precision and coordination [48,83]. PP exercises can improve a surgeon's ability to handle delicate tissues, suture needles, and other instruments without causing damage.

Both KT and PP are essential for developing the fundamental skills required for laparoscopic surgery [51]. They provide a solid foundation for more advanced procedures and help to reduce the risk of complications. As such, these tasks are commonly included in laparoscopic surgery training programs [51]. In this study, the term 'real world' refers to task performance on physical box trainers or robotic systems in simulated operating room conditions, rather than live patient surgery.

2.2.1 Laparoscopic Surgery Training Protocol

The study methodology involved a four stage process laparoscopic simulation and box training with real laparoscopic instrumentation. Initially, participants underwent familiarization training with both the laparoscopic simulators and box trainer to ensure familiarization with the equipment. Subsequently, baseline measurements were collected using fNIRS to establish for physiological responses. Following this, participants engaged in PP (Figure 2.2) using the laparoscopic simulator (the Xperience Team Trainer (Mimic Technologies, Seattle, Wash., USA), focusing on mastering the fundamental movements and techniques required for laparoscopic surgery. Finally, participants transitioned to the laparoscopic box trainer (Figure 2.3), replicating the PP in a more realistic environment. This sequential approach allowed for a systematic evaluation of the effectiveness of both simulation and box training in developing laparoscopic skills and understanding the associated physiological responses. The workflow of the first part of the study is shown in

Figure 2.4.

Figure 2.2 PP module on the laparoscopic surgery simulator.

Figure 2.3 PP module on the laparoscopic surgery box trainer.

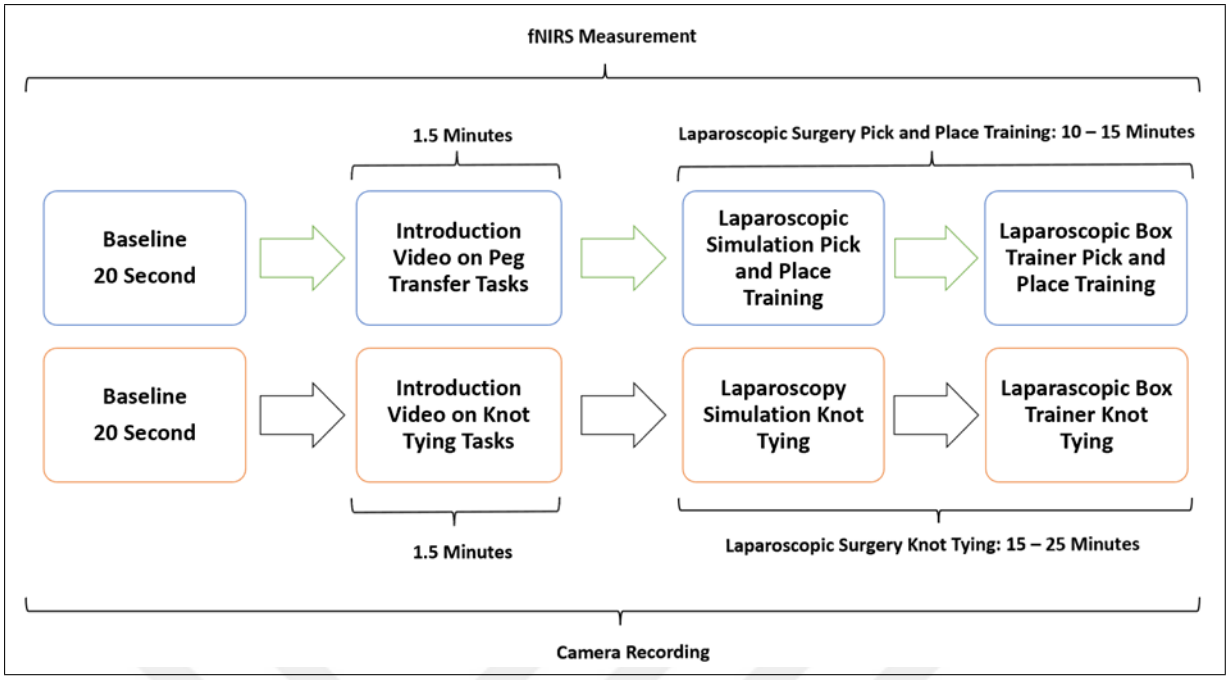

Figure 2.4 Workflow of the of PP and KT using both a laparoscopic surgery simulator and a real life laparoscopic box trainer.

The second part of the study focused on comparing the effectiveness of laparoscopic surgery simulation and box trainer training in improving KT skills. Participants underwent baseline measurements using fNIRS to assess their physiological responses during cognitive load. Subsequently, they received training in KT (Figure 2.5) using both laparoscopic surgery simulation (LapVR simulator, CAE Healthcare, Saint-Laurent, Quebec, Canada) and box trainer platforms. The training sessions were designed to familiarize participants with the equipment, techniques, and challenges associated with laparoscopic KT. Following the training, participants performed KT (Figure 2.6) using both platforms, allowing for a direct comparison of their performance and physiological responses. The workflow of the second part of the study is shown in Figure 2.4.

Figure 2.5 KT module on the laparoscopic surgery simulator.

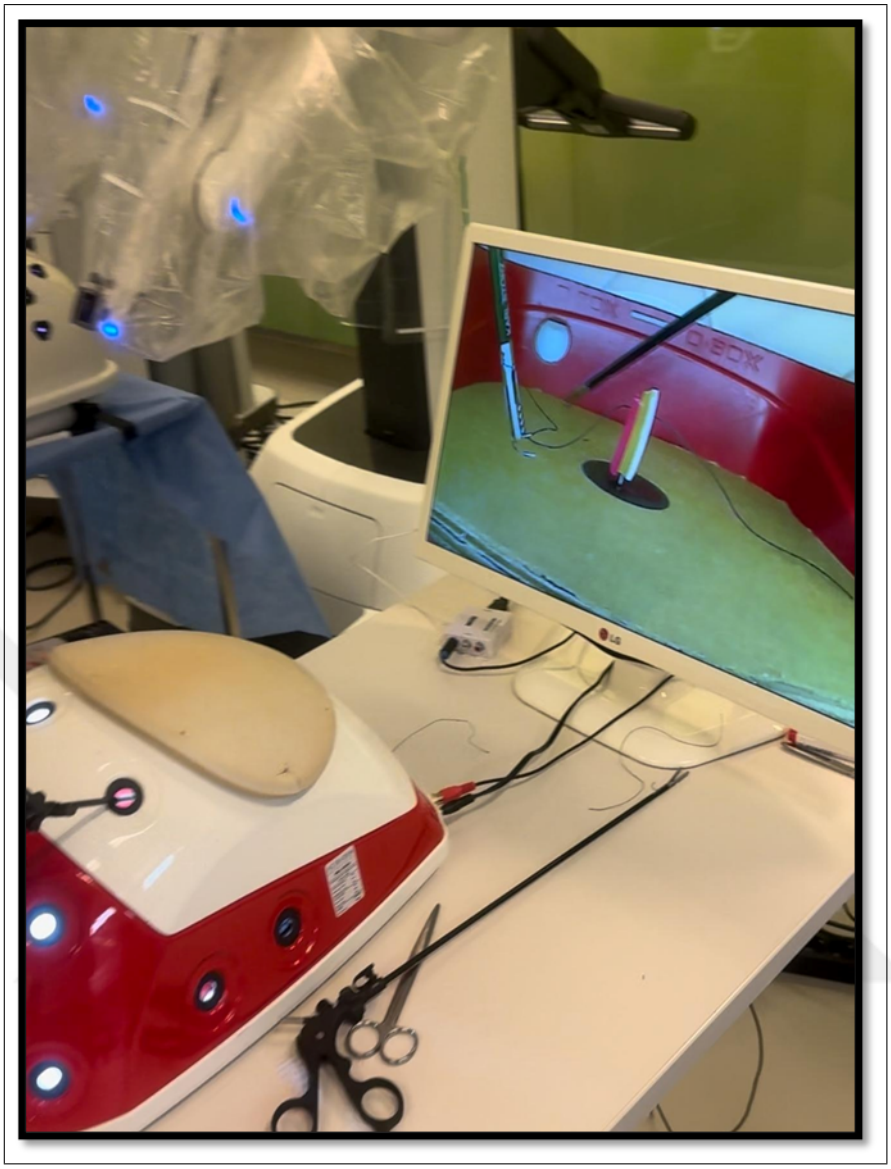

Figure 2.6 KT module on the laparoscopic surgery box trainer.

2.2.2 Robotic Surgery Training Protocol

The third phase of the study involved PP using both a robotic simulator and a real robotic surgery robot. This training aimed to familiarize participants with robotic surgery's interface and procedures to assess their performance and physiological responses. Participants were first trained on the robotic surgery simulator (Figure 2.7), Mimic® dV-Trainer (Mimic Technologies, Seattle, Wash., USA) , to develop fundamental skills and to become accustomed to the robotic interface. Subsequently, they transitioned to the real robotic surgery robot (Figure 2.8) , da Vinci Surgical System (Intuitive Surgical,

Sunnyvale, Calif., USA), where they performed the same PP in a simulated surgical environment. This sequential approach allowed for gradual progression in complexity and ensured that participants were adequately prepared before engaging in the actual robotic system. The workflow of the third part of the study is shown in Figure 2.9.

Figure 2.7 PP module on the robotic surgery simulator.

Figure 2.8 PP training module on the robotic surgery platform.

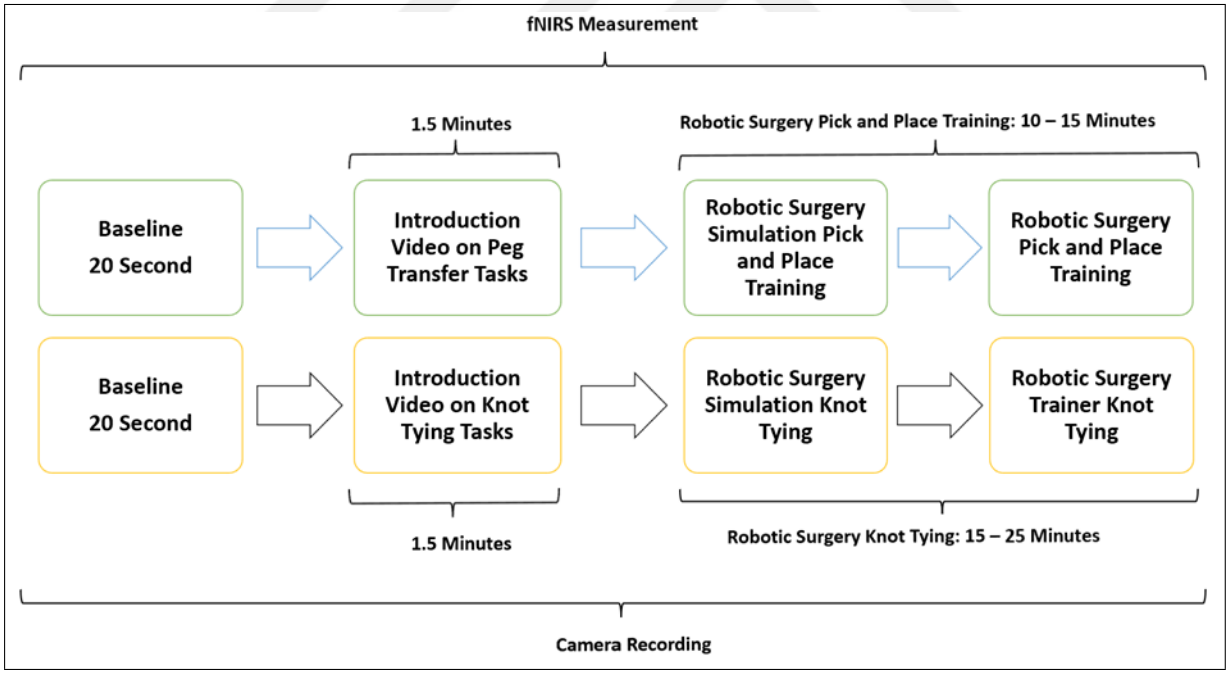

Figure 2.9 The workflow of the study included PP and KT using both a robotic surgery simulator and a real life robotic surgery platform.

The fourth phase of the study involved conducting KT training sessions using both

the robotic simulator (Figure 2.10) and the robotic surgery robot (Figure 2.11). Participants were initially trained on the robotic simulator to familiarize themselves with the system's controls, haptic feedback, and visual interface. Once comfortable with the simulator, they transitioned to the robotic surgery robot for hands on KT practice. Throughout the training sessions, physiological data, fNIRS, were continuously measured to assess participants' cognitive load, and brain activation patterns during the task. The workflow of the fourth part of the study is shown in Figure 2.9.

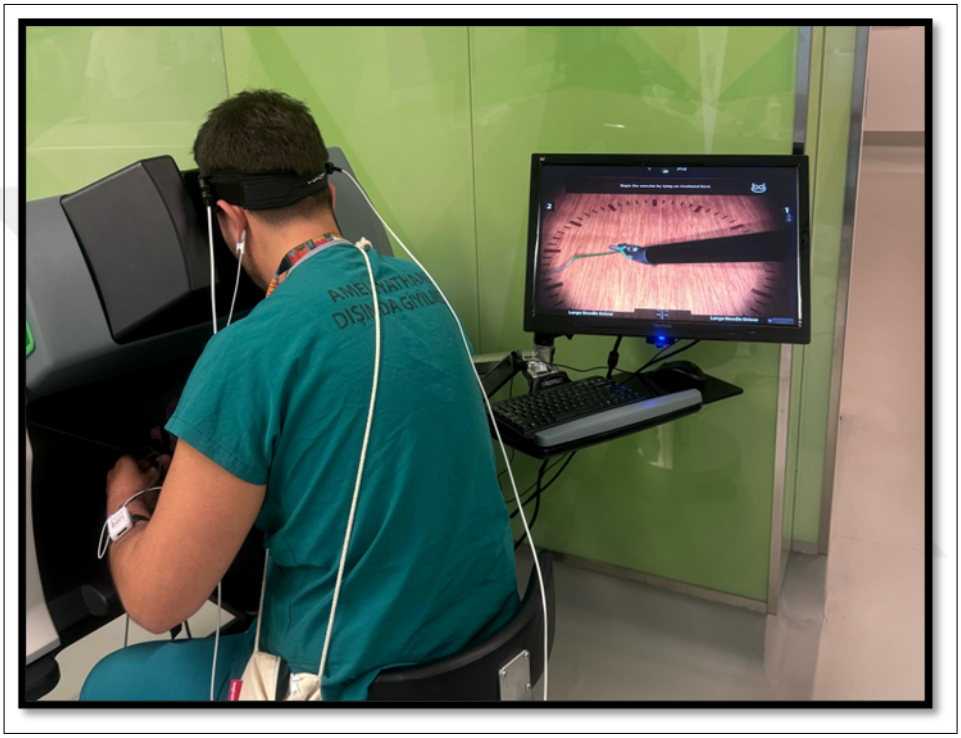

Figure 2.10 KT module on the robotic surgery simulator.

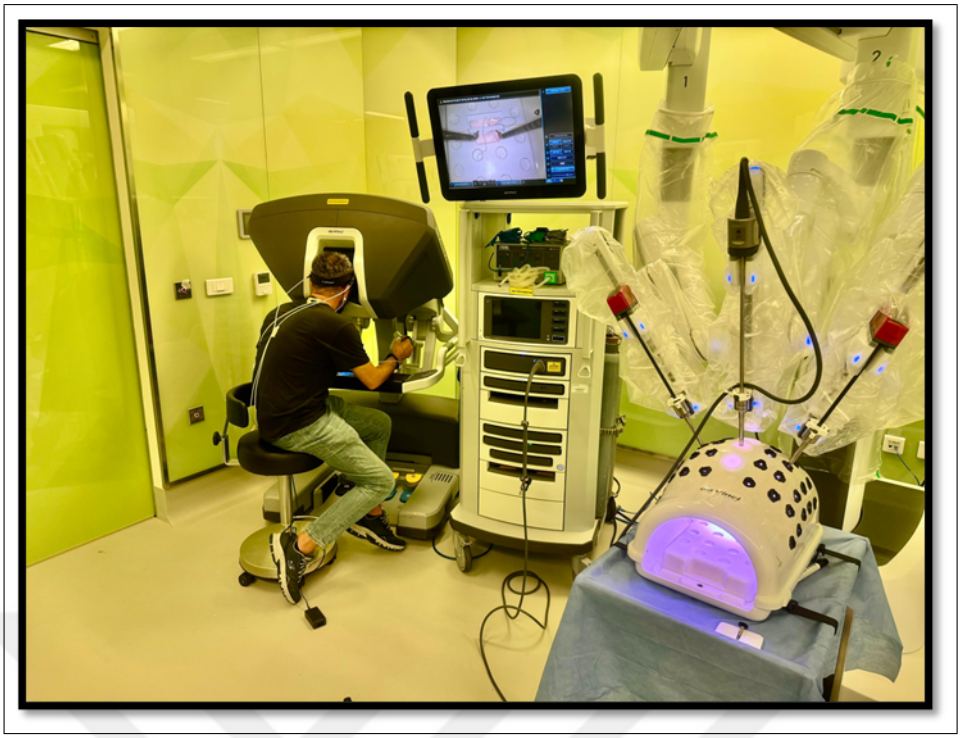

Figure 2.11 KT training module on the robotic surgery platform.

2.3 fNIRS Recording and Preprocessing

To monitor hemodynamic responses in the prefrontal cortex (PFC), participants were equipped with a continuous wave functional near infrared spectroscopy (fNIRS) device (fNIRS Devices, LLC, Potomac, MD). This device utilized a 18 channel probe positioned over the PFC (Figure 2.12 (a)), comprising 4 LED sources emitting light at 730 nm and 850 nm wavelengths and 10 photodetectors as shown in Figure 2.12 (b). To ensure accurate time synchronization during post processing, the start and end times of each session and simulation task were marked and time synchronized. This study employed continuous wave (CW) fNIRS, which estimates relative concentration changes based on modulated light intensity. Unlike frequency or time domain systems, CW systems cannot independently resolve absolute concentrations or pathlength but offer practical benefits in portability and cost.

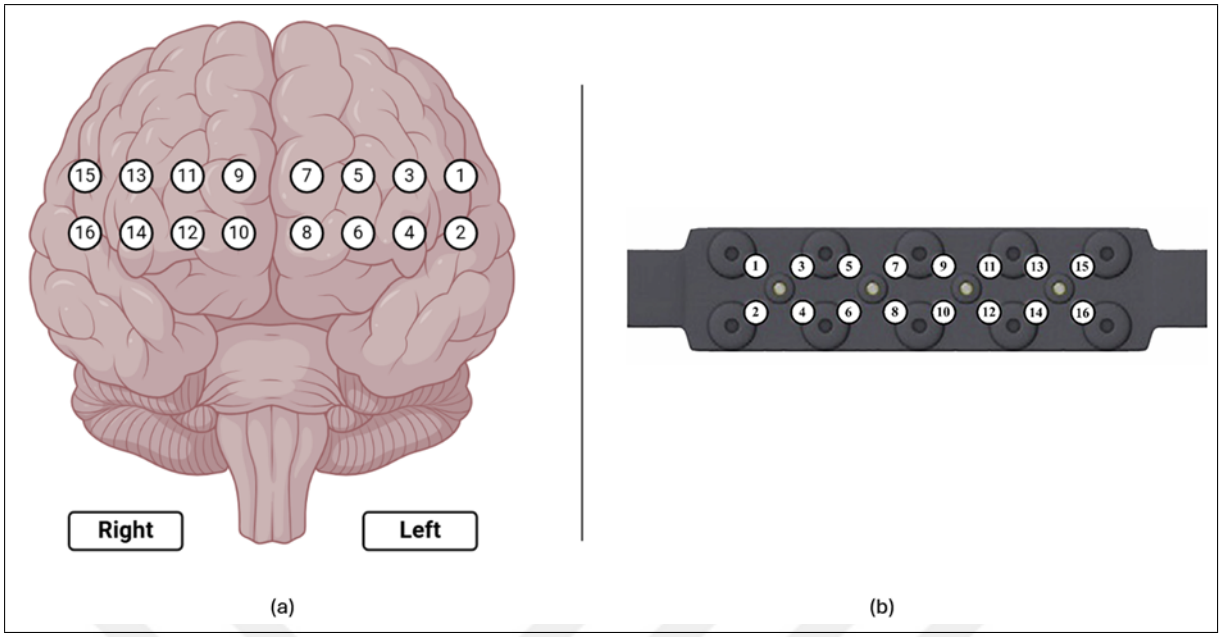

Figure 2.12 (a): Corresponding 16 optodes locations over the prefrontal cortex. (b): fNIRS probe with 4 light sources and 10 detectors.

The quality of fNIRS signals can be affected by various noise sources, including instrument-related noise (e.g., light source instability, electronic noise), physiological interference (e.g., respiration, heartbeat), and motion artifacts. Therefore, effective noise removal is a crucial preliminary step in the data processing pipeline [84–86].

MATLAB (MathWorks, R2022b) was used to preprocess the light intensity signals obtained from fNIRS and to compute changes in oxyhemoglobin (ΔHbO) during the simulation tasks. [87]. Head movements, causing relative shifts between source detector pairs and the scalp, can introduce motion artifacts into fNIRS signals, manifesting as rapid, large magnitude spikes that significantly exceed tissue related hemodynamic changes [88, 89]. Channel rejection and baseline normalization ensured cleaner signals for analysis. Physiological signals like heart rate (over 0.5 Hz) and respiration (over 0.2 Hz) exhibit higher frequency ranges than hemodynamic responses and instrument degradation induced noise (3-5 Hz) [84, 86].

Channels exhibiting saturation ($\lambda_{850} > 4300$) or unusual intensity patterns ($\lambda_{730} > \lambda_{850}$) were manually rejected prior to further analysis to ensure data integrity. To reduce

physiological artifacts such as cardiac, respiratory, and Mayer waves [90], the light intensity signals were low pass filtered using a 250th order finite impulse response (FIR) filter with a cutoff frequency of 0.09 Hz. Additionally, oscillatory artifacts were identified using the Multiscale Oscillatory Dynamics Analysis (MODA) toolbox [91], which applies Morlet wavelet ridge extraction to detect oscillation patterns [92]. Targeted band stop filtering was then used to remove the identified oscillations. If no oscillatory artifacts were present, only the FIR low pass filter was applied. Following preprocessing, the modified Beer Lambert Law was used to compute the changes in oxyhemoglobin (ΔHbO) and deoxyhemoglobin (ΔHbR) concentrations for each channel.

Following artifact removal and channel rejection procedures, for each participant, changes in oxy-Hb and deoxy-Hb concentrations ($\mu mol\,L^{-1}$) over time were calculated using the modified Beer Lambert Law [93] based on optical density (OD) changes measured at 730 nm and 850 nm [84, 85]. This allowed for the quantitative estimation of key hemodynamic parameters on a channel specific basis. These parameters included oxyhemoglobin (HbO), deoxyhemoglobin (HbR), relative changes in oxygen availability (HbDiff or Oxy), and total hemoglobin concentration (HbTotal, approximated as the sum of (HbO and HbR). A 10 s baseline measurement [17,85] was acquired from each participant after a 20 s relaxation period, prior to laparoscopic and robotic surgery training. Changes in HbO_2 and HHb concentrations were calculated relative to this baseline throughout the duration of each task.

An 18 channel fNIRS system, sampling at 10 Hz, was employed, targeting the prefrontal cortex (PFC) with a sensor configuration comprising LEDs, detectors, and short separation channels. This system comprised 16 long separation channels for cortical activity measurements and 2 short separation channels for assessing superficial layer hemodynamics. This study focused on the left and right prefrontal cortex as regions of interest (ROIs) as well as the identification and specification of a feature of interest of the fNIRS biomarkers, given their established roles in brain activity during learning and training [17, 25, 37, 94].

The areas corresponding to the 18 fNIRS channels included the right and left

superior and inferior frontal cortices. These channels were grouped into four regions of interest (ROIs), located as follows: the left dorsolateral prefrontal cortex (DLPFC; optodes 1-4), the left anterior medial prefrontal cortex (aMPFC; optodes 5-8), the right anterior medial prefrontal cortex (optodes 9-12), and the right dorsolateral prefrontal cortex (optodes 13-16), as illustrated in Figure 2.13.

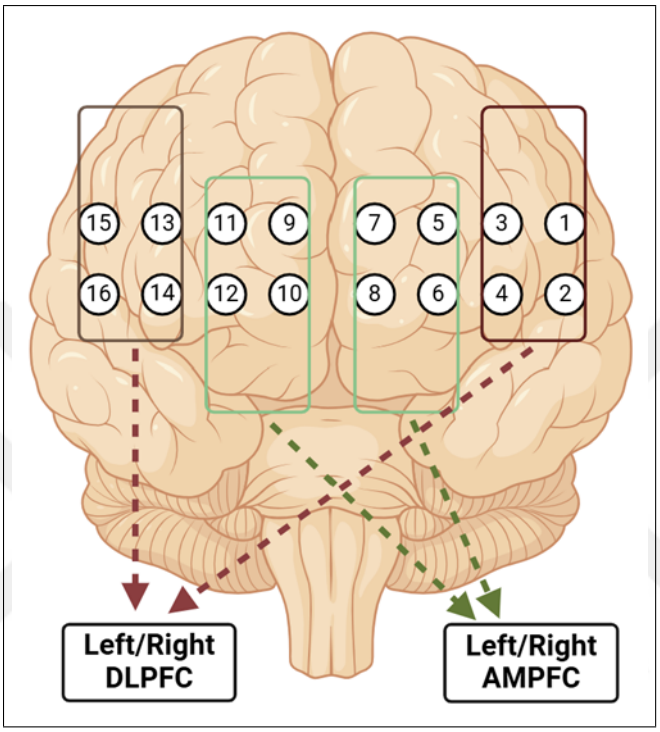

Figure 2.13 Placement of fNIRS device and location of 18 channels categorized as four regions of interest.

2.4 Statistical Analysis

All statistical analyses were performed to evaluate the effects of surgical modality (laparoscopic vs. robotic), task complexity PP vs. KT, training environment (simulation vs. real world), and experience level (resident vs. specialist) on behavioral and neurophysiological outcomes.

To determine whether the missing data in the dataset occurred at random, Little's Missing Completely at Random (MCAR) test was performed. This test evaluates the null hypothesis that the pattern of missingness is completely random and unrelated to any

observed or unobserved variables. The results of Little's MCAR test were non-significant ($\chi^2 = 5.163$, $df = 441$, $p = 1.000$), indicating that the missing data can be considered MCAR. This outcome implies that the missing values are not systematically related to any measured variables, thereby justifying the assumption that parameter estimates derived from subsequent analyses are unlikely to be biased due to missingness.

Given the presence of missing data and the repeated measures structure, a linear mixed-effects regression (LME) approach was employed to analyze the dependent variables: Completion Time and Mean ΔHbO. Mean ΔHbO represents the average change in oxygenated hemoglobin from baseline, a validated neurophysiological index of cognitive effort. LME models were used to account for both fixed effects (Session, Task Complexity, Modality, Environment) and random effects (individual participants). Each participant was modeled with a random intercept to control for baseline inter-individual differences. It is an important neurophysiological marker of cerebral oxygenation and neural activation, particularly in cognitive workload assessment, skill acquisition analysis, and task efficiency evaluation.

Since surgery requires high cognitive load, motor coordination, and decision making, tracking Mean ΔHbO in different brain regions can help understand how trainees develop expertise and manage cognitive resources. Before model fitting, missing values were checked and imputed when necessary, and outliers were removed using the Z-score method within each session separately to avoid excessive data exclusion. Additionally, log10 transformation was applied to dependent variables when normality assumptions were violated.

To account for within-subject dependencies, two linear mixed-effects (LME) models were tested. Each participant was modeled with a random intercept to control for baseline inter-individual differences, and fixed effects were included for Session, Modality, Environment, and Task Complexity. The first model included 'Session' as a fixed effect, where Session was a categorical variable representing the combination of task repetitions and simulation type. The second model treated Modality (Laparoscopic vs. Robotic), Media (Real vs. Simulation), and Scenario (KT vs. PP) as separate fixed effects to as-

sess their individual contributions. In both models, participants were modeled as random intercepts to account for inter-individual variability in task performance.

Model selection was carried out using maximum likelihood estimation (ML), which is suitable for comparing nested models. To evaluate model fit while accounting for complexity, both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were used—AIC focuses on balancing model fit and complexity, while BIC applies a stronger penalty for overly complex models. When some models failed to converge, the Powell optimization method was used instead, as it performs well in handling complex, non-linear problems.

To improve the reliability of parameter estimates, robust standard errors were used in cases where heteroscedasticity (unequal variance across observations) was detected. The best fitting model was determined by comparing several metrics: AIC, BIC, log-likelihood, deviance, marginal R^2 (explained variance due to fixed effects), conditional R^2 (explained variance due to both fixed and random effects), and the intraclass correlation coefficient (ICC), which reflects how much of the variance can be attributed to differences between participants.

Model assumptions were assessed using the Shapiro–Wilk test for normality, Levene's test for homogeneity of variances, and visual inspection of residual diagnostic plots. Multicollinearity between predictors was evaluated using the Variance Inflation Factor (VIF). Where necessary, robust standard errors were again applied to account for any violations of the assumption of homoscedasticity.

For variables with significant fixed effects, pairwise post hoc comparisons were performed using estimated marginal means (EMMs), with p-values adjusted via the Benjamini-Hochberg False Discovery Rate (FDR) method.In addition to the LME-based analyses, paired t-tests were used where appropriate, or example, when comparing mean completion times or hemodynamic responses within a single task or modality. For these analyses, both Bonferroni-corrected and raw p-values were reported to transparently demonstrate the influence of multiple comparison correction on significance. Effect sizes

for each contrast were reported using Cohen's d. Although no continuous covariates were included in the current models, the linear mixed effects (LME) framework supports their inclusion.

To evaluate differences in task performance and neurophysiological responses across conditions, paired-samples t-tests were performed to compare simulation versus real environments and peg transfer (PP) versus knot tying (KT) tasks. Additionally, independent-samples t-tests were conducted to assess differences between surgical assistants and expert surgeons.

To control for multiple comparisons, the Benjamini-Hochberg procedure was applied to adjust all p-values, thereby reducing the false discovery rate (FDR) while preserving statistical power. Corrected p-values are reported alongside Cohen's d effect sizes, with 0.2, 0.5, and 0.8 interpreted as small, medium, and large effects, respectively.

All t-test analyses were conducted using SPSS version 27 (IBM Corp., Armonk, NY), while the Benjamini Hochberg correction was implemented in Python using the `multipletests` function from the `statsmodels` package. Statistical significance was determined at $\alpha = 0.05$. Cohen's d was computed to determine effect sizes, where $d = 0.2$ was considered a small effect, $d = 0.5$ a medium effect, and $d \geq 0.8$ a large effect [95]. All statistical analyses were performed in Python (version 3.13) using the `statsmodels` package [96] for mixed effects modeling, `pingouin` for missing data analysis and effect size calculations [97], `scipy` for statistical testing [98], and `matplotlib` and `seaborn` [99] for visualization and assumption checking.

Boxplots represent the median values of the corresponding measures (e.g., task completion time, mean HbO/HbR levels). The boxes indicate the interquartile range (IQR), spanning from the first to the third quartile. The whiskers extend to the smallest and largest data points within 1.5 times the IQR, representing the typical spread of the data. Data points beyond this range are plotted individually as circles, indicating outliers and illustrating variability beyond the IQR.

3. RESULTS

This chapter presents the outcomes of the study, focusing on cognitive workload and task performance across laparoscopic and robotic surgical modalities. Comparisons were made between simulation and real surgical environments, as well as between two task types: pick and place and knot tying . Cognitive workload was assessed using fNIRS derived data from the prefrontal cortex, and performance was measured by task completion time. Electrodermal Activity (EDA) data were also collected but are not included here due to the lack of stress inducing elements in the task design.

To evaluate the effects of surgical modality, task complexity, and training environment, two main outcome measures were analysed: task completion time and mean change in oxygenated hemoglobin concentration (ΔHbO). These were examined using linear mixed effects models that accounted for repeated measurements within participants. Where appropriate, post hoc comparisons were made, and the false discovery rate was controlled using the Benjamini-Hochberg procedure. The results are presented in the sections that follow.

To analyze the effects of different surgical conditions on participants' performance, we used the log-transformed task completion time as the dependent variable. This transformation was applied to better meet the assumptions of normality and homoscedasticity required for linear mixed effects modeling.

Two distinct model structures were tested to evaluate how best to explain variance in completion time. The first model, referred to as the Session model, treated each unique combination of Surgical Modality (Laparoscopic vs. Robotic), Task Environment (Simulation vs. Real), and Task Type (knot tying vs. pick and place) as a single categorical variable labeled "Session." This approach allowed for the evaluation of performance differences across specific task conditions as discrete units, without assuming any interaction structure among the factors.

The second model, termed the Full Factorial model, included Modality, Environment, and Task Type as separate fixed effects. It was structured to account for both main effects and possible interaction effects among these variables, thereby assuming additive and interactive contributions to task performance.

Model fit comparisons were conducted using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and a likelihood ratio test. The Session model demonstrated superior fit, with a lower AIC (279.589) and BIC (312.216) compared to the Full Factorial model (AIC = 307.826; BIC = 327.402). Additionally, the likelihood ratio test indicated that the Session model significantly outperformed the Full Factorial model ($\chi^2 = 36.232$, $p < 0.001$). These results suggest that modeling the condition as a single categorical factor (Session) more effectively captured the nuances of task performance across different surgical contexts. A detailed comparison of model performance metrics is provided in Table 3.1.

Finally, the Intraclass Correlation Coefficient (ICC) for the model was calculated at 0.120. This value indicates that approximately 12% of the total variance in log-transformed completion time was attributable to individual differences between participants, with the remaining 88% attributed to within participant variance due to task condition effects.

Table 3.1
Model performance comparisons between the Session-based model and the Full Factorial model.

Metric	**Completion Time (Session)**	**Completion Time (Full Factorial)**	**Mean ΔHbO (Session)**	**Mean ΔHbO (Full Factorial)**
AIC	279.589	307.826	842.519	882.962
BIC	312.216	327.402	875.093	902.507
Log Likelihood	129.794	-147.913	-411.259	-435.481
Deviance	259.589	295.826	822.519	870.962
$R^2_{\text{Conditional}}$	0.305	0.094	0.491	0.351
R^2_{Marginal}	1.000	1.000	0.564	0.416
ICC	0.120	0.074	0.144	0.099

Post hoc analyses were conducted to further explore the significant effects identified in the primary comparisons. These analyses revealed pairwise differences in task completion time (log-transformed) and mean changes in oxygenated hemoglobin (ΔHbO) across surgical modalities, task complexities, and environments. The results, presented in Table 3.2, highlight statistically significant contrasts ($p < 0.05$) alongside large effect sizes (Cohen's $d \geq 0.8$), particularly between robotic and laparoscopic modalities during complex tasks. These findings provide a more granular understanding of how cognitive workload and performance efficiency vary across experimental conditions. Abbreviated labels are used in the table for clarity and brevity. Log_{10} indicates the mean difference in log-transformed completion time, while *Sec* refers to the mean difference in raw seconds. *Adj. p** denotes the adjusted p-value, and d represents Cohen's d effect size. The label *Diff. Adj. p* d* is a compact format combining the mean difference, adjusted p-value, and Cohen's d.

Table 3.2
Post hoc comparisons for completion time (log_{10} transformed) and mean ΔHbO.

Contrast		**Completion Time**				**Mean ΔHbO**		
		Log_{10}	Sec	Adj. p*	d	Diff.	Adj. p*	d
Lap_Real_KT	Lap_Real_PP	0.299	1.99	0.417	-0.436	-1.808	0.077	0.791
Lap_Real_KT	Lap_Sim_KT	-0.318	0.48	0.322	0.558	-1.248	0.468	0.523
Lap_Real_KT	Rob_Real_KT	-0.219	0.60	0.795	0.374	**-5.520**	**<0.001**	**2.235**
Lap_Real_KT	Rob_Real_PP	0.074	1.19	1.000	-0.107	**-3.607**	**<0.001**	**1.791**
Lap_Real_KT	Rob_Sim_KT	0.378	2.39	0.142	-0.680	**-6.100**	**<0.001**	**2.269**
Lap_Real_PP	Lap_Sim_KT	**-0.617**	0.24	**0.001**	**1.238**	0.560	0.985	-0.256
Lap_Real_PP	Lap_Sim_PP	-0.249	0.56	0.654	0.456	**-2.134**	**0.017**	**1.195**
Lap_Real_PP	Rob_Real_PP	-0.225	0.6	0.749	0.359	-1.799	0.075	1.011
Lap_Real_PP	Rob_Sim_PP	-0.333	0.46	0.278	0.680	**-2.467**	**0.003**	**1.277**
Lap_Sim_KT	Lap_Sim_PP	0.368	2.33	0.157	-0.942	**-2.695**	**0.001**	**1.413**
Lap_Sim_KT	Rob_Sim_KT	**0.696**	4.97	**<0.001**	**-2.342**	**-4.852**	**<0.001**	**1.862**
Lap_Sim_PP	Rob_Real_PP	0.024	1.06	1.000	-0.043	0.336	0.999	-0.237
Lap_Sim_PP	Rob_Sim_KT	0.328	2.13	0.295	-0.885	**-2.157**	**0.015**	**0.948**
Lap_Sim_PP	Rob_Sim_PP	-0.084	0.82	0.999	0.221	-0.333	1.000	0.207
Rob_Real_KT	Rob_Real_PP	0.292	1.96	0.448	-0.567	**1.913**	**0.049**	**-0.952**
Rob_Real_KT	Rob_Sim_KT	**0.597**	3.95	**0.001**	**-1.844**	-0.580	0.984	0.216
Rob_Real_KT	Rob_Sim_PP	0.185	1.53	0.903	-0.555	1.244	0.514	-0.580
Rob_Real_PP	Rob_Sim_KT	0.304	2.01	0.379	-0.630	**-2.493**	**0.002**	**1.097**
Rob_Real_PP	Rob_Sim_PP	-0.108	0.78	0.995	0.219	-0.669	0.962	0.418
Rob_Sim_KT	Rob_Sim_PP	-0.412	0.39	0.079	**1.465**	1.824	0.079	-0.762

Independent samples t-tests were conducted to compare task performance and pre-

frontal hemodynamic responses (HbO and HbR) across different experimental conditions, including simulation versus real environments, laparoscopic versus robotic modalities, and PP versus KT tasks. For each comparison, the results include the t-value, uncorrected p-value, FDR-adjusted p-value (p^*), and Cohen's d effect size. A significance threshold of 0.05 was applied, with p-values corrected for multiple comparisons using the False Discovery Rate (FDR) method.

Tables 3.3 to 3.14 summarize the outcomes of these t-tests, presenting comparisons of task completion times and hemodynamic responses across the various groups. Statistically significant differences ($p < 0.05$, FDR-corrected) were found in several conditions, particularly in oxygenated hemoglobin (HbO) responses and task durations.

Table 3.3
Laparoscopic Surgery PP Task: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between simulation and real environments.

	Simulation	**Real**	**t**	**p**	**p***	**Effect size**
	Mean±SD	**Mean±SD**				
Completion time	84.01±52.16 62.76 (32.82–251.81)	124.01±108.92 82.59 (25.59–492.3)	-1.513	0.143	0.165	-0.303
Mean Left HbO	0.19±2.15 -0.15 (-2.94–6.28)	2.48±2.42 2.38 (-0.45–6.93)	-5.103	<0.001	0.001	-1.021
Mean Left HbR	0.07±1.79 0.49 (-6.52–2.25)	0.29±1.94 0.35 (-5.67–2.53)	-0.817	0.422	0.452	-0.163
Mean Right HbO	1.04±2.00 0.88 (-2.34–5.97)	2.82±3.09 3.07 (-2.82–10.27)	-3.126	0.005	0.015	-0.625
Mean Right HbR	-0.76±2.23 -1.39 (-5.73–4.4)	1.05±3.53 0.04 (-4.63–13.25)	-2.964	0.007	0.017	-0.593

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.4
Laparoscopic Surgery KT Task: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between simulation and real environments.

	Simulation	**Real**	**t**	**p**	**p***	**Effect size**
	Mean±SD	**Mean±SD**				
Completion time	50.5±16.68 46.83 (24.39–88.83)	99.31±95.19 65.61 (17.19–414.68)	-2.667	0.014	0.030	-0.533
Mean Left HbO	2.8±2.57 2.69 (-1.22–8.23)	4.12±3.31 4.16 (-1.44–10.86)	-3.476	0.002	0.010	-0.695
Mean Left HbR	0.45±1.94 1.06 (-4.73–4.53)	0.23±2.04 0.8 (-4.04–3.41)	1.240	0.227	0.378	0.248
Mean Right HbO	3.7±3.06 3.42 (-1.54–12.97)	4.98±2.73 5.1 (-0.82–11.93)	-3.558	0.002	0.010	-0.712
Mean Right HbR	0.69±3.37 0.22 (-4.46–13.12)	0.55±3.32 0.22 (-4.09–12.92)	0.484	0.633	0.678	0.097

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.5
Laparoscopic Surgery Simulation Environment: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between PP and KT tasks.

	PP Task	**KT Task**	**t**	**p**	**p***	**Effect size**
	Mean±SD	**Mean±SD**				
Completion time	84.01±52.16 62.76 (32.82–251.81)	50.5±16.68 46.83 (24.39–88.83)	3.434	0.002	0.004	0.687
Mean Left HbO	0.19±2.15 -0.15 (-2.94–6.28)	2.8±2.57 2.69 (-1.22–8.23)	-5.420	<0.001	<0.001	-1.084
Mean Left HbR	0.07±1.79 0.49 (-6.52–2.25)	0.45±1.94 1.06 (-4.73–4.53)	-1.257	0.221	0.237	-0.251
Mean Right HbO	1.04±2.00 0.88 (-2.34–5.97)	3.7±3.06 3.42 (-1.54–12.97)	-4.481	<0.001	<0.001	-0.896
Mean Right HbR	-0.76±2.23 -1.39 (-5.73–4.4)	0.69±3.37 0.22 (-4.46–13.12)	-2.663	0.014	0.023	-0.533

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.6
Laparoscopic Surgery Real Environment: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between PP and KT tasks.

	PP Task	KT Task	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	124.01±108.92 82.59 (25.59–492.3)	99.31±95.19 65.61 (17.19–414.68)	0.981	0.336	0.458	0.196
Mean Left HbO	2.48±2.42 2.38 (-0.45–6.93)	4.12±3.31 4.16 (-1.44–10.86)	-3.876	0.001	0.003	-0.775
Mean Left HbR	0.29±1.94 0.35 (-5.67–2.53)	0.23±2.04 0.8 (-4.04–3.41)	0.196	0.846	0.906	0.039
Mean Right HbO	2.82±3.09 3.07 (-2.82–10.27)	4.98±2.73 5.1 (-0.82–11.93)	-4.546	<0.001	0.001	-0.909
Mean Right HbR	1.05±3.53 0.04 (-4.63–13.25)	0.55±3.32 0.22 (-4.09–12.92)	1.887	0.071	0.122	0.377

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.7
Robotic Surgery PP Task: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between simulation and real environments.

	Simulation	Real	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	69.5±23.74 68.79 (37.23–144.6)	85.74±56.34 79.41 (23.58–249.2)	-1.314	0.201	0.302	-0.263
Mean Left HbO	-0.55±3.29 -0.32 (-9.51–4.88)	0.36±3.02 0.01 (-6.74–5.85)	-1.756	0.092	0.291	-0.358
Mean Left HbR	0.67±2.72 0.47 (-5.39–9.53)	0.51±2.4 0.25 (-4.36–7.07)	0.805	0.429	0.515	0.164
Mean Right HbO	0.47±2.61 0.23 (-7–4.12)	1.47±2.36 1.71 (-3.33–5.04)	-1.912	0.068	0.291	-0.382
Mean Right HbR	0.68±4.49 -0.02 (-2.73–20.32)	-0.34±3.1 -0.73 (-6.78–7.31)	0.879	0.388	0.515	0.176

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.8
Robotic Surgery KT Task: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between simulation and real environments.

	Simulation	Real	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	46.96±16.57 41.88 (26.41–92.17)	70.72±49.95 53.17 (17.7–214.86)	-2.264	0.032	0.064	-0.462
Mean Left HbO	1.6±2.8 1.74 (-5.13–6.9)	2.32±2.88 2.57 (-3.61–9.47)	-1.179	0.246	0.295	-0.241
Mean Left HbR	0.31±2.44 0.47 (-4.25–8.82)	0.38±2.57 0.61 (-4.48–6.26)	-0.086	0.932	0.932	-0.017
Mean Right HbO	1.62±3.13 1.57 (-6.16–7.39)	2.77±2.3 2.67 (-1.05–7.25)	-2.261	0.032	0.064	-0.461
Mean Right HbR	-0.13±3.05 -0.35 (-5.62–6.53)	-0.15±2.49 -0.5 (-5.47–5.28)	0.066	0.948	0.948	0.013

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.9
Robotic Surgery Simulation Environment: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between PP and KT tasks.

	PP Task	KT Task	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	69.5±23.74 68.79 (37.23–144.6)	46.96±16.57 41.88 (26.41–92.17)	3.051	0.004	0.007	0.611
Mean Left HbO	-0.55±3.29 -0.32 (-9.51–4.88)	1.6±2.8 1.74 (-5.13–6.9)	-3.426	0.001	0.003	-0.686
Mean Left HbR	0.67±2.72 0.47 (-5.39–9.53)	0.31±2.44 0.47 (-4.25–8.82)	0.824	0.414	0.414	0.165
Mean Right HbO	0.47±2.61 0.23 (-7–4.12)	1.62±3.13 1.57 (-6.16–7.39)	-2.108	0.043	0.064	-0.423
Mean Right HbR	0.68±4.49 -0.02 (-2.73–20.32)	-0.13±3.05 -0.35 (-5.62–6.53)	1.462	0.152	0.228	0.293

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.10
Robotic Surgery Real Environment: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between PP and KT tasks.

	PP Task	KT Task	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	85.74±56.34 79.41 (23.58–249.2)	61.93±37.63 51.63 (23.66–215.99)	1.858	0.075	0.187	0.372
Mean Left HbO	0.36±3.02 0.01 (-6.74–5.85)	-0.86±3.61 -0.49 (-7.4–7.58)	1.876	0.073	0.187	0.383
Mean Left HbR	0.51±2.4 0.25 (-4.36–7.07)	-1.02±5.95 -0.28 (-20.81–9.11)	0.820	0.421	0.590	0.167
Mean Right HbO	1.47±2.36 1.71 (-3.33–5.04)	-1.08±3.71 0.16 (-11.48–4.38)	3.304	**0.003**	**0.023**	0.674
Mean Right HbR	-0.34±3.1 -0.73 (-6.78–7.31)	-0.75±4.96 -0.51 (-18.98–10.42)	0.538	0.596	0.745	0.110

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.11
Real PP Task: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between laparoscopic and robotic surgery.

	Laparoscopy	Robotic	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	124.01±108.92 82.59 (25.59–492.3)	85.74±56.34 79.41 (23.58–249.2)	1.553	0.133	0.222	0.311
Mean Left HbO	2.48±2.42 2.38 (-0.45–6.93)	0.36±3.02 0.01 (-6.74–5.85)	2.290	0.032	0.192	0.467
Mean Left HbR	0.29±1.94 0.35 (-5.67–2.53)	0.51±2.4 0.25 (-4.36–7.07)	-0.363	0.720	0.720	-0.074
Mean Right HbO	2.82±3.09 3.07 (-2.82–10.27)	1.47±2.36 1.71 (-3.33–5.04)	1.634	0.115	0.216	0.327
Mean Right HbR	1.05±3.53 0.04 (-4.63–13.25)	-0.34±3.1 -0.73 (-6.78–7.31)	1.170	0.254	0.381	0.234

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.12
PP Task in Simulation: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between laparoscopic and robotic surgery.

	Laparoscopy	Robotic	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	84.01±52.16 62.76 (32.82–251.81)	69.5±23.74 68.79 (37.23–144.6)	1.268	0.217	0.670	0.254
Mean Left HbO	0.19±2.15 -0.15 (-2.94–6.28)	-0.55±3.29 -0.32 (-9.51–4.88)	0.940	0.357	0.670	0.188
Mean Left HbR	0.07±1.79 0.49 (-6.52–2.25)	0.67±2.72 0.47 (-5.39–9.53)	-0.928	0.363	0.670	-0.186
Mean Right HbO	1.04±2.00 0.88 (-2.34–5.97)	0.47±2.61 0.23 (-7–4.12)	0.768	0.450	0.675	0.154
Mean Right HbR	-0.76±2.23 -1.39 (-5.73–4.4)	0.68±4.49 -0.02 (-2.73–20.32)	-1.552	0.134	0.670	-0.310

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.13
Real KT Task: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between laparoscopic and robotic surgery.

	Laparoscopy	Robotic	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	99.31±95.19 65.61 (17.19–414.68)	61.93±37.63 51.63 (23.66–215.99)	1.915	0.067	0.101	0.383
Mean Left HbO	4.12±3.31 4.16 (-1.44–10.86)	-0.86±3.61 -0.49 (-7.4–7.58)	5.056	<0.001	0.003	1.011
Mean Left HbR	0.23±2.04 0.8 (-4.04–3.41)	-1.02±5.95 -0.28 (-20.81–9.11)	0.994	0.330	0.381	0.199
Mean Right HbO	4.98±2.73 5.1 (-0.82–11.93)	-1.08±3.71 0.16 (-11.48–4.38)	6.194	<0.001	0.003	1.264
Mean Right HbR	0.55±3.32 0.22 (-4.09–12.92)	-0.75±4.96 -0.51 (-18.98–10.42)	1.001	0.327	0.381	0.204

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

Table 3.14
Simulation KT Task: Comparison of task completion time and prefrontal hemodynamic responses (HbO/HbR) between laparoscopic and robotic surgery.

	Laparoscopy	Robotic	t	p	p*	Effect size
	Mean±SD	Mean±SD				
Completion time	50.5±16.68 46.83 (24.39–88.83)	104.13±34.66 96.33 (57.66–208.01)	-6.861	<0.001	0.003	-1.372
Mean Left HbO	2.8±2.57 2.69 (-1.22–8.23)	-1.6±3.73 -2.16 (-11.67–7.38)	4.960	<0.001	0.003	0.992
Mean Left HbR	0.45±1.94 1.06 (-4.73–4.53)	-0.26±4.56 -0.36 (-13.66–7.75)	0.713	0.483	0.603	0.143
Mean Right HbO	3.7±3.06 2.62 (-1.54–12.97)	-1.25±4.56 -0.67 (-12.7–6.05)	4.612	<0.001	0.003	0.941
Mean Right HbR	0.69±3.37 -0.22 (-4.46–13.12)	-2.14±3.97 -1.38 (-17.71–2.37)	2.494	0.020	0.030	0.509

*Reported p-values are corrected using the Benjamini–Hochberg False Discovery Rate (FDR) method.

3.1 Completion Time Analysis

Completion times varied across surgical modalities, task complexity, and training environments. Performance differences were evident when comparing simulated and real environments across both laparoscopic and robotic modalities. As illustrated in Figure 3.1, For the laparoscopy PP task, the median completion time increased substantially from 93.3 seconds [85.0 - 102.6] in simulation to 125.2 seconds [113.4 - 143.7] in the real environment. Similarly, for the laparoscopy KT task, the median time rose from 49.0 seconds [45.2 - 70.6] in simulation to 99.2 seconds [89.4 - 115.0] in the real setting, highlighting the increased task complexity and cognitive burden experienced in authentic surgical conditions. In the robotic PP task, the median time increased from 71.3 seconds [61.2 - 83.6] in simulation to 90.8 seconds [84.2 - 104.8] in real operations, indicating that even for simpler tasks, real world settings pose additional demands. Interestingly, a contrasting trend was observed in the robotic KT task, where the median completion time dropped from 101.9 seconds [90.6 - 126.4] in simulation to 66.3 seconds [60.3 - 74.1] in the real setting. This decrease may be attributed to improved performance facilitated by the ergonomic and visual advantages of the robotic platform, and potentially the effects

of task familiarity developed during simulation.

These findings are further supported by the linear mixed effects model results, which demonstrated that session based grouping better explained completion time variations than the independent effects of modality, environment, and task complexity. Notably, the reduction in task duration for robotic real KT tasks suggests a positive transfer of skills from simulated practice to real world application, underlining the efficacy of structured simulation protocols. However, the general trend of longer completion times in real laparoscopic tasks implies that simulation alone may not sufficiently replicate the mental and physical demands of actual surgical environments. This performance gap reinforces the need for enhanced simulation fidelity and supplemental real environment exposure to fully prepare surgical trainees. Overall, these observations emphasize the combined influence of surgical modality, task complexity, and training context on operative efficiency, further validating the inclusion of neuroergonomic insights in curriculum design (Figure 3.1).

Figure 3.1 supports these findings, with wider interquartile ranges and more outliers in real conditions, particularly for laparoscopic tasks, indicating greater variability and cognitive demand. In contrast, robotic tasks demonstrated tighter distributions and lower mean durations, emphasizing their efficiency and consistency, especially under complex conditions.

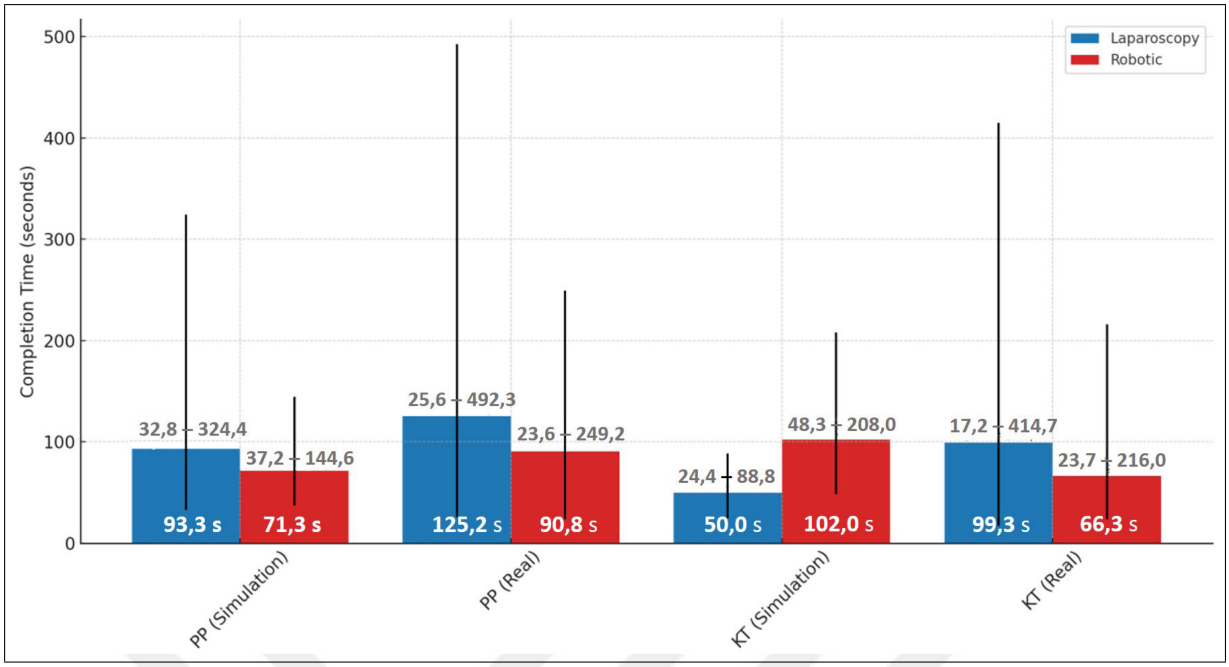

Figure 3.1 Median completion times with min-max ranges for laparoscopic and robotic tasks in simulation and real environments.

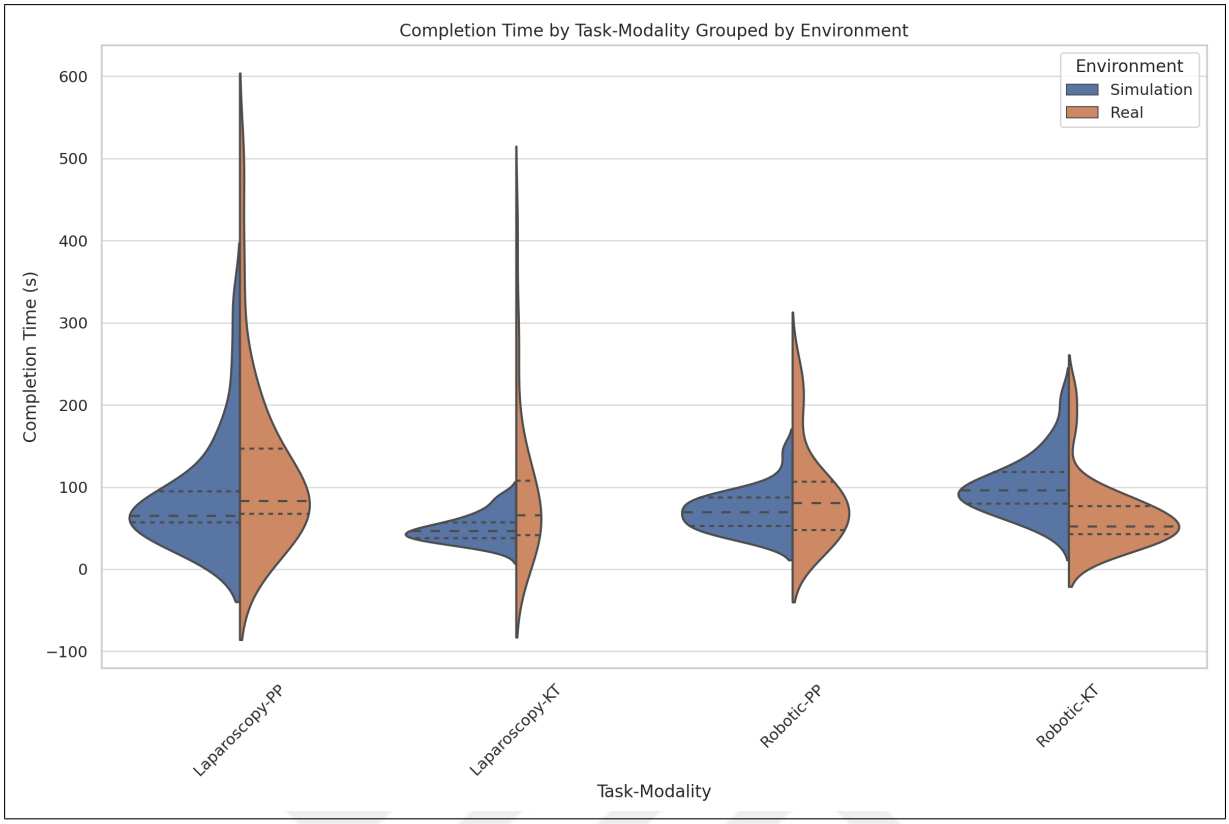

Figure 3.2 Violin plot showing completion time distributions across Task Modality combinations (Lap-PP, Lap-KT, Robot-PP, Robot-KT), grouped by Environment (Simulation vs. Real). Quartile lines indicate median and interquartile range. The violin shape reveals underlying distribution spread for each group.

To further explore the variability and distributional characteristics of completion times across different task modality combinations and environments, a violin plot was created (Figure 3.2). Unlike boxplots, violin plots illustrate the full probability density of the data, offering insights into the spread and symmetry of performance times. Each violin represents a unique task modality pairing (e.g., Lap-PP, Lap-KT), with distributions split by environment (Simulation vs. Real). This visualization revealed that real laparoscopic tasks (both PP and KT) exhibited broader and more skewed distributions, indicating high variability among participants. In contrast, robotic tasks,especially those performed in simulation,displayed more compact and symmetric distributions, reflecting more consistent performance.

3.2 Modality Effects

Performance and cognitive workload varied substantially between laparoscopic and robotic surgical modalities across both simulated and real environments. These differences are visualized in below figures, which compare completion time and fNIRS derived prefrontal cortex activation metrics across task types and environments.

In the real world setting, robotic surgery demonstrated superior performance efficiency compared to laparoscopy across both task types. According to Figure 3.1, median completion times for laparoscopic tasks were consistently longer than those for robotic procedures. For the pick and place task, the median time increased from 93.3 seconds [85.0-102.6] in simulation to 125.2 seconds [113.4-143.7] in the real environment, whereas robotic PP tasks saw a smaller increase, rising from 71.3 [61.2-83.6] to 90.8 seconds [84.2-104.8]. The knot tying task presented a more striking contrast: laparoscopic KT times escalated from 49.0 seconds to 99.2 seconds [89.4-115.0], while robotic KT completion time decreased significantly from 101.9 [90.6-126.4] to 66.3 seconds [60.3-74.1]. This reversal in performance likely reflects a transfer of learning from simulation to real settings, particularly for robotic KT, supported by post hoc results in Table 3.2 showing a significant time reduction ($\Delta = -3.95$ s, $p = 0.001$, $d = 1.844$). These findings underscore the efficiency and ergonomic advantage of the robotic platform under realistic surgical demands. Figure 3.3 presents the mean completion times for all tasks performed in the real environment, highlighting consistent advantages for the robotic modality.

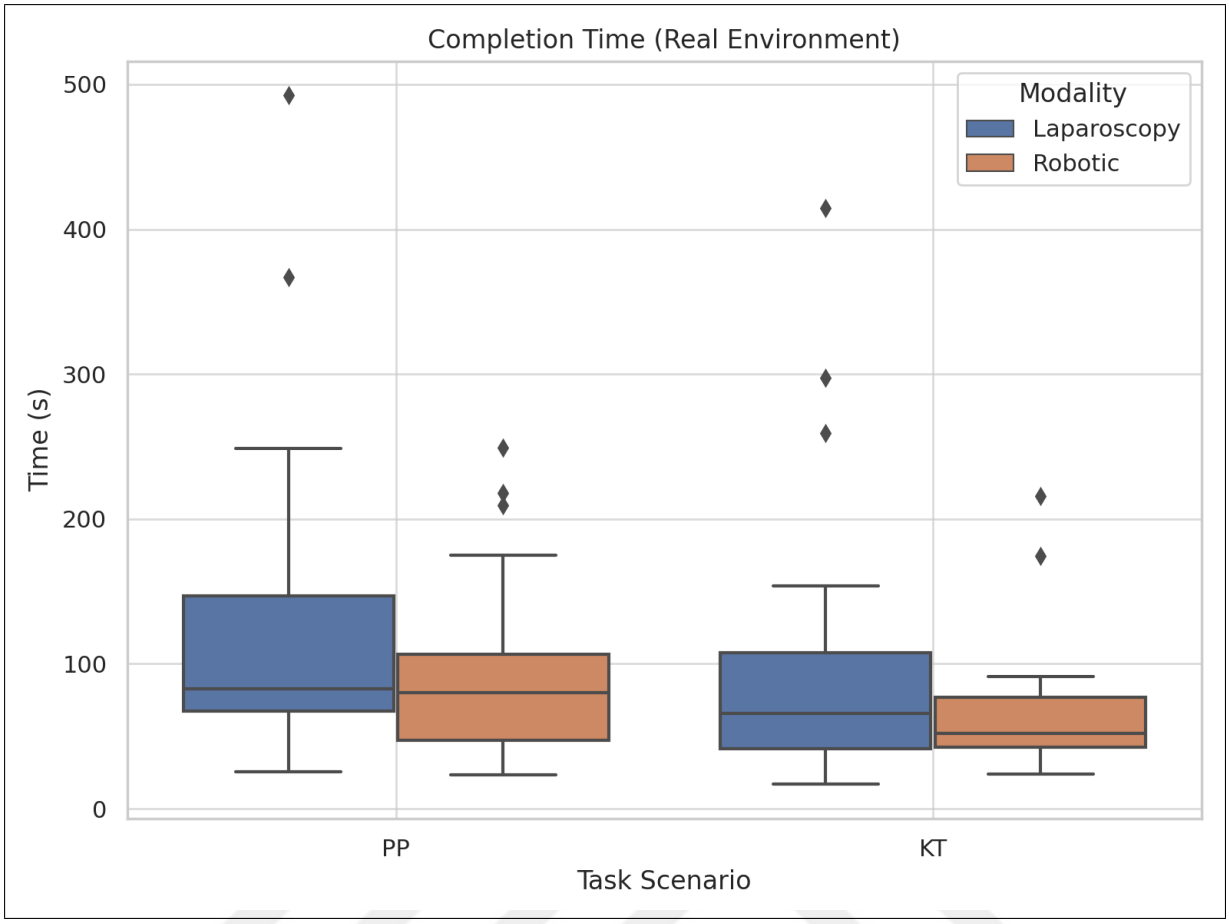

Figure 3.3 Completion Time in Real Environment.

A similar trend was observed in the simulation environment, with a slightly lower magnitude compared to the real setting. In the pick and place task, laparoscopic surgery required a median completion time of 93.3 seconds [85.0-102.6], whereas robotic surgery was completed faster at 71.3 seconds [61.2-83.6]. In the knot tying task, laparoscopic surgery was completed in a median of 49.0 seconds [45.2-70.6], while robotic KT required a longer median time of 101.9 seconds [90.6-126.4].

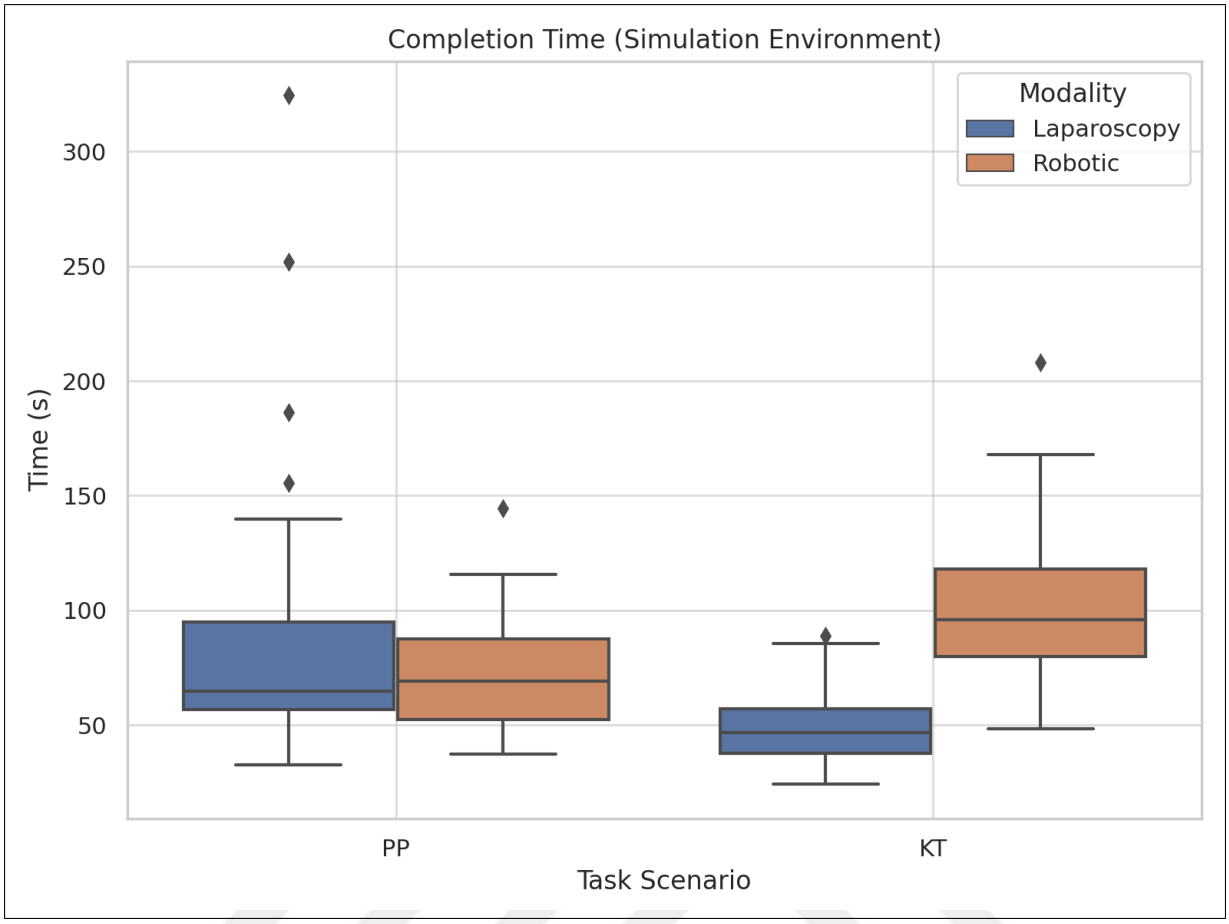

Figure 3.4 Completion Time in Simulation Environment.

In the simulation knot tying phase, there is a significant difference in mean left HbO levels between laparoscopy (4.12 ± 3.31) and robotic (-0.86 ± 3.61), with a t-value of 5.056 and a p-value < 0.001 (Bonferroni corrected $p^* = 0.003$). The effect size is 1.011, indicating a large effect. This suggests that laparoscopy significantly increased oxygenation levels in the left hemisphere during simulation based knot tying tasks compared to the robotic platform. This might imply better cortical activation or reduced cognitive load during Laparoscopy, potentially linked to familiarity or procedural efficiency.

A significant difference is also observed during the real KT phase, with laparoscopy (2.8 ± 2.57) showing higher oxygenation compared to Robotic (-1.6 ± 3.73), with a t-value of 4.960 and $p < 0.001$ (Bonferroni corrected $p^* = 0.003$). The effect size is 0.992, reflecting a large impact. This indicates that the laparoscopy modality enhances left

hemisphere oxygenation during real knot tying, suggesting improved cognitive processing or motor control under this modality.

There is also a statistically significant difference during the real pick and place task, with Laparoscopy (2.48 ± 2.42) showing higher oxygenation compared to robotic (0.36 ± 3.02), with a t-value of 2.290 and $p = 0.032$ (Bonferroni corrected $p^* = 0.192$, not significant after correction). The effect size is 0.467, suggesting a medium effect, although the significance is not preserved after correction. This indicates a trend toward higher activation in the left hemisphere during real PP under laparoscopy, but the corrected p-value suggests caution in interpretation.

These findings are illustrated in Figure 3.5, which shows the comparison of mean left HbO between laparoscopy and robotic modalities during simulation and real tasks for both knot tying and pick and place.

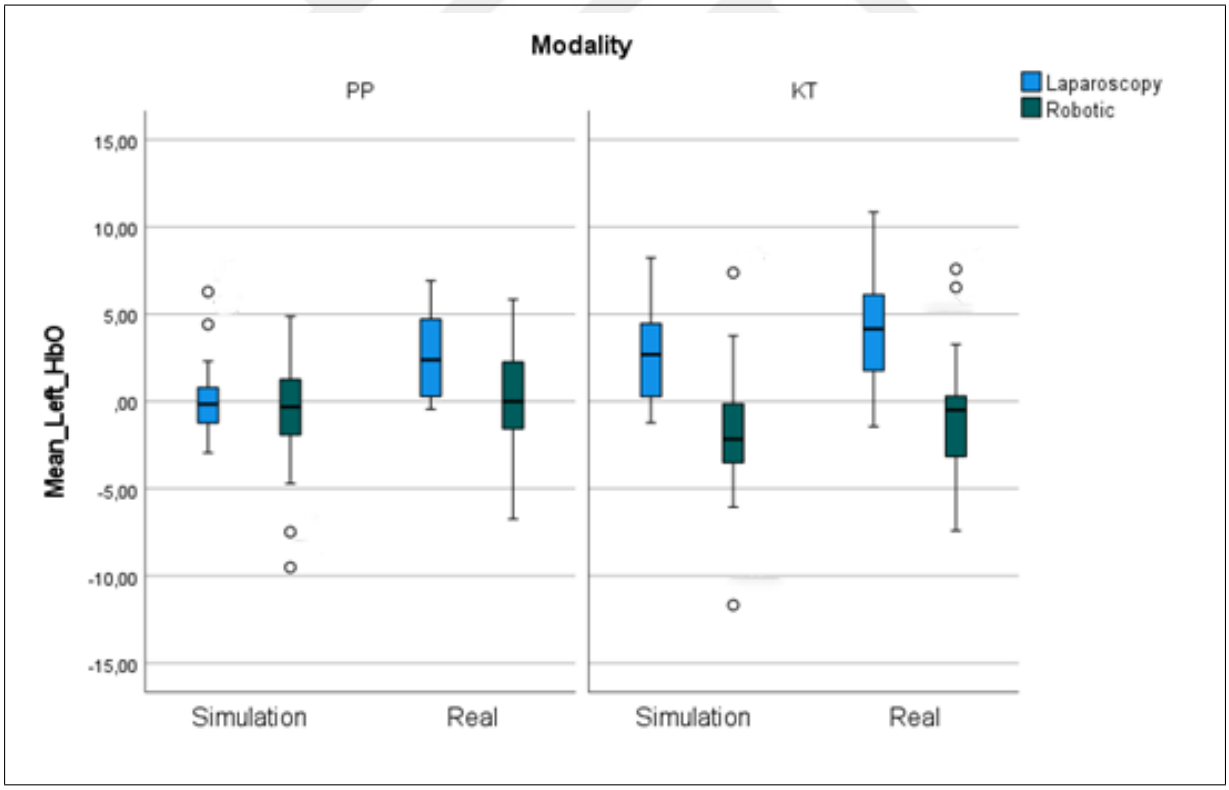

Figure 3.5 Comparison of Mean Left HbO between Laparoscopy and Robotic modalities during Simulation and Real tasks for both knot tying and pick and place.

There is also a significant difference for mean right HbO during simulation knot tying , with Laparoscopy (4.98 ± 2.73) showing much higher oxygenation levels compared to robotic (-1.08 ± 3.71), with a t-value of 6.194, $p < 0.001$ (Bonferroni corrected $p^* = 0.003$), and an effect size of 1.264. This is a very strong effect, indicating higher activation in the right prefrontal cortex during laparoscopy, possibly reflecting greater spatial awareness or motor control when compared to the robotic platform.

Similarly, mean right HbO showed significant differences in the Real KT phase: Laparoscopy (3.7 ± 3.06) versus Robotic (-1.25 ± 4.56), with a t-value of 4.612, $p < 0.001$ (Bonferroni corrected $p^* = 0.003$), and an effect size of 0.941. This reinforces that laparoscopy significantly improves right hemisphere oxygenation during real tasks compared to the robotic method, potentially due to more intuitive hand eye coordination and direct manipulation.

These findings are illustrated in Figure 3.6, which shows the comparison of mean right HbO between laparoscopy and robotic modalities during simulation and real tasks for both knot tying and pick and place.

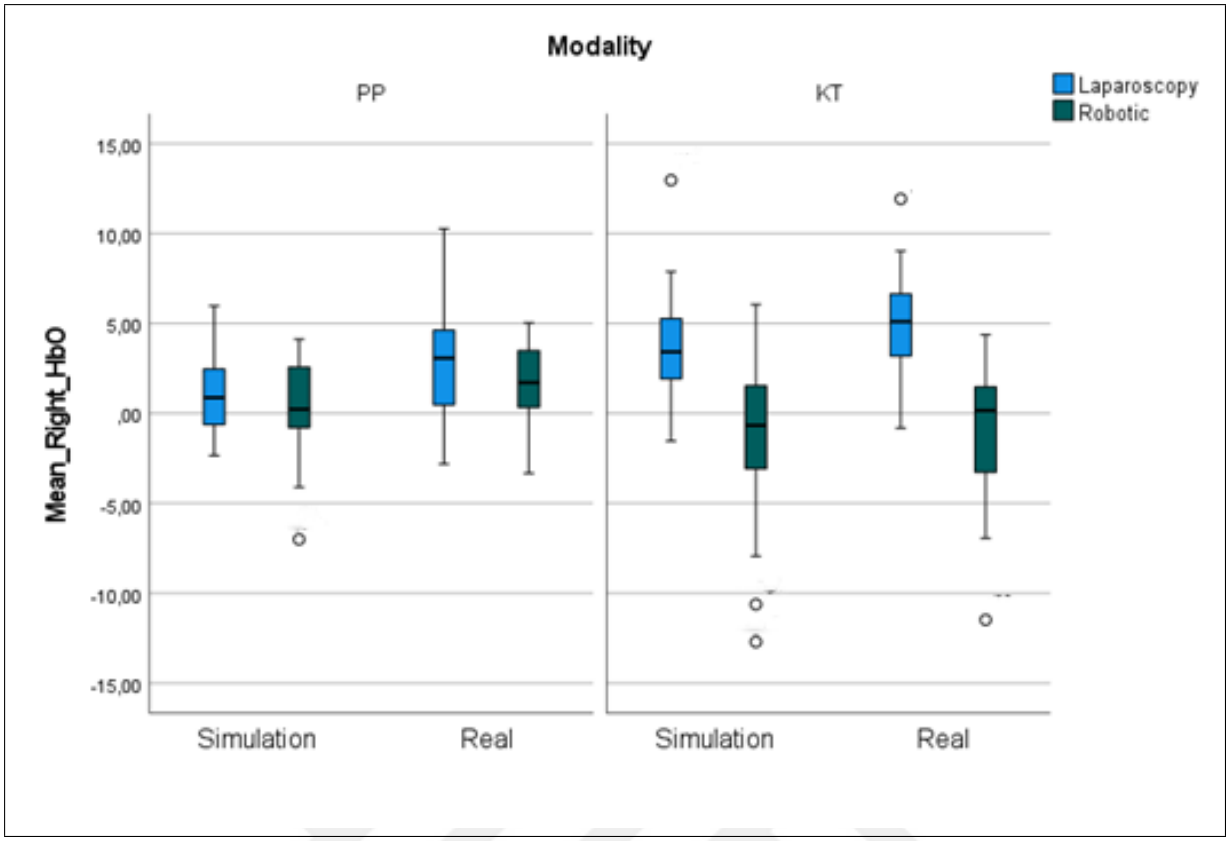

Figure 3.6 Comparison of Mean Right HbO between Laparoscopy and Robotic modalities during Simulation and Real tasks for both knot tying and pick and place.

Mean right HbR showed a significant difference: Laparoscopy (0.69 ± 3.37) versus Robotic (-2.14 ± 3.97), with a t-value of 2.494, p = 0.020 (Bonferroni corrected p* = 0.030), and an effect size of 0.509. The negative shift in HbR during the robotic procedure suggests a relative reduction in deoxygenated hemoglobin, indicating decreased right hemisphere activation or altered neurovascular coupling in robotic manipulation. This may imply cognitive strain or different neural engagement when using robotic tools compared to the manual approach. There were no significant differences observed in mean left HbR between the modalities.

These results are summarized in Figure 3.7, which shows the comparison of Mean Right HbR between laparoscopy and robotic modalities during simulation and real tasks for both knot tying and pick and place.

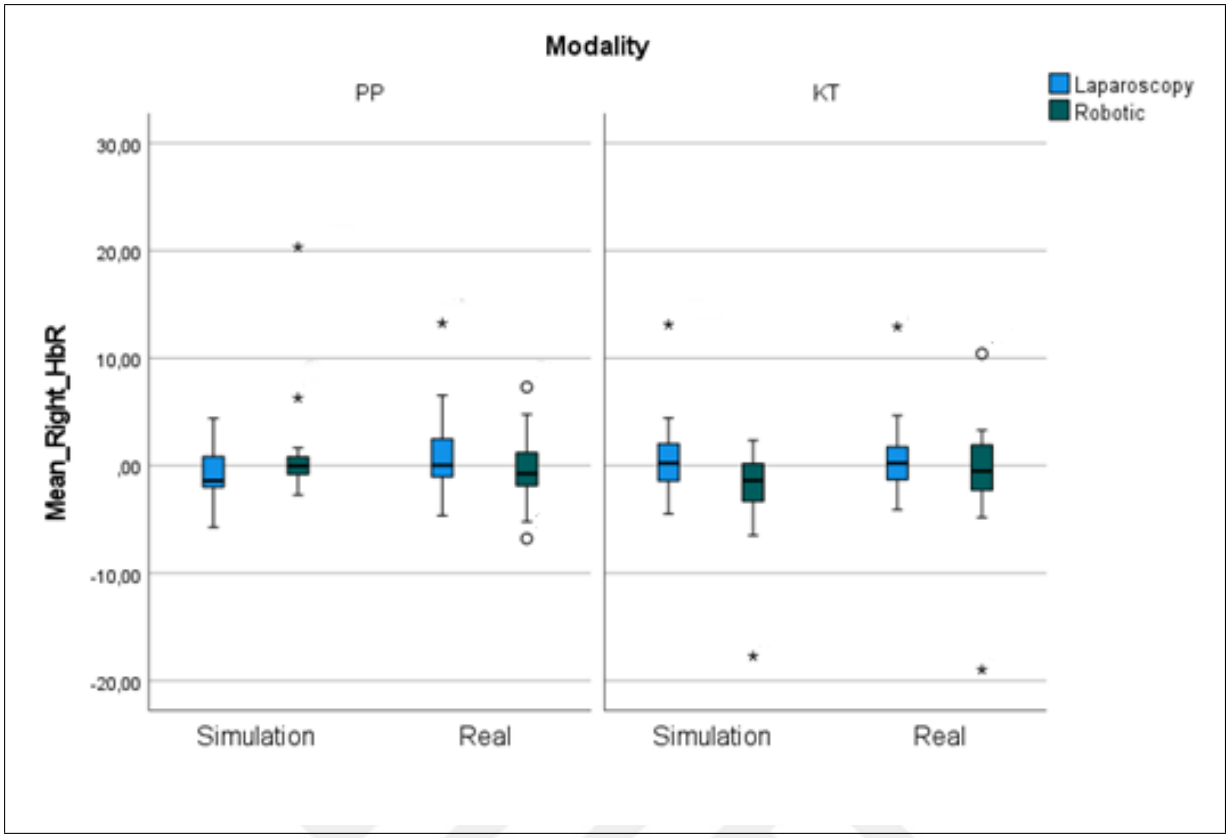

Figure 3.7 Comparison of Mean Right HbR between Laparoscopy and Robotic modalities during Simulation and Real tasks for both knot tying and pick and place.

These findings indicate that during simulation, laparoscopic surgery, particularly knot tying, imposed significantly higher cognitive demands and was performed more efficiently in terms of time compared to robotic surgery.

3.3 Task Complexity

To evaluate the impact of task complexity on surgical performance and cognitive workload, we compared the pick and place task, representing a lower complexity procedure with the knot tying task, which is cognitively and technically more demanding. Analyses were conducted across both laparoscopic and robotic modalities in both simulation based and real life environments. Completion time was used to assess performance, while cognitive workload was evaluated via hemodynamic responses (HbO and HbR) in

the left and right prefrontal cortices as measured by fNIRS.

3.3.1 Completion Time

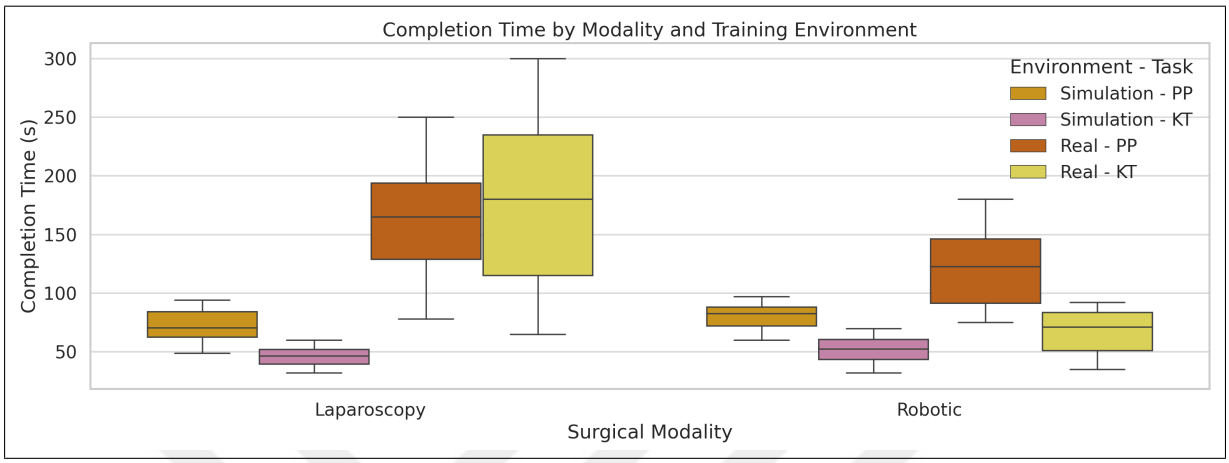

Figure 3.8 Completion Time by Task Complexity and Modality.

Analysis of completion times revealed distinct patterns associated with task complexity across surgical modalities and environments.

In laparoscopic simulations, paradoxically, the more complex KT task was completed faster than the PP task, with median times of 49.0 seconds [45.2-70.6] for KT and 93.3 seconds [85.0-102.6] for PP. However, in the real laparoscopic environment, this pattern shifted: KT took longer than in simulation (99.2 seconds [89.4-115.0]), and PP completion time also increased substantially (125.2 seconds [113.4-143.7]). Despite these increases, no statistically significant difference between KT and PP times was observed in real laparoscopy .

For robotic surgery, a more intuitive trend emerged. In robotic simulations, KT required longer times than PP (101.9 seconds [90.6-126.4] vs. 71.3 seconds [61.2-83.6], respectively), consistent with the expected increase in duration for more complex tasks. Interestingly, in the real robotic environment, the KT task completion time decreased significantly compared to simulation, dropping to 66.3 seconds [60.3-74.1], while PP in-

creased slightly to 90.8 seconds [84.2-104.8]. This reversal in KT performance suggests a transfer of learning and highlights the ergonomic and visual advantages offered by the robotic platform under real surgical demands .

Overall, these results underscore that task complexity influenced completion time differently depending on the surgical modality and environment. Robotic systems, particularly for complex tasks like KT, demonstrated a significant efficiency advantage in real settings compared to simulation, whereas laparoscopic performance showed a general trend of prolonged task times under real world conditions.

3.3.2 fNIRS Based Cognitive Workload

To assess the effect of task complexity pick and place vs. knot tying on cognitive workload, changes in oxygenated hemoglobin (ΔHbO) measured by fNIRS were compared within each modality and environment condition.

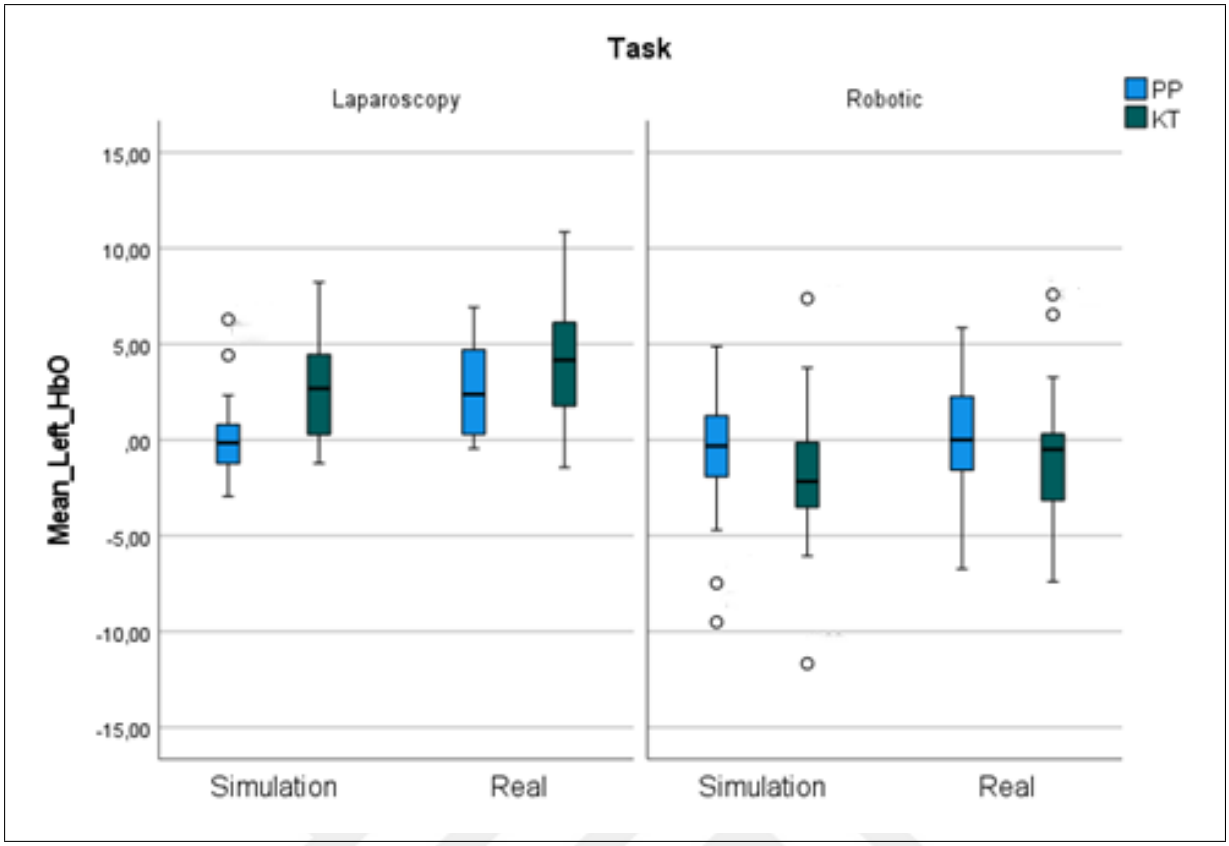

Figure 3.9 Comparison of mean left HbO between knot tying and pick and place tasks during simulation and real settings in both laparoscopy and robotic modalities.

In the simulation setting, mean left HbO was significantly higher in KT (2.8±2.57) compared to PP (0.19 ± 2.15) with a t-value of -5.420, $p < 0.001$, and a corrected $p^* < 0.001$. The effect size of -1.084 indicates a large effect.

This finding suggests that laparoscopy based KT tasks evoke greater left hemisphere oxygenation, reflecting heightened cortical activity and cognitive load associated with knot tying under manual control (see Figure 3.9).

In real surgical settings, mean left HbO was significantly higher during KT (4.12± 3.31) compared to PP (2.48 ± 2.42) with a t-value of -3.876, $p = 0.001$, and corrected $p^* = 0.003$. The effect size of -0.775 represents a strong effect.

This indicates that real laparoscopic KT tasks demand more prefrontal activa-

tion, potentially due to the fine motor control and spatial awareness required in real environments (see Figure 3.9).

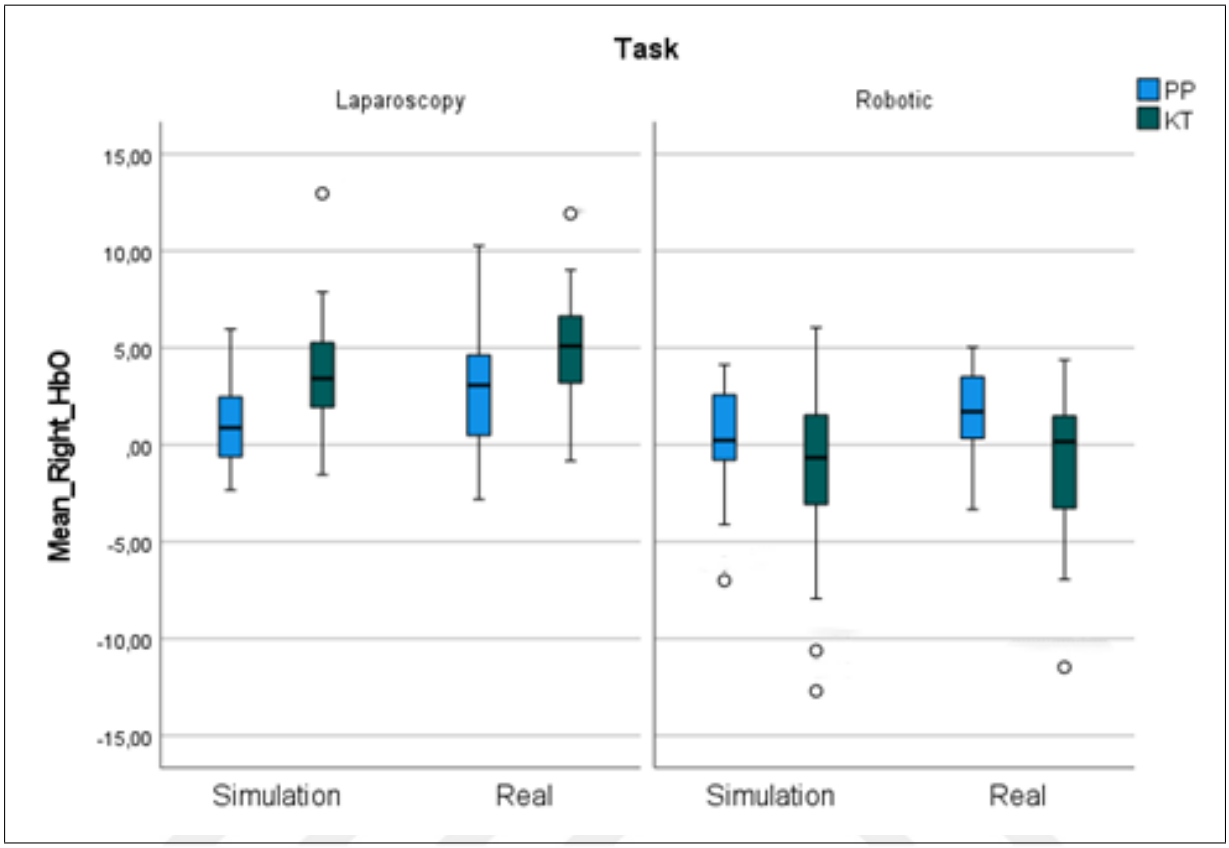

Figure 3.10 Comparison of mean right HbO between knot tying and pick and place tasks during simulation and real settings in both laparoscopy and robotic modalities.

Similarly, mean right HbO was significantly higher during KT (3.7 ± 3.06) compared to PP (1.04 ± 2.00) in the simulation setting, with a t-value of -4.481, $p < 0.001$, and a corrected $p^* < 0.001$. The effect size of -0.896 indicates a substantial impact.

This demonstrates that the right prefrontal cortex is more active during KT, likely due to its role in visuospatial processing and motor control in laparoscopic procedures (see Figure 3.10).

A significant difference was also observed in mean right HbO during real settings, where KT (4.98 ± 2.73) showed higher activation compared to PP (2.82 ± 3.09) with a t-value of -4.546, $p < 0.001$, and corrected $p^* = 0.001$. The effect size of -0.909 suggests

a strong neural activation.

This may imply that KT under real laparoscopic conditions requires substantial visuospatial processing and motor coordination, activating the right prefrontal cortex more intensively (see Figure 3.10).

In the robotic modality, a significant difference in mean right HbO was found between PP (1.47 ± 2.36) and KT (-1.08 ± 3.71), with a t-value of 3.304, $p = 0.003$, and a corrected $p^* = 0.023$. The effect size of 0.674 indicates a moderate to strong effect.

This result suggests that KT tasks under robotic assistance reduce right hemisphere oxygenation compared to PP, possibly due to the assistance provided by the robotic interface, reducing the need for intense right hemisphere activation (see Figure 3.11).

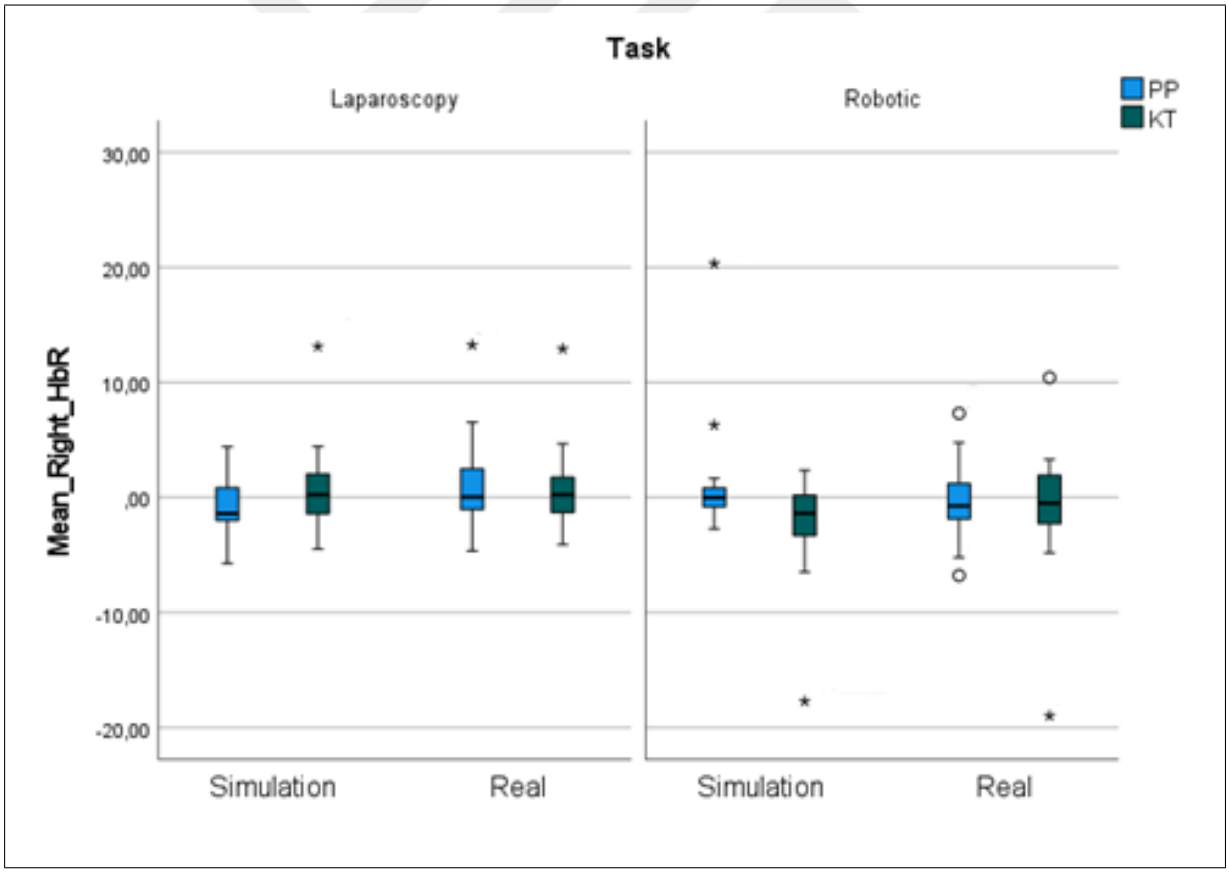

Figure 3.11 Comparison of mean right HbR between knot tying and pick and place tasks during simulation and real settings in both laparoscopy and robotic modalities.

A significant difference was observed in mean right HbR between PP (-0.76 ± 2.23) and KT (0.69 ± 3.37) in the simulation setting, with a t-value of -2.663, $p = 0.014$, and a corrected $p^* = 0.023$. The effect size of -0.533 suggests a moderate effect.

The higher HbR levels in KT suggest a greater demand for oxygenated blood, reflecting more intensive neural engagement during knot tying (see Figure 3.11).

A significant difference was also noted for Mean Right HbR in the real setting, where PP (0.68 ± 4.49) was higher than KT (-2.14 ± 3.97), with a t-value of 2.223, $p = 0.036$, and a corrected $p^* = 0.135$ (not significant after correction). The effect size of 0.454 indicates a moderate effect.

Although not significant after Bonferroni correction, the trend indicates that KT under robotic assistance may reduce deoxygenated hemoglobin, suggesting that different cognitive or motor strategies are employed during robotic knot tying (see Figure 3.11).

3.4 Environment (Simulation vs Real)

The effects of training environment on surgical performance and cognitive workload. By comparing simulation based tasks with their real world counterparts, we aimed to determine whether simulated training environments elicit comparable behavioral and neurophysiological responses. Both completion time and fNIRS derived measures of prefrontal cortex activity were analyzed to evaluate differences in cognitive demand between simulation and real settings.

3.4.1 Completion Time

The training environment had a significant effect on surgical performance as measured by completion time.

For **laparoscopic surgery**, completion times increased when transitioning from simulation to the real environment. In the PP task, the median time rose from 93.3 seconds [85.0–102.6] in simulation to 125.2 seconds [113.4–143.7] in real settings. Similarly, for the KT task, completion times increased from 49.0 seconds [45.2–70.6] to 99.2 seconds [89.4–115.0]. The linear mixed effects model indicated a significant main effect of environment, particularly for the KT task, suggesting that the real environment imposes greater cognitive and motor demands.

In **robotic surgery**, the trends differed. For the PP task, completion times increased slightly from 71.3 seconds [61.2–83.6] in simulation to 90.8 seconds [84.2–104.8] in real operations, although this difference was not statistically significant. In contrast, for the KT task, a decrease in completion time was observed, from 101.9 seconds [90.6–126.4] in simulation to 66.3 seconds [60.3–74.1] in the real environment. The LME model identified a significant interaction between environment and task type, indicating that robotic systems may better facilitate complex procedural performance in real settings compared to simulations. Consistent with previous findings, simulation based robotic training environments have been associated with lower cognitive workload and greater neural efficiency compared to laparoscopic simulations. In a cohort of surgical residents with no prior robotic experience, robotic simulators elicited lower Mean ΔHbO levels and shorter task completion times, underscoring the cognitive advantages conferred by the ergonomic and visual enhancements of robotic platforms even in purely simulated settings [39].

Overall, these findings suggest that while the real environment generally increases task difficulty for laparoscopic surgery, robotic surgery may mitigate or even reverse this effect, particularly for more complex tasks such as knot tying. These results are visualized in Figure 3.12.

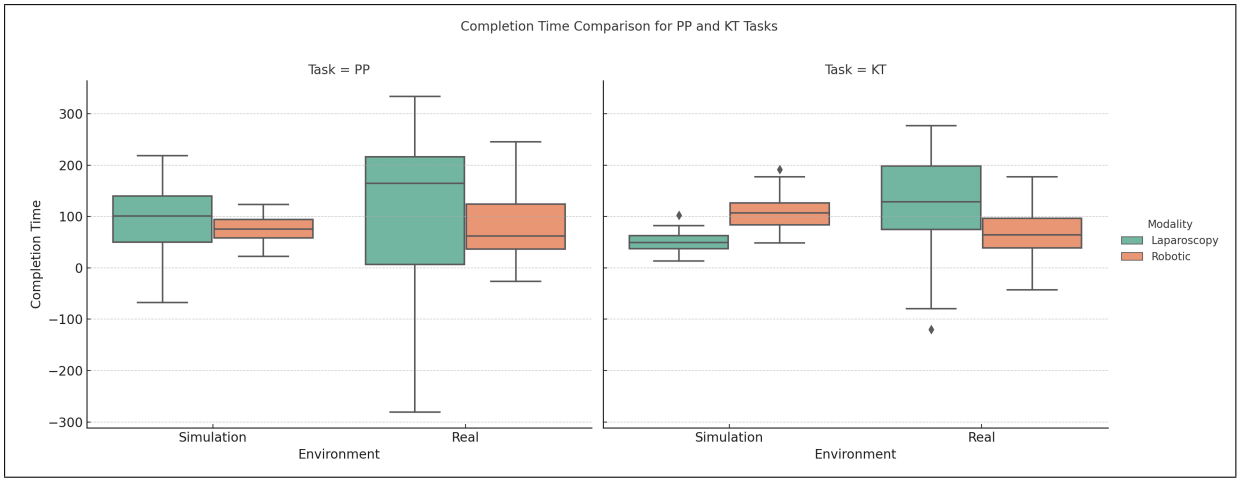

Figure 3.12 Completion Time by environment (Simulation vs. Real).

3.4.2 fNIRS based Cognitive Workload

When comparing simulation based and real world environments, significant differences in cognitive workload measured via functional near infrared spectroscopy (fNIRS) were evident predominantly within laparoscopic surgery tasks. A direct comparison of simulation and real world environments for each surgical modality and task revealed notable differences in prefrontal activation, as measured by Mean ΔHbO values.

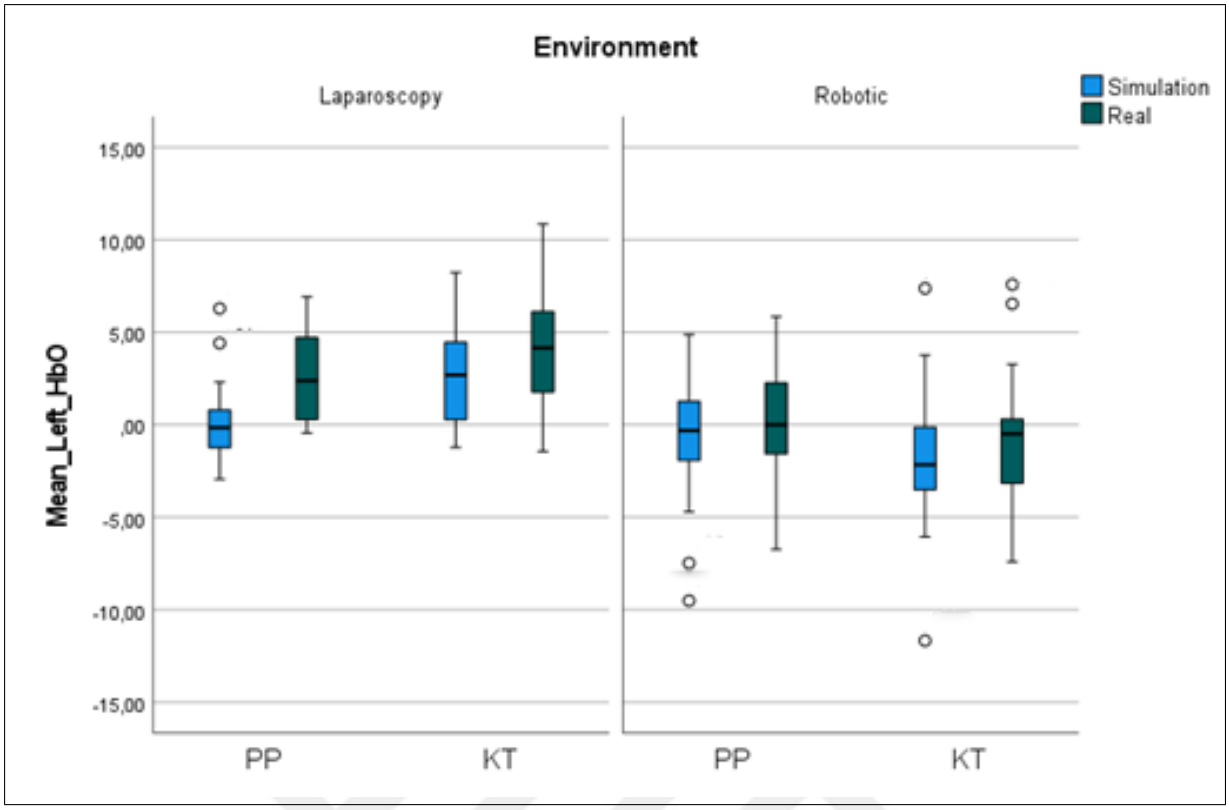

Figure 3.13 Comparison of mean left HbO between simulation and real settings during pick and place tasks in the laparoscopy and robotic modalities.

There was a significant increase in mean left HbO from simulation (0.19 ± 2.15) to Real (2.48 ± 2.42) in the pick and place (PP) task, with a t-value of -5.103, $p < 0.001$, and a corrected $p^* = 0.001$. The effect size of -1.021 suggests a large effect.

This result indicates a substantial increase in oxygenated hemoglobin in the left prefrontal cortex during real PP in laparoscopy, reflecting greater cortical engagement likely due to the higher cognitive and motor demands of real world manipulation (see Figure 3.13).

Similarly, there was a significant increase in mean left HbO from simulation (2.8 ± 2.57) to Real (4.12 ± 3.31) in the knot tying (KT) task, with a t-value of -3.476, $p = 0.002$, and a corrected $p^* = 0.010$. The effect size of -0.695 indicates a moderate-to-large effect.

This result suggests heightened activation in the left prefrontal cortex during real

KT tasks, reflecting the complex hand eye coordination and decision making required in live scenarios (see Figure 3.13).

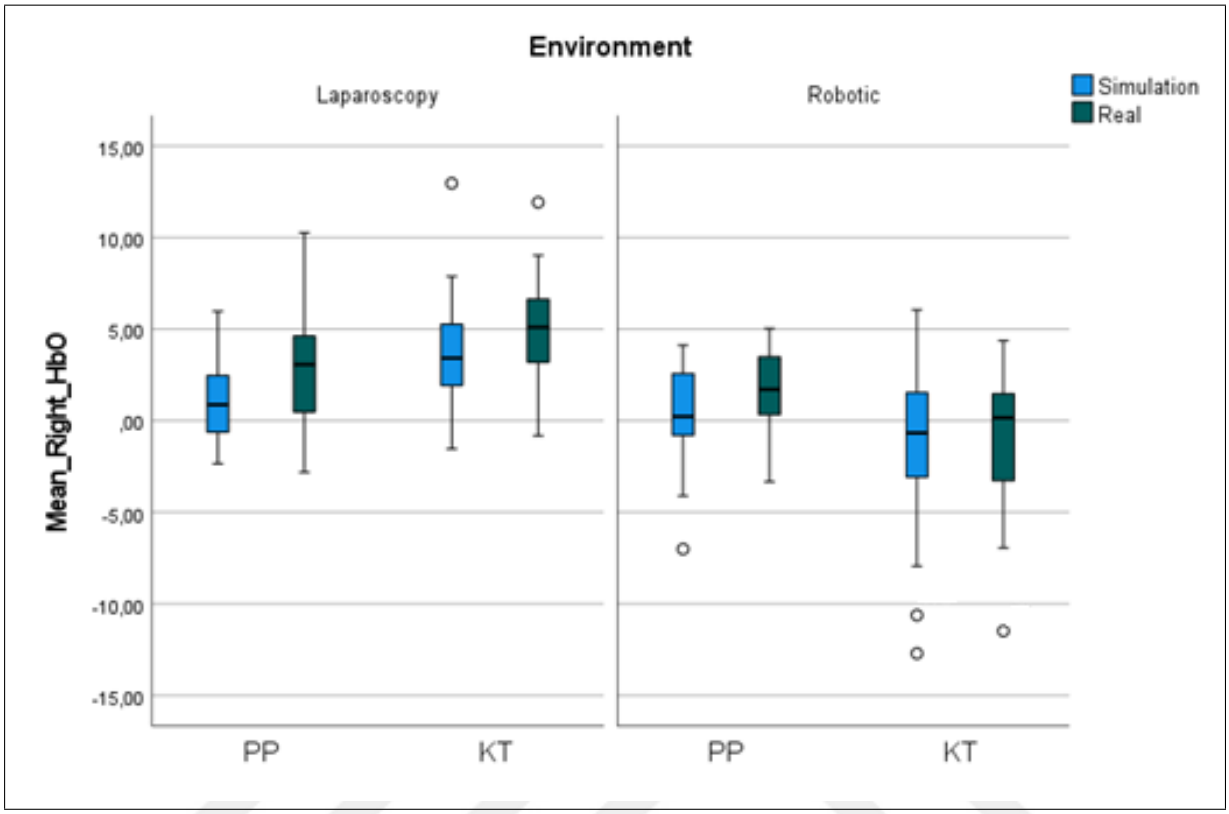

Figure 3.14 Comparison of mean right HbO between simulation and real settings during pick and place tasks in the laparoscopy and robotic modalities.

A significant increase was also observed in mean right HbO from simulation (1.04± 2.00) to Real (2.82 ± 3.09) in the pick and place (PP) task, yielding a t-value of -3.126, $p = 0.005$, and a corrected $p^* = 0.015$. The effect size of -0.625 suggests a moderate to large effect.

This indicates that real PP tasks activate the right hemisphere significantly more than in simulations, possibly due to enhanced visuospatial and motor processing (see Figure 3.14).

A similar pattern was observed for mean right HbO in the knot tying task, with simulation (3.7 ± 3.06) compared to real (4.98 ± 2.73), producing a t-value of -3.558,

$p = 0.002$, and a corrected $p^* = 0.010$. The effect size of -0.712 indicates a moderate to large impact.

This suggests that right hemisphere activation is notably elevated during real KT, reflecting enhanced spatial awareness and precision required in real settings compared to simulation (see Figure 3.14).

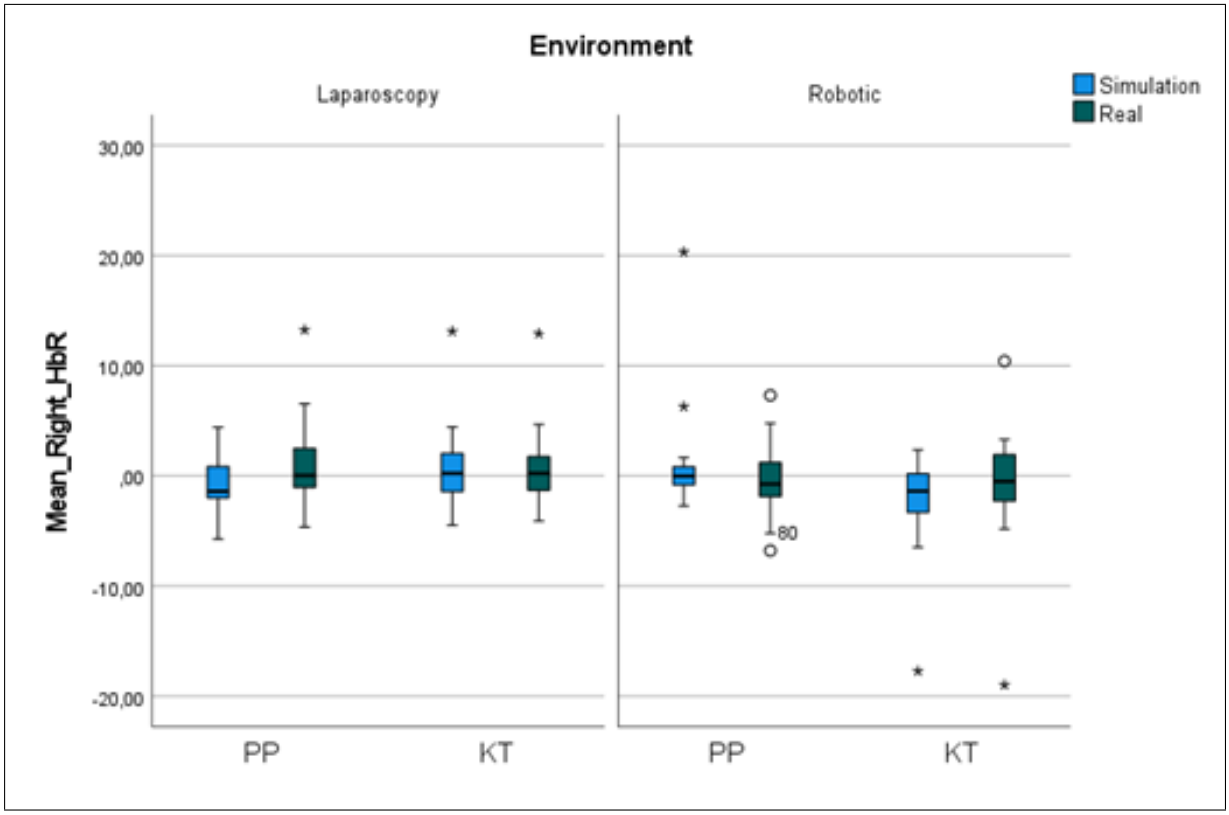

Figure 3.15 Comparison of mean right HbR between simulation and real settings during pick and place tasks in the laparoscopy and robotic modalities.

Mean Right HbR levels increased significantly from simulation (-0.76 ± 2.23) to Real (1.05 ± 3.53) in the pick and place (PP) task, with a t-value of -2.964, $p = 0.007$, and a corrected $p^* = 0.017$. The effect size of -0.593 indicates a moderate effect.

The increase in HbR suggests higher neuronal activity demands in the right hemisphere during real PP in laparoscopy, reflecting more intensive cognitive and motor resource allocation (see Figure 3.15).

A significant difference was observed for Mean Right HbR in the knot tying (KT) task, with levels changing from Simulation (-2.14 ± 3.97) to Real (-0.75 ± 4.96), yielding a t-value of -2.429, $p = 0.023$, and a corrected $p^* = 0.160$ (not significant after correction). The effect size of -0.496 points to a moderate effect that did not reach statistical significance.

Although not significant after correction, this suggests that real world robotic knot tying slightly reduces deoxygenated hemoglobin in the right hemisphere, potentially indicating a difference in cognitive or motor processing during real KT tasks (see Figure 3.15).

There were no significant differences observed in mean Left HbR between simulation and real settings across both PP and KT tasks.

4. DISCUSSION

This study investigated how minimally invasive surgical modality (laparoscopic vs. robotic surgery) and task complexity (pick and place vs. Knot Tying) influence cognitive workload using functional near infrared spectroscopy. The data revealed clear differences in prefrontal cortex activation and task completion times across surgical modalities, task types, and settings (simulation vs. real). These findings offer significant insights into cognitive demands during surgical training and provide implications for designing targeted training interventions.

4.1 Modality Dependent Cognitive Load

One of the most consistent findings was that laparoscopic surgery elicited significantly higher prefrontal cortical activation (particularly in HbO levels) than robotic surgery, especially during knot tying tasks. This was reflected in both increased oxygenated hemoglobin (HbO) levels in the PFC and shorter task completion times in robotic procedures [61]. These results align with prior research suggesting that robotic platforms, with their ergonomic design, wristed instrumentation, and enhanced 3D visualization, reduce the physical and mental demands on the surgeon [2,3,53]. The elevated neural activation observed in laparoscopic tasks likely stems from the more challenging psychomotor coordination required when operating with limited degrees of freedom, 2D imaging, and indirect camera control [18,20]. This activation was particularly pronounced in the right hemisphere of the prefrontal cortex, where significantly higher HbO levels were observed during laparoscopic knot tying in real settings, suggesting heightened spatial and attentional demands. Laparoscopy exhibited higher cortical activation (HbO levels) in multiple brain regions, including the left and right prefrontal cortex, suggesting increased cognitive and attentional demands. Moreover, laparoscopic tasks generally required longer completion times, particularly under real world conditions. The higher HbR levels in laparoscopy further indicate greater oxygen consumption and task related neuronal activity. These findings suggest that laparoscopy requires greater cognitive engagement and task related brain activity, whereas robotic procedures, while taking longer to complete,

may demand less cognitive effort. This was especially evident in complex tasks such as knot tying, where left and right PFC HbO levels were significantly higher in laparoscopic procedures compared to robotic ones.

Post hoc pairwise comparisons showed large effect sizes (Cohen's d $\geq$ 0.8) in cognitive workload differences between laparoscopic and robotic modalities during complex tasks, reinforcing the ergonomic and technological advantages associated with robotic platforms Interestingly, while robotic surgery showed lower cognitive workload overall, the longer completion times observed in robotic KT tasks suggest that efficiency and comfort may not always translate into faster performance, particularly for novice users still acclimating to the robotic interface [61]. This dissociation between neural effort and performance time suggests that cognitive ease in robotic environments might initially support comfort but not necessarily speed. While HbO differences were statistically significant, some HbR values did not remain significant after multiple comparison corrections, suggesting variability in deoxygenation responses across tasks. These findings underscore the importance of considering both cognitive and behavioral performance indicators when evaluating surgical modality effectiveness. Robotic platforms appear to ease cognitive load through ergonomic design and simplified control interfaces.

Our results revealed that laparoscopic surgery induced significantly higher activation in the prefrontal cortex compared to robotic surgery, suggesting a higher cognitive workload associated with the laparoscopic modality. These findings are consistent with previous research, where robotic surgical simulation tasks were associated with reduced prefrontal cortex activation compared to laparoscopic tasks, indicating a lower cognitive workload [39]. This alignment with the existing literature reinforces the hypothesis that robotic surgical systems, due to their ergonomic advantages and improved control, may ease the cognitive burden on surgeons during complex minimally invasive tasks.

Moreover, completion time distributions for robotic modalities appeared less variable than their laparoscopic counterparts (Figure 3.2), reinforcing the idea that robotic systems may reduce performance inconsistency and support more stable task execution under both simulation and real conditions.

These findings are consistent with broader neuroimaging research summarized by Andersen et al. (2024), who reported that Robotic assisted surgery is often associated with reduced PFC activation compared to laparoscopy [68]. This pattern is attributed to enhanced ergonomics and motor control in robotic systems, reducing the cognitive demands placed on the surgeon, especially during complex procedures like knot tying.

The observed lower cognitive workload during robotic procedures in our study aligns with literature suggesting a more intuitive learning curve. Leijte et al. (2020) demonstrated that participants using robotic systems required fewer repetitions to achieve stable instrument control and efficient visualization, which likely contributed to reduced cognitive demands [21]. These ergonomic and design related advantages support the use of robotic systems for novice surgeons, not only in reducing workload, but also in helping them achieve procedural milestones more efficiently, particularly during complex tasks in real settings, despite initial compromises in knot quality.

4.2 Effects of Task Complexity

Task complexity had a robust and statistically significant effect on cognitive workload, consistent across both modalities and environments. The linear mixed effects model indicated that knot tying tasks imposed higher cognitive demands than pick and place tasks, as evidenced by increased mean ΔHbO levels in the prefrontal cortex. KT tasks, which require advanced bimanual coordination, force manipulation, spatial planning, and decision making, resulted in greater activation in both hemispheres of the PFC compared to PP tasks, particularly in real laparoscopic settings, where both left and right HbO levels were significantly elevated. These findings confirm the hypothesis that more complex motor cognitive integrations demand heightened executive function, particularly in the dorsolateral and anterior medial regions of the PFC, which are associated with planning, attention, and strategy development [53–55]. The higher cognitive demand observed in laparoscopic real PP tasks posits the necessity for enhanced training protocols tailored to task complexity. Task complexity significantly influenced performance, as observed in the differences between knot tying and pick and place tasks.

In robotic procedures, while right PFC activation was lower overall, the contrast between KT and PP tasks remained consistent, suggesting that task complexity independently elevates cognitive load regardless of modality.

Completion time analysis similarly showed that participants required significantly longer durations to complete KT tasks compared to PP tasks, consistent with the hypothesis that more complex psychomotor activities demand greater cognitive and temporal resources.

Post hoc comparisons confirmed that both completion time and cognitive workload were significantly elevated during knot tying tasks across modalities and environments, with moderate to large effect sizes (Cohen's d ranging from 0.5 to 0.8), especially during real laparoscopic KT tasks where the highest ΔHbO values and prolonged completion times were observed. These findings align with previous research highlighting the procedural and attentional challenges inherent to intracorporeal suturing compared to basic object transfer tasks [100–102].

Our results demonstrated that task complexity significantly influenced cognitive workload, with the knot tying task eliciting greater prefrontal activation compared to the simpler pick and place task. This finding is consistent with previous brain in the loop studies using functional near infrared spectroscopy (fNIRS), which revealed that more complex simulated surgical tasks, such as cholecystectomy procedures, induced higher cognitive demands and lower relative neural efficiency than simpler coordination tasks during virtual reality based surgical training [103]. These parallels suggest that task complexity not only increases behavioral demands but also imposes measurable neurophysiological burden on trainees, emphasizing the need to tailor surgical training programs according to task difficulty.

This result reinforces the value of including hierarchical task difficulty in surgical training programs. By identifying which tasks provoke higher cognitive loads, educators can tailor curricula to progressively challenge cognitive resources, monitor overload, and enhance learning retention.

4.3 Effects of Training Environments

The training environment (simulation versus real world surgical settings) significantly affected both cognitive workload and task performance. Participants exhibited higher prefrontal activation (ΔHbO), particularly in the right hemisphere, and longer completion times in the real world environment compared to simulation, suggesting increased visuospatial and attentional demands during actual surgical procedures. This effect was most pronounced during real laparoscopic knot tying, where both left and right PFC HbO levels increased significantly compared to simulation, reflecting the compounded demands of task complexity and real world stressors.

The linear mixed effects analysis demonstrated that the environment factor contributed substantially to the variance in both cognitive workload and performance outcomes. Post hoc tests revealed statistically significant differences (adjusted $p < 0.05$) between simulation and real world conditions across both surgical modalities, with moderate to large effect sizes observed for cognitive workload measures (Cohen's d between 0.6 and 0.9).

This suggests that the cognitive demands of real world surgical settings, driven by physical constraints, unfamiliar ergonomics, and heightened cognitive stress, are not fully replicated in simulation, particularly in laparoscopic procedures. These results suggest that for a basic motor task like pick and place, modality may not strongly impact cognitive engagement in robotic surgery training. These findings are in line with existing literature indicating that while simulators are invaluable tools for skill acquisition, they lack certain fidelity aspects that are critical for evaluating true performance readiness [17, 18]. real world robotic surgery conditions facilitated improved efficiency, as seen in reduced completion times and lower Mean ΔHbO levels. This suggests that while simulation based training is essential for initial skill acquisition, real world exposure is crucial for refining sensorimotor skills and reducing cognitive load. The difficulty in transferring these skills to real world applications underscores the need for enhanced training strategies.

Beyond mean differences, the distribution of completion times provides critical

insight into performance variability across training environments. As illustrated in the violin plots (Figure 3.2), tasks performed in real settings, particularly laparoscopic PP and KT showed broad, asymmetric distributions with long upper tails. This pattern suggests that while some participants adapted well, others struggled significantly under real conditions, likely due to increased cognitive load, unfamiliar ergonomics, or higher stress. In contrast, robotic tasks, especially those performed in simulation, displayed narrower, more symmetric distributions, indicating more uniform performance. These findings suggest that robotic platforms may offer a more forgiving learning environment, while real laparoscopic procedures amplify individual differences in skill and cognitive adaptability. Such distributional insights underscore the importance of using comprehensive visual analytics in cognitive workload research.

The simulation environment, despite its controlled conditions, may not fully replicate the multifactorial demands of real clinical settings. This aligns with findings from the Virtual on Call (iVOC) case study, which emphasized that even high fidelity simulations struggle to elicit the full spectrum of non-technical cognitive stressors seen in real on-call shifts [104].

The gap between simulated and real world cognitive workload observed in this study highlights the importance of incorporating higher fidelity simulation designs and cognitive stressors into training curricula to better prepare surgical trainees for clinical practice.

Interestingly, real robotic knot tying showed a decrease in completion time compared to simulation, with lower or unchanged HbO values. This suggests that real world performance may improve under robotic conditions, possibly due to ergonomic familiarity developed in simulation and reduced reliance on reactive visuomotor strategies.

Thus, while simulators should continue to be integrated into surgical education, this study supports the inclusion of real world or high fidelity environments for performance assessment and final stage training. Moreover, it highlights the utility of fNIRS as an objective adjunct to conventional assessment tools, capable of capturing subtle mental

state differences beyond task time or accuracy.

4.4 Limitations of the Study

While the present study provides valuable insights into the cognitive demands associated with laparoscopic and robotic surgical modalities using functional near infrared spectroscopy (fNIRS), several limitations must be acknowledged.

Firstly, the sample size (n = 26) was relatively small and predominantly comprised novice users of robotic systems. Although participants had comparable experience in laparoscopic surgery, their limited exposure to robotic platforms may have influenced the cognitive workload patterns observed. Future studies should consider recruiting a more diverse participant pool, including experienced robotic surgeons, to enhance the generalizability of the findings.

Secondly, the experimental design did not include live patient procedures. Instead, task performance was assessed in simulated environments and on physical box trainers designed to emulate real world surgical scenarios. While these settings provide a high level of realism, they do not fully replicate the pressures, unpredictability, and environmental factors inherent in actual operating rooms. As such, the conclusions drawn regarding "real world" cognitive workload should be interpreted within the bounds of simulation based training.

Thirdly, the task selection was limited to two standardized procedures: pick and place and knot tying . Although these tasks are widely recognized and represent varying levels of complexity in surgical training, they may not capture the full spectrum of cognitive challenges encountered during complex or unexpected intraoperative situations. Future research should consider including a broader array of surgical tasks, including those that require intraoperative decision making and multitasking under time constraints.

Additionally, the absence of subjective workload measures such as the NASA-TLX

or SURG-TLX limits the study's ability to provide a multidimensional understanding of cognitive load. While fNIRS offers valuable objective insights into neural activity, combining it with subjective assessments and behavioral performance metrics would yield a more comprehensive evaluation of trainee mental workload.

Finally, the use of a single neuroimaging modality presents another limitation. Although fNIRS is well-suited for real time, noninvasive monitoring of cortical activity, incorporating additional physiological measures,such as electrodermal activity (EDA), heart rate variability (HRV), and pupilometry, could enrich the dataset and improve the interpretation of cognitive states. These multimodal approaches may provide more robust and reliable assessments of mental workload in complex surgical environments.

Addressing these limitations in future research will enhance the ecological validity, clinical applicability, and overall impact of cognitive workload assessments in minimally invasive surgical training.

5. CONCLUSION

This thesis systematically investigated the influence of minimally invasive surgical modalities, laparoscopic and robotic, on cognitive workload in relation to task complexity and training environments, employing functional near infrared spectroscopy (fNIRS) to assess prefrontal cortex activation. By adopting linear mixed effects modeling with log-transformed task completion times and applying robust statistical procedures, this study provides empirically grounded insights into the cognitive demands of surgical training.

The results compellingly demonstrate that robotic surgery imposes consistently lower cognitive demands than laparoscopic surgery, particularly during complex tasks such as knot tying. This was especially evident in real world conditions, where both left and right PFC HbO levels were significantly elevated in laparoscopic knot tying compared to robotic procedures, indicating modality and hemisphere specific workload differences. This finding is supported by significant contrasts with large effect sizes, underscoring the ergonomic and technological advantages inherent in robotic platforms. Task complexity emerged as a robust determinant of cognitive burden, with knot tying tasks eliciting higher activation across dorsolateral and anterior medial prefrontal regions compared to simpler pick and place tasks. This aligns with Holper et al. [79], who demonstrated that task complexity is positively associated with increased cortical oxygenation, particularly in motor and prefrontal areas, as measured by fNIRS. These results support the premise that the mental demands of surgical tasks scale with procedural complexity, and that fNIRS based neuroimaging can sensitively detect these differences. Additionally, real world surgical environments were associated with elevated cognitive workload and prolonged task completion times relative to simulation based environments, emphasizing the need for more ecologically valid simulation protocols.

The observed differences, quantified through post hoc comparisons with large Cohen's d values, reinforce the need for differentiated training pathways that account for both modality specific and task specific cognitive demands. Moreover, the successful application of fNIRS in this study highlights its potential as a practical and sensitive tool

for real time cognitive workload assessment in surgical trainees.

Moreover, this study reveals a substantial disparity between simulated and real world training environments. Real surgical settings elicited higher prefrontal hemodynamic responses and longer task completion times, indicating elevated cognitive engagement and executive functioning demands. Although simulation based training remains indispensable for foundational skill acquisition, it fails to fully replicate the cognitive conditions encountered in live procedural environments. This underscores a pressing need for the refinement of simulation protocols and the integration of advanced neuroergonomic assessment tools into surgical training curricula.

In summary, the confluence of findings presented in this work affirms that modality, task complexity, and environment independently and interactively shape cognitive workload during minimally invasive surgery. fNIRS proves to be a viable, field deployable technique for objectively assessing mental workload and offers a promising direction for enhancing surgical education, skill assessment, and patient safety. Its portability, motion tolerance, and real world applicability have been well documented in clinical neuroscience literature, such as the work by Irani et al. [105], which underscores its utility in applied and ecologically valid environments. By incorporating cognitive workload monitoring into training frameworks, educators and institutions can tailor instruction to the learner's cognitive state, thereby fostering more efficient learning trajectories and better operative outcomes.

5.1 Clinical Implications and Future Applications

The results of this study hold important implications for both surgical training and clinical practice. First, the observed reduction in cognitive workload during robotic surgery, particularly during complex tasks, reinforces the ergonomic and cognitive advantages of robotic platforms. These findings support broader adoption and integration of robotic systems, especially in high stakes or cognitively demanding procedures, to enhance surgical performance and potentially reduce error rates.

Second, the demonstrated sensitivity of fNIRS in detecting task related cognitive workload highlights its value as a neuroergonomic assessment tool. The integration of fNIRS into surgical education could enable real time monitoring of cognitive load, offering educators a novel metric to tailor feedback, optimize learning curves, and identify cognitive overload before it manifests as performance degradation. As proposed in neuroergonomics literature [106], incorporating real time cognitive monitoring via fNIRS into surgical training may support adaptive feedback systems that respond to individual trainee workload, improving learning efficiency and reducing risk.

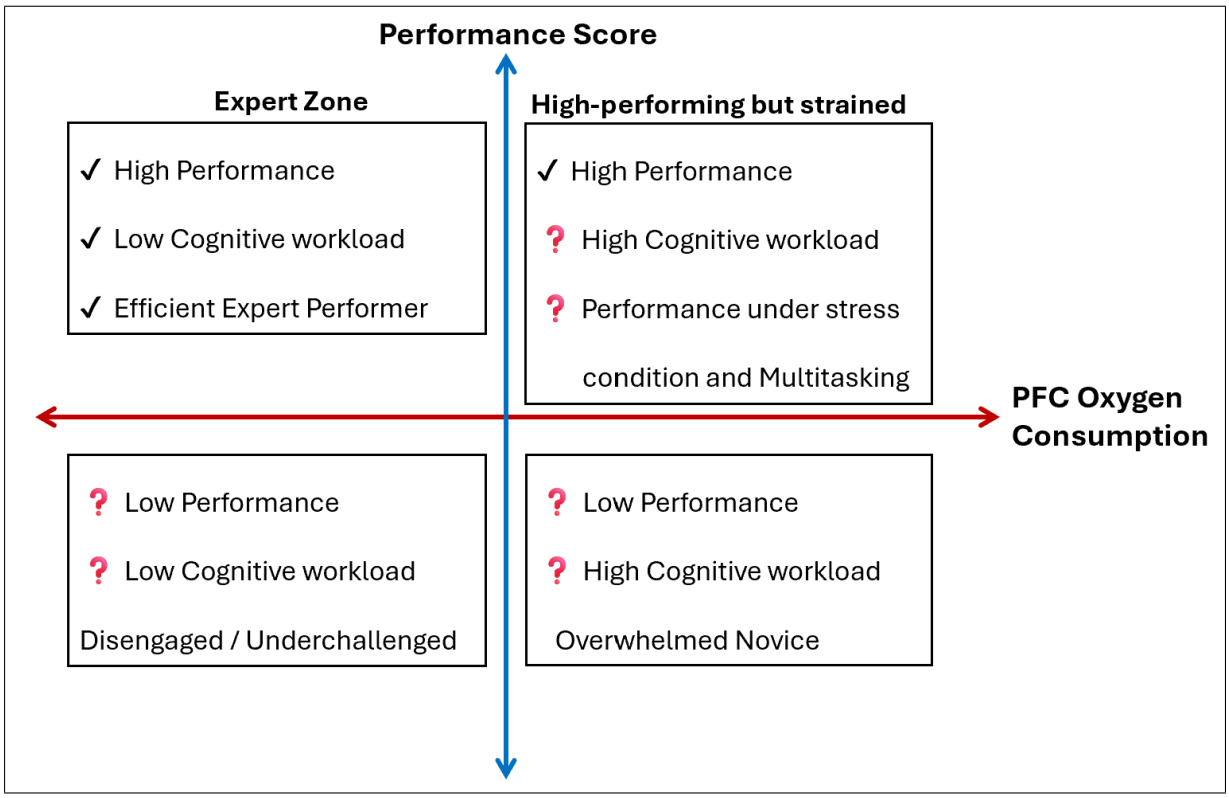

Figure 5.1 Cognitive performance quadrant model for surgical trainees, based on fNIRS-derived prefrontal cortex activation and task completion time.

An important clinical implication of this study is the potential use of neurophysiological data to assess surgical cognitive efficiency. Figure 5.1 presents a cognitive efficiency framework that maps individual performance against prefrontal cortex (PFC) oxygen consumption, as measured by fNIRS. This two dimensional model enables identification of four distinct cognitive states:

- **Expert Zone**: High performance with low cognitive workload, indicating efficient, automated skill execution.
- **High-Performing but Strained**: High performance with high workload, suggesting potential cognitive overload or stress.
- **Overwhelmed Novice**: Low performance with high workload, typical of early-stage learners.
- **Disengaged or Underchallenged**: Low performance with low workload, often due to task simplicity or lack of engagement.

This model supports the development of adaptive training systems and real time feedback loops. By combining fNIRS based neural metrics with behavioral performance, surgical educators can personalize learning experiences, identify trainees in need of intervention, and eventually improve patient safety by reducing cognitive overload in the operating room.

As highlighted by Andersen et al. (2024), neuroimaging holds promise not only for evaluating individual skill acquisition but also for tracking longitudinal progress and tailoring instruction to a learner's cognitive profile [68]. These tools may complement traditional assessments by offering objective, biologically grounded metrics of mental effort and skill proficiency. The future of surgical education may benefit from adaptive training systems that use real time neurophysiological data to guide feedback and optimize learning.

Furthermore, the significant discrepancy in workload between simulated and real world environments underscores the need for more ecologically valid training models. Prior research by Aksoy et al. [60] demonstrated how robotic surgery simulators, assessed with fNIRS, can provide meaningful insights into workload and training effectiveness, highlighting the value of integrating neuroergonomic evaluation tools into simulation curricula. Future simulation systems should strive to better mimic real operative conditions, not only in terms of technical fidelity, but also in cognitive demands, to ensure smoother skill transfer and better preparedness for clinical practice.

Looking ahead, fNIRS based cognitive profiling may be extended beyond training to intraoperative monitoring in the operating room, opening pathways for personalized workload management, adaptive assistance systems, and safety driven interventions. As neuroimaging technologies become more accessible, their adoption into clinical workflows may ultimately improve both surgeon well being and patient outcomes.

5.2 Summary of Key Findings

This thesis explored how different minimally invasive surgical modalities "laparoscopic and robotic" interact with task complexity and environment to influence cognitive workload during skill acquisition and task performance. Using functional near infrared spectroscopy (fNIRS) to monitor prefrontal cortex (PFC) activation, the study yielded several key insights:

- **Surgical Modality:** Robotic surgery was consistently associated with lower cognitive workload compared to laparoscopic surgery. This effect was especially pronounced during complex tasks, highlighting the ergonomic and technological advantages of robotic systems.
- **Task Complexity:** Knot tying tasks imposed significantly greater cognitive demands than simpler pick and place tasks. This was reflected in increased activation in multiple regions of the PFC, particularly in dorsolateral and anterior medial subregions.
- **Training Environment:** Tasks performed in real world surgical environments led to higher cognitive engagement and longer completion times compared to their simulation based counterparts. These findings underscore the limitations of simulation alone in replicating true operative conditions.
- **Neuroergonomic Assessment:** fNIRS proved to be a sensitive, practical, and field deployable neuroimaging tool for quantifying mental workload in surgical trainees. Its use enables deeper understanding of the cognitive demands imposed by different training configurations.

Collectively, these findings validate the multidimensional nature of surgical cognitive workload and offer empirical support for incorporating cognitive assessments into surgical education. They also highlight the need to refine simulation curricula and consider personalized, modality specific training pathways to optimize surgeon performance and patient safety.

5.3 List of Publications

1. The Impact of Minimally Invasive Surgical Modality and Task Complexity on Cognitive Workload: An fNIRS Study, Ücrak, F.; Izzetoglu, K.; Polat, M.D.; Gür, Ü.; Şahin, T.; Yöner, S.I.; İnan, N.G.; Aksoy, M.E.; Öztürk, C. , Brain Sci. 2025, 15, 387. https://doi.org/10.3390/brainsci15040387

2. Functional Brain Activity Measures to Quantify Cognitive Workload During Laparoscopic and Robotic Surgery Techniques, Ücrak, F., Izzetoglu, K., Polat, M.D., Gür, Ü., Şahin, T., Yöner, S.I., İnan, N.G., Aksoy, M.E., Öztürk, C., HCI International 2025 . HCII 2025,Springer, Accepted.

REFERENCES

1. Buia, A., F. Stockhausen, and E. Hanisch, "Laparoscopic surgery: A qualified systematic review," *World Journal of Methodology*, Vol. 5, no. 4, 2015.

2. Gamal, A., M. C. Moschovas, A. R. Jaber, S. Saikali, and et al., "Clinical applications of robotic surgery platforms: a comprehensive review," *Journal of Robotic Surgery*, Vol. 18, no. 1, 2024.

3. Shugaba, A., J. E. Lambert, T. M. Bampouras, H. E. Nuttall, C. J. Gaffney, and et al., "Should all minimal access surgery be robot-assisted? a systematic review into the musculoskeletal and cognitive demands of laparoscopic and robot-assisted laparoscopic surgery," *Journal of Gastrointestinal Surgery*, Vol. 26, no. 7, pp. 1520–1530, 2022.

4. Alkatout, I., U. Mechler, L. Mettler, J. Pape, and et al., "The development of laparoscopy a historical overview," *Frontiers in Surgery*, Vol. 8, 2021.

5. De Marchi, D., G. Mantica, A. Tafuri, G. Giusti, and et al., "Robotic surgery in urology: a review from the beginning to the single-site," *AME Medical Journal*, Vol. 7, 2021.

6. Singh, H., H. N. Modi, S. Ranjan, J. W. R. Dilley, and et al., "Robotic surgery improves technical performance and enhances prefrontal activation during high temporal demand," *Annals of Biomedical Engineering*, Vol. 46, no. 10, pp. 1621–1636, 2018.

7. Biswas, P., S. Sikander, and P. Kulkarni, "Recent advances in robot-assisted surgical systems," *Biomedical Engineering Advances*, Vol. 6, 2023.

8. Herron, D. M., and M. Marohn, "A consensus document on robotic surgery," *Surgical Endoscopy*, Vol. 22, no. 2, pp. 313–325, 2007.

9. Zorn, K. C., "Robotic radical prostatectomy: advantages of an initial posterior dissection," *Journal of Robotic Surgery*, Vol. 2, no. 3, pp. 135–137, 2008.

10. Shah, A. A., J. Bandari, D. Pelzman, B. J. Davies, and et al., "Diffusion and adoption of the robot in urology," *Translational Andrology and Urology*, Vol. 10, no. 5, pp. 2151–2157, 2019.

11. Zendejas, B., R. Brydges, S. J. Hamstra, and D. A. Cook, "State of the evidence on simulation-based training for laparoscopic surgery," *Annals of Surgery*, Vol. 257, no. 4, pp. 586–593, 2013.

12. Agha, R. A., and A. J. Fowler, "The role and validity of surgical simulation," *International Surgery*, Vol. 100, no. 2, pp. 350–357, 2015.

13. Qian, K., J. Bai, X. Yang, J. Pan, and et al., "Essential techniques for laparoscopic surgery simulation," *Computer Animation and Virtual Worlds*, Vol. 28, no. 2, 2016.

14. Gordon, C. J., T. Ryall, and B. Judd, "Simulation-based assessments in health professional education: a systematic review," *Journal of Multidisciplinary Healthcare*, 2016.

15. Ziv, A., P. R. Wolpe, S. D. Small, and S. Glick, "Simulation-based medical education: An ethical imperative," *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, Vol. 1, no. 4, pp. 252–256, 2006.

16. Hong, M., J. W. Rozenblit, and A. J. Hamilton, "Simulation-based surgical training systems in laparoscopic surgery: a current review," *Virtual Reality*, Vol. 25, no. 2, pp. 491–510, 2020.

17. Ayaz, H., P. A. Shewokis, S. Bunce, K. Izzetoglu, and et al., "Optical brain monitoring for operator training and mental workload assessment," *NeuroImage*, Vol. 59, no. 1, pp. 36–47, 2012.

18. Hur, H.-C., D. Arden, L. E. Dodge, B. Zheng, and et al., "Fundamentals of laparoscopic surgery: A surgical skills assessment tool in gynecology," *JSLS : Journal of the Society of Laparoendoscopic Surgeons*, Vol. 15, no. 1, pp. 21–26, 2011.

19. Mizota, T., V. G. Dodge, and D. Stefanidis, *Fundamentals of Robotic Surgery*, book section Chapter 16, pp. 215–225. 2018.

20. Cullinan, D. R., M. R. Schill, A. DeClue, A. Salles, and et al., "Fundamentals of laparoscopic surgery: Not only for senior residents," *Journal of Surgical Education*, Vol. 74, no. 6, pp. e51–e54, 2017.

21. Leijte, E., I. de Blaauw, F. Van Workum, C. Rosman, and et al., "Robot assisted versus laparoscopic suturing learning curve in a simulated setting," *Surgical Endoscopy*, Vol. 34, no. 8, pp. 3679–3689, 2019.

22. Roy, R. N., A. Moly, F. Dehais, and S. Scannella, *EEG and FNIRS Connectivity Features for Mental Workload Assessment*, pp. 327–328. 2018.

23. Hancock, P. A., and P. A. Desmond, *Stress, Workload, and Fatigue*, 2000.

24. Causse, M., Z. Chua, V. Peysakhovich, N. Del Campo, and N. Matton, "Mental workload and neural efficiency quantified in the prefrontal cortex using fnirs," *Scientific Reports*, Vol. 7, no. 1, 2017.

25. Polat, M. D., K. Izzetoglu, M. E. Aksoy, D. Kitapcioglu, and et al., *Cognitive Load Quantified via Functional Near Infrared Spectroscopy During Immersive Training with VR Based Basic Life Support Learning Modules in Hostile Environment*, book section Chapter 23, pp. 359–372. Lecture Notes in Computer Science, 2023.

26. Benerradi, J., H. A. Maior, A. Marinescu, J. Clos, and et al., "Exploring machine learning approaches for classifying mental workload using fnirs data from hci tasks," 2019.

27. Longo, L., C. D. Wickens, G. Hancock, and P. A. Hancock, "Human mental workload: A survey and a novel inclusive definition," *Frontiers in Psychology*, Vol. 13, 2022.

28. Lohani, M., B. R. Payne, and D. L. Strayer, "A review of psychophysiological measures to assess cognitive states in real-world driving," *Frontiers in Human Neuroscience*, Vol. 13, 2019.

29. Jin, K., A. Rubio-Solis, R. Naik, T. Onyeogulu, and et al., "Identification of cognitive workload during surgical tasks with multimodal deep learning," *arXiv preprint arXiv:2209.06208*, 2022.

30. Almukhtar, A., V. Caddick, R. Naik, M. Goble, and et al., "Objective assessment of cognitive workload in surgery," *Annals of Surgery*, 2024.

31. Scerbo, M. W., F. G. Freeman, and P. J. Mikulka, "A brain-based system for adaptive automation," *Theoretical Issues in Ergonomics Science*, Vol. 4, no. 1-2, pp. 200–219, 2003.

32. Gevins, A., "High-resolution eeg mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice," *Cerebral Cortex*, Vol. 7, no. 4, pp. 374–385, 1997.

33. Wilson, G. F., and C. A. Russell, "Operator functional state classification using multiple psychophysiological features in an air traffic control task," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 45, no. 3, pp. 381–389, 2003.

34. Soltanlou, M., M. A. Sitnikova, H.-C. Nuerk, and T. Dresler, "Applications of functional near-infrared spectroscopy (fnirs) in studying cognitive development: The case of mathematics and language," *Frontiers in Psychology*, Vol. 9, 2018.

35. Das Chakladar, D., and P. P. Roy, "Cognitive workload estimation using physiological measures: a review," *Cognitive Neurodynamics*, Vol. 18, no. 4, pp. 1445–1465, 2024.

36. Crosson, B., A. Ford, K. M. McGregor, M. Meinzer, and et al., "Functional imaging and related techniques: An introduction for rehabilitation researchers," *The Journal of Rehabilitation Research and Development*, Vol. 47, no. 2, 2010.

37. Getchell, N., and P. Shewokis, "Understanding the role of cognitive effort within contextual interference paradigms: Theory, measurement, and tutorial," *Brazilian Journal of Motor Behavior*, Vol. 17, no. 1, pp. 59–69, 2023.

38. Pinti, P., I. Tachtsidis, A. Hamilton, J. Hirsch, and et al., "The present and future use of functional near-infrared spectroscopy (fnirs) for cognitive neuroscience," *Annals of the New York Academy of Sciences*, Vol. 1464, no. 1, pp. 5–29, 2020.

39. Aksoy, M. E., K. Izzetoglu, N. Z. Utkan, A. Agrali, and et al., "Comparing behavioral and neural activity changes during laparoscopic and robotic surgery trainings," *Journal of Surgical Education*, Vol. 82, no. 5, 2025.

40. Zou, X., X. Liu, P. Zhang, K. Wang, and et al., "Cognitive load in novice uav pilots: A preliminary fnirs investigation," *SSRN*, 2025.

41. Izzetoglu, M., S. C. Bunce, K. Izzetoglu, B. Onaral, and et al., "Functional brain imaging using near-infrared technology," *IEEE Engineering in Medicine and Biology Magazine*, Vol. 26, no. 4, pp. 38–46, 2007.

42. Richards, D., K. Izzetoglu, and J. Armstrong, "Using functional near infrared spectroscopy to assess cognitive performance of uav sensor operators during route scanning," 2018.

43. Chen, Y.-H., and M. Sawan, *Monitoring Brain Activities Using fNIRS to Avoid Stroke*, book section Chapter 6. 2023.

44. Hoshi, Y., *Hemodynamic signals in fNIRS*, pp. 153–179. Progress in Brain Research, 2016.

45. Izzetoglu, K., S. Bunce, B. Onaral, K. Pourrezaei, and et al., "Functional optical brain imaging using near-infrared during cognitive tasks," *International Journal of Human-Computer Interaction*, Vol. 17, no. 2, pp. 211–227, 2004.

46. Naseer, N., and K.-S. Hong, "fnirs-based brain-computer interfaces: a review," *Frontiers in Human Neuroscience*, Vol. 9, 2015.

47. Ferrari, M., and V. Quaresima, "A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application," *NeuroImage*, Vol. 63, no. 2, pp. 921–935, 2012.

48. Rashidi Fathabadi, F., J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, "Autonomous sequential surgical skills assessment for the peg transfer task in a laparoscopic box-trainer system with three cameras," *Robotica*, Vol. 41, no. 6, pp. 1837–1855, 2023.

49. EkÃ§i, B., "A simple technique for knot tying in single incision laparoscopic surgery (sils)," *Clinics*, Vol. 65, no. 10, pp. 1055–1057, 2010.

50. Corporation, S., "Peg transfer board with triangles," 2024.

51. Loukas, C., "Surgical simulation training systems: Box trainers, virtual reality and augmented reality simulators," *International Journal of Advanced Robotics and Automation*, Vol. 1, no. 2, pp. 1–9, 2016.

52. Wilson, C. R. E., D. Gaffan, P. G. F. Browning, and M. G. Baxter, "Functional localization within the prefrontal cortex: missing the forest for the trees?," *Trends in Neurosciences*, Vol. 33, no. 12, pp. 533–540, 2010.

53. Aron, A. R., T. W. Robbins, and R. A. Poldrack, "Inhibition and the right inferior frontal cortex," *Trends in Cognitive Sciences*, Vol. 8, no. 4, pp. 170–177, 2004.

54. Miller, E. K., and J. D. Cohen, "An integrative theory of prefrontal cortex function," *Annual Review of Neuroscience*, Vol. 24, no. 1, pp. 167–202, 2001.

55. Duncan, J., "The multiple-demand (md) system of the primate brain: mental programs for intelligent behaviour," *Trends in Cognitive Sciences*, Vol. 14, no. 4, pp. 172–179, 2010.

56. Chen, Y., Z. Cao, M. Mao, W. Sun, and et al., "Increased cortical activation and enhanced functional connectivity in the prefrontal cortex ensure dynamic postural balance during dual-task obstacle negotiation in the older adults: A fnirs study," *Brain and Cognition*, Vol. 163, 2022.

57. Harrison, B. J., H. Koshino, T. Minamoto, T. Ikeda, and et al., "Anterior medial prefrontal cortex exhibits activation during task preparation but deactivation during task execution," *PLoS ONE*, Vol. 6, no. 8, 2011.

58. Koechlin, E., G. Basso, P. Pietrini, S. Panzer, and et al., "The role of the anterior prefrontal cortex in human cognition," *Nature*, Vol. 399, no. 6732, pp. 148–151, 1999.

59. Izzetoglu, K., H. Ayaz, A. Merzagora, M. Izzetoglu, and et al., "The evolution of field deployable fnir spectroscopy from bench to clinical settings," *Journal of Innovative Optical Health Sciences*, Vol. 04, no. 03, pp. 239–250, 2011.

60. Aksoy, M. E., K. Izzetoglu, A. Agrali, D. Kitapcioglu, M. Gungor, and A. Simsek, *Effect of Robotic Surgery Simulators in Training Assessed by Functional Near-Infrared Spectroscopy (fNIRs)*, book section Chapter 18, pp. 271–278. Lecture Notes in Computer Science, 2020.

61. Izzetoglu, K., M. E. Aksoy, A. Agrali, D. Kitapcioglu, and et al., "Studying brain activation during skill acquisition via robot-assisted surgery training," *Brain Sciences*, Vol. 11, no. 7, 2021.

62. Aksoy, M. E., B. Kocaoglu, K. Ä°zzetoglu, A. Agrali, and et al., "Assessment of learning in simulator based arthroscopy training with the diagnostic arthroscopy skill score (dass) and neurophysiological measures," *Knee Surgery, Sports Traumatology, Arthroscopy*, Vol. 31, no. 12, pp. 5332–5345, 2023.

63. Nemani, A., M. A. YÃ¼cel, U. Kruger, D. W. Gee, and et al., "Assessing bimanual motor skills with optical neuroimaging," *Science Advances*, Vol. 4, no. 10, 2018.

64. Fu, Y., P. Walia, S. D. Schwaitzberg, X. Intes, and et al., "Changes in functional neuroimaging measures as novices gain proficiency on the fundamentals of laparoscopic surgery suturing task," *Neurophotonics*, Vol. 10, no. 02, 2023.

65. Holper, L., R. H. W. ten Brincke, M. Wolf, and R. O. Murphy, "fnirs derived hemodynamic signals and electrodermal responses in a sequential risk-taking task," *Brain Research*, Vol. 1557, pp. 141–154, 2014.

66. Kawaguchi, K., Y. Nikai, S. Yomota, A. Kawashima, and et al., "Effects of age and flight experience on prefrontal cortex activity in airline pilots: An fnirs study," *Heliyon*, Vol. 10, no. 9, 2024.

67. Mark, J. A., A. E. Kraft, M. D. Ziegler, and H. Ayaz, "Neuroadaptive training via fnirs in flight simulators," *Frontiers in Neuroergonomics*, Vol. 3, 2022.

68. Andersen, A. G., A. C. Riparbelli, H. R. Siebner, L. Konge, and et al., "Using neuroimaging to assess brain activity and areas associated with surgical skills: a systematic review," *Surgical Endoscopy*, Vol. 38, no. 6, pp. 3004–3026, 2024.

69. Sklivanioti Greenfield, M., Y. Wang, and M. Msghina, "Similarities and differences in the induction and regulation of the negative emotions fear and disgust: A functional near infrared spectroscopy study," *Scandinavian Journal of Psychology*, Vol. 63, no. 6, pp. 581–593, 2022.

70. Critchley, H. D., "Review: Electrodermal responses: What happens in the brain," *The Neuroscientist*, Vol. 8, no. 2, pp. 132–142, 2002.

71. Greco, A., A. Lanata, G. Valenza, E. P. Scilingo, and et al., "Electrodermal activity processing: A convex optimization approach," 2014.

72. Sequeira, H., P. Hot, L. Silvert, and S. Delplanque, "Electrical autonomic correlates of emotion," *International Journal of Psychophysiology*, Vol. 71, no. 1, pp. 50–56, 2009.

73. Greco, A., G. Valenza, A. Lanata, E. Scilingo, and et al., "cvxeda: a convex optimization approach to electrodermal activity processing," *IEEE Transactions on Biomedical Engineering*, pp. 1–1, 2016.

74. Movisens, "Electrodermal activity," 18/06/2024 2024.

75. Braithwaite, J. J., D. G. Watson, R. Jones, and M. Rowe, "A guide for analysing electrodermal activity (eda) and skin conductance responses (scrs) for psychological experiments," *Psychophysiology*, Vol. 49, no. 1, pp. 1017–1034, 2013.

76. Wass, S. V., K. de Barbaro, and K. Clackson, "Tonic and phasic co-variation of peripheral arousal indices in infants," *Biological Psychology*, Vol. 111, pp. 26–39, 2015.

77. Winter, M., R. Pryss, T. Probst, and M. Reichert, "Towards the applicability of measuring the electrodermal activity in the context of process model comprehension: Feasibility study," *Sensors*, Vol. 20, no. 16, 2020.

78. Sarchiapone, M., C. Gramaglia, M. Iosue, V. Carli, and et al., "The association between electrodermal activity (eda), depression and suicidal behaviour: A systematic review and narrative synthesis," *BMC Psychiatry*, Vol. 18, no. 1, 2018.

79. Holper, L., F. Scholkmann, and M. Wolf, "The relationship between sympathetic nervous activity and cerebral hemodynamics and oxygenation: A study using skin conductance measurement and functional near-infrared spectroscopy," *Behavioural Brain Research*, Vol. 270, pp. 95–107, 2014.

80. Topoglu, Y., J. Watson, R. Suri, and H. Ayaz, *Electrodermal Activity in Ambulatory Settings: A Narrative Review of Literature*, book section Chapter 10, pp. 91–102. Advances in Intelligent Systems and Computing, 2020.

81. Holper, L., and R. O. Murphy, "Hemodynamic and affective correlates assessed during performance on the columbia card task (cct)," *Brain Imaging and Behavior*, Vol. 8, no. 4, pp. 517–530, 2013.

82. Watson, J., A. Sargent, Y. Topoglu, H. Ye, and et al., *Using fNIRS and EDA to Investigate the Effects of Messaging Related to a Dimensional Theory of Emotion*, book section Chapter 6, pp. 59–67. Advances in Intelligent Systems and Computing, 2020.

83. Bogar, P. Z., M. Virag, M. Bene, P. Hardi, and et al., "Validation of a novel, low-fidelity virtual reality simulator and an artificial intelligence assessment approach for peg transfer laparoscopic training," *Scientific Reports*, Vol. 14, no. 1, 2024.

84. Ayaz, H., P. A. Shewokis, A. Curtin, M. Izzetoglu, and et al., "Using mazesuite and functional near infrared spectroscopy to study learning in spatial navigation," *Journal of Visualized Experiments*, no. 56, 2011.

85. Brigadoi, S., L. Ceccherini, S. Cutini, F. Scarpa, and et al., "Motion artifacts in functional near-infrared spectroscopy: A comparison of motion correction techniques applied to real cognitive data," *NeuroImage*, Vol. 85, pp. 181–191, 2014.

86. Li, R., D. Yang, F. Fang, K.-S. Hong, and et al., "Concurrent fnirs and eeg for brain function investigation: A systematic, methodology-focused review," *Sensors*, Vol. 22, no. 15, p. 5865, 2022.

87. The MathWorks, I., "Matlab," 2022.

88. Dans, P. W., S. D. Foglia, and A. J. Nelson, "Data processing in functional near-infrared spectroscopy (fnirs) motor control research," *Brain Sciences*, Vol. 11, no. 5, 2021.

89. Tak, S., and J. C. Ye, "Statistical analysis of fnirs data: A comprehensive review," *NeuroImage*, Vol. 85, pp. 72–91, 2014.

90. Pinti, P., F. Scholkmann, A. Hamilton, P. Burgess, and et al., "Current status and issues regarding pre-processing of fnirs neuroimaging data: An investigation of diverse signal filtering methods within a general linear model framework," *Frontiers in Human Neuroscience*, Vol. 12, 2019.

91. Iatsenko, D., P. V. E. McClintock, and A. Stefanovska, "Linear and synchrosqueezed time-frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms," *Digital Signal Processing*, Vol. 42, pp. 1–26, 2015.

92. Iatsenko, D., P. V. E. McClintock, and A. Stefanovska, "Extraction of instantaneous frequencies from ridges in time-frequency representations of signals," *Signal Processing*, Vol. 125, pp. 290–303, 2016.

93. Villringer, A., and B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function," *Trends in Neurosciences*, Vol. 20, no. 10, pp. 435–442, 1997.

94. Kreplin, U., C. Burns, and S. H. Fairclough, "Fnirs activity in the prefrontal cortex and motivational intensity: impact of working memory load, financial reward, and correlation-based signal improvement," *Neurophotonics*, Vol. 5, no. 03, 2018.

95. Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*, 2013.

96. Seabold, S., and J. Perktold, "Statsmodels: Econometric and statistical modeling with python," 2010.

97. Vallat, R., "Pingouin: statistics in python," *Journal of Open Source Software*, Vol. 3, no. 31, 2018.

98. Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, and et al., "Scipy 1.0: fundamental algorithms for scientific computing in python," *Nature Methods*, Vol. 17, no. 3, pp. 261–272, 2020.

99. Hunter, J. D., "Matplotlib: A 2d graphics environment," *Computing in Science and Engineering*, Vol. 9, no. 3, pp. 90–95, 2007.

100. Seymour, N. E., A. G. Gallagher, S. A. Roman, M. K. O'Brien, and et al., "Virtual reality training improves operating room performance: Results of a randomized, double-blinded study," *Annals of Surgery*, Vol. 236, no. 4, pp. 458–464, 2002.

101. Mina, A. A., J. F. Knipfer, D. Y. Park, H. A. Bair, and et al., "Intracranial complications of preinjury anticoagulation in trauma patients with head injury," *Journal of Trauma and Acute Care Surgery*, Vol. 53, no. 4, pp. 668–672, 2002.

102. Scott, D. J., W. N. Young, S. T. Tesfay, W. H. Frawley, and et al., "Laparoscopic skills training," *The American Journal of Surgery*, Vol. 182, no. 2, pp. 137–142, 2001.

103. Shewokis, P. A., H. Ayaz, L. Panait, Y. Liu, and et al., "Brain-in-the-loop learning using fnir and simulated virtual reality surgical tasks: Hemodynamic and behavioral effects," Foundations of Augmented Cognition, pp. 324–335, Springer International Publishing.

104. Emin, E. I., E. Emin, A. Bimpis, M. Pierides, and et al., "Teaching and assessment of medical students during complex multifactorial team based tasks: The virtual on call case study," *Advances in Medical Education and Practice*, Vol. Volume 13, pp. 457–465, 2022.

105. Irani, F., S. M. Platek, S. Bunce, A. C. Ruocco, and et al., "Functional near infrared spectroscopy (fnirs): An emerging neuroimaging technology with important applications for the study of brain disorders," *The Clinical Neuropsychologist*, Vol. 21, no. 1, pp. 9–37, 2007.

106. Dehais, F., A. Lafont, R. Roy, and S. Fairclough, "A neuroergonomics approach to mental workload, engagement and human performance," *Frontiers in Neuroscience*, Vol. 14, 2020.