Motor Learning Induced Neuroplasticity in Minimally Invasive Surgery

Kunal Shetty MBBS, MRCS (Eng), MSc

The Hamlyn Centre for Robotic Surgery Division of Surgery,
Department of Surgery and Cancer, St Mary's Hospital,
Imperial College London

2016

This thesis is submitted to Imperial College of Science, Technology and Medicine in partial fulfilment of the requirements for the degree of Doctor of Philosophy. The presented work except where indicated describes the results of my own research.

ABSTRACT

Technical skills in surgery have become more complex and challenging to acquire since the introduction of technological aids, particularly in the arena of Minimally Invasive Surgery. Additional challenges posed by reforms to surgical careers and increased public scrutiny, have propelled identification of methods to assess and acquire MIS technical skills. Although validated objective assessments have been developed to assess motor skills requisite for MIS, they poorly understand the development of expertise. Motor skills learning, is indirectly observable, an internal process leading to relative permanent changes in the central nervous system. Advances in functional neuroimaging permit direct interrogation of evolving patterns of brain function associated with motor learning due to the property of neuroplasticity and has been used on surgeons to identify the neural correlates for technical skills acquisition and the impact of new technology. However significant gaps exist in understanding neuroplasticity underlying learning complex bimanual MIS skills. In this thesis the available evidence on applying functional neuroimaging towards assessment and enhancing operative performance in the field of surgery has been synthesized.

The purpose of this thesis was to evaluate frontal lobe neuroplasticity associated with learning a complex bimanual MIS skill using functional near-infrared spectroscopy an indirect neuroimaging technique. Laparoscopic suturing and knot-tying a technically challenging bimanual skill is selected to demonstrate learning related reorganisation of cortical behaviour within the frontal lobe by shifts in activation from the prefrontal cortex (PFC) subserving attention to primary and secondary motor centres (premotor cortex, supplementary motor area and primary motor cortex) in which motor sequences are encoded and executed. In the cross-sectional study, participants of varying expertise demonstrate frontal lobe neuroplasticity commensurate with motor learning. The longitudinal study involves tracking evolution in cortical behaviour of novices in response to receipt of eight hours distributed training over a fortnight. Despite novices achieving expert like performance and stabilisation on the technical task, this study demonstrates that novices displayed persistent PFC activity. This study establishes for complex bimanual tasks, that improvements in technical performance do not accompany a reduced reliance in attention to support performance. Finally, least-squares support vector machine is used to classify expertise based on frontal lobe functional connectivity. Findings of this thesis demonstrate the value of interrogating cortical behaviour towards assessing MIS skills development and credentialing.

ACKNOWLEDGEMENTS

"Never say never, because limits like fears, are often just an illusion"

Michael Jordan

Firstly, I would like to thank my supervisors *Professor the Lord Ara Darzi of Denham* and *Professor Guang-Zhong Yang* for providing me a path to pursue this research in the scintillating field of surgical innovation and technology. Their grand vision, unwavering commitment and alliance in forging two fields to form the Hamlyn Centre for Robotic Surgery to be the cynosure for surgical innovation is truly inspirational. It has certainly been a privilege to undertake research under their guidance in this distinguished academic centre.

The prospect of undertaking research that involved neuroscience and functional neuroimaging was a 'fish out of water' experience to me as a surgical trainee. If it not were for the unwavering support and guidance of *Mr Daniel Leff* from beginning until end, the task in hand would seem insurmountable. His tenacity, ability to simplify complex problems and availability despite juggling multiple roles is truly inspirational. I am grateful for the countless discussions, the vast of amount of time spent in supervision and his constant motivation through the highs and lows of my academic pursuit.

Undertaking research in the department has enabled access to *Professor Thanos Athanasiou* who has been a pillar of support in terms of statistical analysis, adherence to rigorous methodology and critical analysis. His ability to understand complex problems, expertise in statistical analysis and vast knowledge is uniquely exceptional. I will fondly look back at the friendship we have forged and I am indebted for the time he has spent providing honest appraisal of my research and providing statistical support.

Analysis of high-density neuroimaging data represented a daunting challenge, scaled with the help of *Felipe Orihuela-Espina* and *Javier Andreu-Perez*. *Felipe's* vast mathematical knowledge and passion for fNIRS was notably infectious. I am grateful for the time he spent in explaining to me mathematical concepts and their application, specific to functional neuroimaging. In addition I am grateful for the analytic support provided throughout this research. I am also indebted to *Javier* for devoting his time in setting up analytic frameworks and meaningful interpretations of data. Several other members of the Hamlyn Centre for Robotic Surgery played an important role in supporting my research. I would like to extend my thanks to *Blair Crewther, Kumuthan Sriskandarajah, and Muzzafer Chaudhery* for their support and friendship.

This tenure has coincided with more than one personally challenging circumstance. One of the values my parents had instilled in me is the value of perseverance and for this I thank them. I would also like to thank my brother *Nikith* for providing me an overall perspective when needed. This period has most certainly had an impact on my wife *Surya* and son *Krishna* for the countless hours I have spent away from them. I am eternally grateful for her understanding, selflessness and steadfast support, without which this would have been impossible.

I dedicate this thesis to my recently born daughter *Maya* who provided me the timely resolve to complete my research.

CONTENTS

CHAPTER 6

EFFECTS OF PRACTICE ON FRONTAL LOBE NEUROPLASTICITY IN MIS: A LONGITUDINAL STUDY……183

CHAPTER 7

FRONTAL LOBE CONNECTIVITY AIDS EXPERTISE CLASSIFICATION……209

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

2D Two-dimensions

3D Three-dimensions

AP Action Potential

ATP Adenosine Triphosphate

A.U. Arbitrary units

BOLD Blood Oxygenation Level Dependent

BP Blood Pressure

BCI Brain Computer Interface

CBF Cerebral Blood Flow

CI Confidence Interval

DCM Dynamic Causal Modelling

DOT Diffuse Optical Topography

DPF Differential Pathlength Factor

dPMC dorsal Premotor Cortex

DTI Diffusion Tensor Imaging

DTK Double Throw Knot

E Expert

ECG Electrocardiogram

EEG Electroencephalogram

EMD Earth Mover’s Distance

EMG Electromyography

F Final

FLS Fundamentals of Laparoscopic Surgery

fMRI Functional Magnetic Resonance Imaging

fNIRS Functional Near Infrared Spectroscopy

F-P Frontoparietal

Fr Friedman’s test

GRS Global Rating Scale

Hb Haemoglobin

HbO_2 Oxyhaemoglobin

HHb Deoyhaemoglobin

HbT Total Haemoglobin

HMM Hidden Markov Models

HR Heart Rate

HRV Heart Rate Variability

HUESAD Hiroshima University Endoscopic Surgical Assessment Device

Hz Hertz

I Initial

ICA Independent Component Analysis

ICNA Imperial College Neuroimage Analysis

ICSAD Imperial College Surgical Assessment Device LF Low Frequency

IQR Interquartile range

KW Kruskall-Wallis test

LDA Linear Discriminate Analysis

LICS Laparoscopic Intra-corporeal suturing

LREC Local Regional Ethics Committee

M Mid

M1 Primary Motor Cortex

MBLL Modified Beer Lambert Law

MEG Magnetoencephalography

MIS Minimally invasive surgery

MISTELS McGill Inanimate System for Training and Evaluation of Laparoscopic Skill

MRI Magnetic Resonance Imaging

MWU Mann Whitney U test

NASA-TLX National Aerospace Agency – Task Load Index

NDI Needle Drive Insertion

NIR Near Infrared NIRS Near Infrared Spectroscopy

NOTES Natural Orifice Translumenal Endoscopic Surgery

OCHRA Observational Clinical Human Reliability Assessment

OSATS Objective Structured Assessment of Technical Skill

OR Operating room

OT Optical Topography

PCA Principal Component Analysis

PET Positron Emission Tomography

PFC Prefrontal Cortex

PMC Premotor Cortex

PL Pathlength

PMC Premotor Cortex

PPC Posterior Parietal Cortex

PRISMA Preferred Statement for Reporting Systematic Reviews and Meta-Analyses

REM Random Effects Model

ROI Region of Interest

S1 Somatosensory Cortex

SD Standard Deviation

SD_{RR} Standard Deviation of the R to R Interval

SEM Structural Equation Modelling

SMA Supplementary Motor Area

SPECT Single Photon Emission Computed Tomography

SPSS Statistical Package for Social Sciences

STAI State Trait Anxiety Inventory

STK Single Throw KnotV1 Visual Cortex VLPFC

SVM Support Vector Machnie

SWI Small-World Index

T Trainee

V1 Visual Cortex

VR Virtual Reality

WSR Wilcoxon sign rank test

Chapter 1 Introduction

Learning skilled movements is of paramount importance for daily human activity whether at work or home environments. The human brain is unique in its remarkable ability to learn a wide variety of skilled movements that require interaction with tools and the environment. In order to perform learnt skilled movements, several complex processes occur in both perceptual and motor domains of the brain[3]. Surgeons represent a unique population in whom to study motor skills acquisition due to the intensity of training, and the precision and complexity of the bimanual motor skills they need to assimilate. Dexterity analysis in surgeons is valuable as surgical skill has been identified to be a key factor in attaining superior operative outcomes[4].

Over the last thirty years, the incorporation of technological aids has revolutionized the field of surgery of which Minimally invasive surgery (MIS) has arguably had the greatest impact[5]. MIS through laparoscopic or robotic platforms has enabled surgeons to perform more complex operations through smaller incisions thereby decreasing post-operative pain and the time required for return to normal function[6]. Presently, several index operations are performed more commonly via MIS over the traditional approach[7]. However, MIS operations are more challenging to perform, manifest by protracted learning curves for adult[8] and paediatric surgery[9], greater operative times[6], higher stress levels induced[10] and a greater number of complications observed during the early phase of the long learning curve[11].

Laparoscopic surgery (LS), the most common MIS approach, induces a number of constraints on the operator that include dependence on an assistant to manipulate the laparoscopic camera, reduction in freedom of movement, adaptation to visuo-motor scaling, tremor amplification at the instrument tips and paradoxical movements stemming from the fulcrum effect, loss of depth perception and distortion of tactile feedback. Robotic surgery visualised by many as the next technological leap in the field of MIS, addressed several of the ergonomic limitations of LS, yet introduced a new visuo-motor disturbance in the form of complete loss of haptic feedback. Despite the improvement in kinematics compared to LS, it is still considerably inferior to unrestricted hand movement (*7 versus 23 degrees of freedom*), and as a result, the learning curve for proficiency in RS is still considerably longer than its open equivalent[12].

Therefore with the advent of MIS, proficiency in skilled motor performance is paramount, yet today's surgical trainees face an insurmountable challenge. Reforms in surgical education have led to a shift from the traditional 'Halstedian' apprenticeship model a result of regulations imposing reduced working hours for doctors[13]. In addition, increased medico-legal claims[14], public scrutiny[15] and the continuous need to improve outcomes have stifled training opportunities. Understanding the glaring need to enhance surgical skills acquisition, methods have been identified to assess and enhance technical performance[16-18], which are outlined in Chapter 2. Although skills acquisition in surgery has been proposed to adopt motor learning models[18], the emphasis has been on surrogate markers of learning. As a result several issues such as overlapping performance[19], inability to transfer skills[20], the effect of feedback and differential requirements for effort invested remain unexplained[21].

Motor learning represents an important and well-researched neuroscience topic[22], which has several applications towards enhancing training and identifying hallmarks of expert performance in healthy individuals whether it is in the field of aviation, music, sport or dance[23]. Established principles, models and factors affecting motor learning are described in Chapter 2.

Concurrently, the field of neuroscience has been revolutionized with the advent of functional neuroimaging, which has now been established as the predominant investigatory technique for behavioural and cognitive neuroscience[24]. Significant insights have been drawn from functional neuroimaging studies, which have enabled mapping of neural correlates that underlie novice progression to expert motor performance[25]. This is possible because practice leads to dynamic changes in functional and structural patterns of the brain, termed as ‘neuroplasticity’. In addition the emergent field of ‘neuroergonomics’, forged by the convergence of two fields, namely neuroscience and ergonomics has enabled unique insights into understanding brain behaviour during interaction with complex instrumentation. This is of particular relevance as the environment and instrumentation surgeons’ interact with have become technologically more complex. Despite availability of investigatory techniques, research into skills acquisition requisite for surgeons remains parsimonious. Perhaps, it is a consequence of the challenges faced in implementation of a research design that incorporates the contextual features necessary to study a surgeon’s brain in appropriate operative environments. Therefore, it is imperative to select the optimal brain regions and neuroimaging modality on the basis of function and feasibility, which is explained in Chapter 3. The justification for selecting functional near-infrared spectroscopy and motor regions in the frontal lobe as the neuroimaging modality and neural correlates, respectively are provided.

Almost a decade has passed since functional neuroimaging has been used to interrogate assessment of various factors that underpin surgical performance. Chapter 4 provides a systematic review on the application of functional neuroimaging in assessment of skill, training induced changes and non-technical factors that impact technical performance.

According to the purpose of this thesis, Chapter 5 (first experimental) employs a cross-sectional design to interrogate frontal lobe neuroplasticity for an advanced surgical skill. Participants are categorized into three groups based on expertise-level of skill. The chosen task is laparoscopic intracorporeal suturing and knot tying (LICS) a key component of LS curriculum[26] where proficiency in its performance serves as a barometer to performing complex laparoscopic procedures[27]. The frontal lobe is the focus of brain interrogation, as it contains primary and secondary motor regions whose engagement is dependent on the phase of skill acquired. Laparoscopic experience is investigated as a factor for differential brain activation during task performance.

Frontal lobe neuroplasticity associated with practice related improvement in performance utilising a longitudinal design is described in Chapter 6. Novice participants are provided intensive distributed training for LICS task until they reach expert-levels of proficiency. They are serially monitored at regular intervals and following cessation of practice to track learning related changes and retention. Despite novices attaining expert-level proficiency on currently validated benchmarks, practice-dependent attenuation in an associative motor region responsible for attention was not observed, contrary to motor learning literature. These results suggest that reduction in attentional demand during learning of advanced motor skills do not necessarily accompany observed improvement in technical performance.

Mapping functional activation can either follow ‘functional segregation’ or ‘functional integration’ principles[24], where the former describes brain activity in isolation and the latter presents neuroimaging data of brain regions interacting with each other. In recent times within the functional neuroimaging community there has been a shift towards interpreting data using connectivity analysis[28]. Aligning with this trend, in Chapter 7 expertise-classification for the same task (LICS) is described using functional connectivity analysis of motor regions in the frontal lobe. More importantly an algorithm for automated discrimination of expertise level using fNIRS data is described.

The thesis is concluded in Chapter 8 with a summary of its achievements, limitations and future directions arising from the studies conducted in this thesis. A potential application of functional neuroimaging is towards automated detection of the operator’s cognitive state induced by task complexity. In chapter 8, the feasibility of Electroencephalography (EEG) towards automatic task load detection in trainee surgeons via a passive-brain computer interface (BCI) is introduced.

Original contributions of this thesis are as follows

1.) Evidence synthesis of the applying functional neuroimaging towards assessment and enhancing operative and non-operative performance in the field of surgery.

2.) Deriving and testing a hypothesis based on established current motor learning literature with regards to mapping frontal lobe neuroplasticity associated with practice and/or expertise on a complex bimanual motor task. The chosen motor task is complex as it not only tests motor sequence learning but also requires sensorimotor adaptation induced by sensorimotor distortions. In accordance with motor learning literature, practice led to attenuation in prefrontal cortex (PFC) activity and increase in activation of secondary motor regions namely the dorsal premotor cortex (PMC) and supplementary motor area (SMA).

3.) The feasibility of fNIRS, a non-invasive functional neuroimaging tool to map multiple functional brain regions situated in the frontal lobe associated with learning complex motor skills required for MIS. The chosen task (LICS) was deconstructed into three sub-tasks to allow feasible scrutiny based on a functional neuroimaging experimental paradigm. Task deconstruction also enabled sub-task comparison based on directionality.

4.) Investigation of expertise related disparity in frontal lobe activation using a cross-sectional experimental design. Commensurate with motor learning literature, novices in comparison to trainees and experts displayed greater PFC excitation, and lower PMC and SMA activation. The advantage of measuring multiple frontal regions enables better classification thus overcoming the limitations of a similar study, which measured the PFC in isolation[29].

5.) Application of graph network analysis for fNIRS data to enable expertise classification. Brain networks derived from functional connectivity of fNIRS data was used to compute network econometrics. In accordance with network science, experts displayed a more mature cortical network than novices measured by small-world index.

6.) To investigate practice dependent neuroplasticity by employing a longitudinal design. Skills retention was measured one month after cessation of practice. Practice led to improvements in technical performance, an increase in activation across secondary motor regions but not concurrent attenuation in PFC activity. The persistently raised PFC activity is an interesting feature, suggesting continuous need for attention to support technical performance despite improvement in it. It highlights the potential of neuroimaging as an adjunct to measure development of expertise during complex motor skills acquisition in surgery.

7.) Application of machine-learning algorithms for functional connectivity data towards enabling automated expertise level classification. Based on seed-based functional connectivity analysis, novices were found to be more dependent on interaction between associative and secondary motor regions (PFC-PMC, PFC-SMA), and between secondary motor regions (PMC-SMA). Associations between primary motor cortex (M1) and SMA, an area associated with storage of learnt sequential skill was observed to be greatest amongst experts. These finding are commensurate neuroplasticity associated changes in motor loops.

The work presented in this thesis has resulted in the following publications in peer-reviewed journals and presentations at international conferences:

Papers

- Shetty K, Leff DR, Orihuela-Espina F, Yang GZ, Darzi A. Persistent Attentional Demands Despite Laparoscopic Skills Acquisition. *JAMA Surgery* 2016 Jul 1; 151(7): 682-4.

- Andreu-Perez J, Leff DR, Shetty K, Darzi A, Yang GZ. Disparity in Frontal Lobe Connectivity on a Complex Bimanual Motor Task Aids Classification of Operator Expertise. *Brain Connectivity* 2016 Jun; 6(5): 375-88.

- Crewther B, Shetty K, Jarchi D, Selvadurai S, Cook CJ, Leff DR, Darzi A, Yang GZ. Skill acquisition and stress adaptations following laparoscopic surgery training and detraining in novice surgeons. *Surgical Endoscopy* 2016 Jul; 30(7): 2961-8.

- Zander T, Shetty K, Lorenz R, Leff DR, Krol LR, Darzi AW, Gramann K, Yang GZ. Automated Task Load Detection with Electroencephalography: Towards Passive Brain-Computer Interfacing in Robotic Surgery. *JMRR* (Accepted May 2016).

Presentations at International Conferences

- Shetty K, Leff DR, Yang GZ, Darzi A. Beware Credentialing Based on Fundamentals of Laparoscopy: "The Eyes are Open, The Hands Move, But the Prefrontal Brain Has Not Departed". Presented at the *8th Hamlyn Symposium 2015, London, UK*.

- Shetty K, Leff DR, Orihuela-Espina F, Selvadurai S, Chaudery M, Athanasiou T, Yang GZ, Darzi A. Looks can be deceiving: Is there value in cognitive assessment of complex surgical skills? *SARS 2015*, Durham, UK.

- Orihuela-Espina F, Leff DR, Herrera-Vega J, Shetty K, James DRC, Darzi AW, Yang GZ. Semi-virtual registration and virtual channel synthesization in fNIRS imaging. Poster presentation at *fNIRS 2014*, Montreal, Canada.

- Shetty K, Leff DR, Andreu- Perez J, Orihuela-Espina F, Sriskandarajah K, Athanasiou T, Darzi A, Yang G-Z. Long term complex motor skills training and evolution in small network world topology. Poster presentation at the *Organization for Human Brain Mapping Conference* 2014, Hamburg, Germany.

- Shetty K, Leff DR, Orihuela-Espina F, Yasuura S, Sriskandarajah K, Dunne R, Andreu-Perez J, Darzi A, Yang G-Z. Clinical Decision Making: Does prefrontal cortex engagement vary with expertise and task difficulty? Accepted as a poster presentation to the *Organization for Human Brain Mapping Conference* 2014, Hamburg, Germany.

- Shetty K, Leff DR, Orihuela-Espina F, Yasuura S, Sriskandarajah K, Cundy TP, Yang GZ, Darzi A. An analysis of prefrontal engagement in advanced laparoscopic skills: resident's do, attending's don't, novices won't. Poster at the *99th Annual Clinical Congress of the American College of Surgeons*, October 2013, Washington DC, USA

- K Shetty, T Zander, DR Leff, R Lorenz, GZ Yang, A Darzi. Automated cognitive load detection with Electroencephalography: Towards Brain-Computer Interfacing in Robotic Surgery. *6th Hamlyn Symposium on Medical Robotics*, London, June 2013. Prize winning presentation.

- K Sriskandarajah, K Shetty**,** M Sodergren, Yang G-Z, Darzi A. Performance and eye behaviour changes associated with visuomotor rotation- Relevance for design of robotic telemanipulators. *6th Hamlyn Symposium on Medical Robotics*, London, June 2013.

- Shetty K, Leff DR, Orihuela-Espina F, Sriskandarajah K, Yasuura S, Cundy TP, Yang GZ, Darzi A. Expertise related disparity in executive control and motor planning on complex surgical skills. Poster at *Organization for Human Brain Mapping 2013*, Seattle, USA.

- Leibrandt K, Wasuntapichaikul P, Lin J, Cakir A, Shetty K**,** Sriskandarajah K, Leff DR, Sodergren M, Darzi A, Yang GZ. Classification of safe and erroneous surgery using visual gaze behaviour and cortical brain function. Poster at *Organization for Human Brain Mapping 2013*, Seattle, USA

- Shetty K**,** Leff DR, Orihuela-Espina F, Koshy A, Sriskandarajah K, Darzi AW, Yang GZ. Effects of visuomotor rotation in laparoscopic surgery on the Prefrontal Cortex. Poster at *fNIRS 2012***,** London, U.K**.**

- Shetty K**,** Leff DR, Orihuela-Espina F, Cundy TP, Darzi AW, Yang GZ. A study of executive control during intracorporeal minimally invasive suturing (ICS). Poster at *fNIRS 2012***,** London, U.K.

- K. Shetty, D.R. Leff, F. Orihuela-Espina, A.W Darzi, G.-Z. Yang. A Study of Executive Control during Intracorporeal Minimally Invasive Suturing using Functional Near Infrared Spectroscopy (fNIRS). Poster at *5th Hamlyn Symposium* July 2012. London U.K.

Chapter 2

Motor Skills Assessment in Surgery

2.1. Introduction

Master surgeons, like expert performers in sport, music or dance display a remarkable ability to perform to skilled motor routines comprising complex movements in repeated, effortless fashion. Manual dexterity is critical to operative performance, and is attained by well-rehearsed deliberate practice[30]. This concept known as the apprenticeship model was popularized by Sir William Halsted an eminent surgeon at John Hopkins Hospital[31] and has been the cornerstone for surgical training for over a century[18]. Unsurprisingly, it takes an estimated 15,000 to 20,000 hours to train a surgeon based on this model[13], a notion which resonates with Ericsson's theoretical model of expertise that it requires 10,000 hours of intensive practice over ten years to achieve elite performance ability[32]. Although parallels have been drawn between surgeons and other expert performers such as musicians[17] or athletes[33] in the pursuit of development of expertise[34], surgeons face inherent challenges in accessing training opportunities and more importantly training commences only in their adult life. Musicians and athletes have the advantage that they often commence training during childhood or adolescence, a stage when brain development is greater.

Surgical training in recent times has undergone major reform in Europe[35] and North America[35]. Regulations on working hours for doctors has led to reduction in weekly working hours that vary based on global location from 80 hours a week in the United States to 60 hours a week in Australia[36] and 48 hours a week in Europe[13]. Ultimately reduced working hours equates to decrease in training opportunities, requiring a radical upheaval of the 'volume-based' training model. The advent of MIS, has revolutionized the field of surgery yet imposes greater training needs due to the increasingly complexity of the task, novel instrumentation and environment. Additional pressures are increasing public scrutiny[15], medical litigation[14] and the need to balance improving patient safety and outcomes.

Considering the magnitude of the challenges in surgical training[13], objective measurement of motor skill is necessary to monitor progression throughout surgical training and credentialing. In order to better track progression of learning, which by definition is a set of not directly observable, highly complex internal processes occurring in the mind of the learner, greater emphasis might be paid to understand the neurocognitive processes that underlie complex motor skills acquisition[18]. In this chapter, the principles and models of motor learning are described. The latter half of the chapter describes methods to assess level of technical surgical skills acquisition.

2.2. Motor skills learning

2.2.1. Motor Skill has several key components[22]:

1. It requires a desirable '*environment goal',* which differentiates it from a 'movement'. For example the goal in basketball can be shooting a free throw, which differs it from idly moving one's hand.
2. The second feature involves meeting this environmental performance goal with '*maximum certainty'*. For example, a basketball player might successful shoot a free throw without much certainty, a chance result amongst numerous unsuccessful attempts. In order to be regarded as 'skilled' it requires the player to consistently reproduce the skill during high-pressure situations, the hallmark of a champion athlete.
3. Conservation of energy is critical for sustained skilled performance, thereby making '*efficiency'* the third feature, a requisite for most but not all skills. This does not apply merely to physiological or physical energy but also incorporates cognitive energy. For example many skills are well honed where the individual pays minimal attention to the action, often described as '*effortless performance'*.

4. The final feature is *‘time’* required to complete the goal driven task. For example expert surgeons are capable of maintaining accuracy, yet also perform technical tasks more rapidly than the inexperienced.

Thus, motor skill can be defined as the ability to achieve a target end result with maximum certainty in minimal time and with minimal expenditure of physical and cognitive energy[22].

2.2.2. Motor Learning

Motor skill learning refers to the process by which movements are executed more quickly and accurately with practice[37]. Schmidt et al[22], defined motor learning as a process incorporated by several key components, namely:

1. It affects ‘capability’ or capacity for skilled performance.
2. Learning occurs from training or experience.
3. Learning is not directly observable as it leads to alterations internally. These internal processes are highly complex phenomena occurring in the central nervous system that are particularly difficult to study. Instead learning is often assumed from identifiable improved performance derived from kinematic assessment.
4. Changes brought about by learning are not transient but relatively permanent in the capability for skilled performance.

Thus in summary motor learning can defined as a “set of processes associated with practice or experience leading to relatively permanent gains in the capability for skilled performance” [22].

2.2.3. Classification of motor skills

Classification is necessary in order to understand the laws of motor behaviour. It can be classified based on type of movements made or environmental predictability of the task[22].

2.2.3.1. Classification based on type of movement

1. **Discrete:** These movements have a recognizable initial and end points for example shooting a basketball through the hoop. Several surgical tasks are discrete such as cutting with a knife or scissor.

2. **Continuous:** Movements that do not have a recognizable beginning and end are defined as continuous movements, for example running or swimming. In the laboratory they are often tested as tracking tasks, which are further sub-divided as pursuit and compensatory tracking. Pursuit tracking tasks essentially involves concurrent display of intended target and performer's movement responses. Compensatory tracking differs in that the intended target and the performer's movement are combined to display a single value. A common example of a pursuit tracking task is driving or in the case of surgery, endoscopic or laparoscopic camera navigation.

3. **Serial:** Composed of a series of individual movements laced together; such as the classic example of playing a piano. In the field of surgery, suturing can be classified as a serial movement as it requires gestures that require stitching and knot tying usually in a set sequence.

2.2.3.2. Classification based on environment predictability

Performer's perception of the environmentally is the basis of classification into closed or open skills.

1. **Closed:** Movements in stable environment conditions are defined as "closed". An example of a closed motor skill is shooting a free throw basket from the foul line where the environment is stable because of intermission of play. Surgical skills developed during simulation are examples of closed motor skills as the environment in which they are practiced is controlled.

2. **Open:** These movements occur in situations in which the environment is dynamic, at times unpredictable, requiring adaptation and does not permit the performer to effectively map out the entire movement in advance. Examples of open skills are shooting a basket against an opponent or surgical manoeuvres during live abdominal surgery where variations in human anatomy are prevalent. The performer's ability to adapt according to the situation is key for success in open motor skills.

2.2.4. Models of Motor skills acquisition

Stages of motor learning was first described by Bryan and Harter[38] in 1899 by examining skilled movements acquired by telegraphic operators. Several models of motor learning have been later proposed from the two stage views of Snoddy[39], Adams[40] and Gentile[41] to three stage view proposed by Fitts[42] and Anderson[43] and five stages by Dreyfus[2]. Considerable overlap is present in the descriptions of various models, therefore only the most salient models are (see Figure 2.1) are described below. The stages should not be viewed as discrete boundaries but more as fuzzy borders[44].

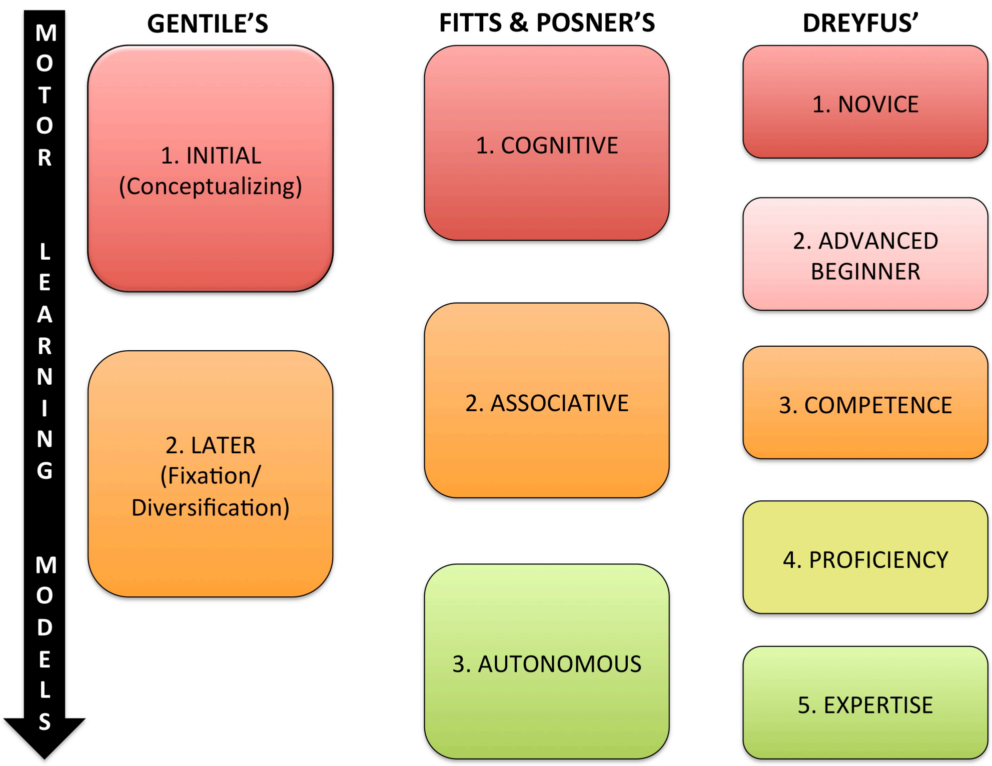

Figure 2.1. Colorimetric illustration of the motor learning stages according to the Gentile, Fitts and Posner and Dreyfus models.

2.2.4.1. Fitts and Posner's three stage learning model

The model of motor learning proposed by Fits and Posner in 1967[42], theorizes progression from novice cognitive phase to a learner's associative phase and ultimately to an expert autonomous phase.

1. **Cognitive:** At the initial cognitive phase, the novice focuses on the elements of the motor task, is sensory feedback dependent and is outwardly observed by the large number of errors and performance is erratic.

2. **Associative:** In this phase the learner after practice has developed the ability to integrate sensory information with appropriate goal directed movement and link movements. They are capable of appraising their performance, in the process eliminating their mistakes and refining skills. Learners' at this stage are still attentionally dependent whilst developing strategies to improve performance. Performance is characterized by reduced variability and few errors than the cognitive phase.
3. **Autonomous:** On reaching the final phase motor tasks become automatic, characterized by consistently high performance with minimal errors, variance in performance and mental load. Experts are capable of making subtle adjustments to performance. Performance in this phase is portrayed by effortless skill, which is stable despite distractions, manifest by ability to multi-task with minimal interference. This ability described later as 'automaticity'[45,46] is therefore representative of achieving expertise.

2.2.4.2. Gentile's two stage learning model

Gentile's motor learning model[41] is similar to the earlier described Fitts' model[42], yet is simplified as follows:

1. **Initial** (getting the idea of the movement): At the beginning, the learner tries to understand the movement requirements by practicing various patterns of movement. Learners must also understand and discriminate between environmental factors that are inherently related to successful motor performance (regulatory) and those that are not (non-regulatory). In other words the beginner must use selective attention on aspects

that are requisite for successful performance.

2. **Later** (Fixation for closed motor skills; Diversification for open motor skill): During this stage the learner focuses on refining the acquired skills, dependent on the type of motor skill. Fixation applies to 'closed skills' where the skill is replicated consistently and accurately in a stable environment. 'Open skills' development unlike closed occurs in a dynamic, unpredictable environment where the learner learns to diversify the skill in order to adapt.

2.2.4.3. Dreyfus and Dreyfus' five-stage model of skill acquisition

This learning model was first proposed by the Dreyfus and Dreyfus in 1980[2] has undergone subsequent modifications[47,48]. It explains progression through a series of five levels from novice to expert and has recently received attention from surgical educators[49] and sports scientist[50]. The five levels are as follows:

1. **Novice:** At the very beginning the task is decomposed into context-free features and the novice is provided rules to determine actions based on set features. The novice is rigidly adherent to taught rules with no exercise of discretionary judgement. For example, the novice merely adheres to instructions received on how to mount and manipulate a needle during a suturing task.

2. **Advanced Beginner:** During this stage, with practice the individual begins to show situational awareness, understands the context, recognizes new aspects and treats all aspects with equal importance. For example in a suturing task the beginner not only focuses on how to manipulate a needle but also the amount and symmetry of tissue taken in each bite based on the type of tissue.

3. **Competence:** The beginner with practice learns to appreciate more aspects relevant to the task and paying attention to them all becomes exhausting. At this stage, the beginner learns to cope with crowdedness by developing plans and routines in selecting salient aspects determined by the situation. The performer can no longer rely on rules alone but will have to use judgement based on a situation. For example the operator will now have to select the right amount of tension to oppose sutured ends in a suturing task.

4. **Proficiency:** At this stage the proficient performer becomes embodied to the task and gains a holistic view of the situation. Actions are no longer dictated by rules but replaced by situational awareness drawn from emotional experiences of successful and unsuccessful outcomes. However the proficient performer will still rely on analytic decision-making, having not yet achieved sufficient experience to be capable of automatic response based on an unfamiliar situation. Put simply, the proficient performer recognizes what action needs to taken based on the situation but deliberates on how to achieve it. For example, the proficient surgeon will instantly recognize bleeding but will decide the best way to control it.

5. **Expertise:** The expert is different from the proficient performer in being able to immediately decide how to achieve the goal based on a situation. This ability is developed from an extensive experience of a variety of situations and being able to draw from those experiences. For example, the expert surgeon will rapidly recognize bleeding in the operative field, judge its significance and if required will intuitively select the best action to control it.

Skill level	Component	Perspective	Decision	Commitment
1. Novice	Context free	None	Analytic	Detached
2. Advanced beginner	Context free and situational	None	Analytic	Detached
3. Competent	Context free and situational	Chosen	Analytic	Detached understanding and deciding; involved outcome
4. Proficient	Context free and situational	Experienced	Analytic	Involved understanding; detached deciding
5. Expert	Context free and situational	Experienced	Intuitive	Involved

Table 2.1. Dreyfus five-stage model of skills acquisition[51]. Component refers to elements of the situation the learner is capable of perceiving. Perspective refers to the ability to selectively focus on recognized components. Decision refers to the type of decision-making ability based on experience, and decision describes the degree to which the leaner is immersed in the task.

2.2.5. Motor Learning Theories

2.2.5.1. Closed Loop Theory

In 1971 Adams[40] based on subjective reinforcement described that movements made during practice were stored in memory and compared to a reference of correctness termed as 'perceptual trace'. Learners actively engaged in verbalization of the motor task, each trial of practice led to increased consistency and correct movements. Each correct movement led to a strengthening of the correct trace and reduction in relative strength of incorrect perceptual trace (see Figure 2.2). Conversely, errors produced from practice were harmful to learning as they not only increased the strength of the incorrect perceptual trace but also relatively degraded the correct perceptual trace. In order for the system to function two memory states must occur, one to produce the action and the other to evaluate the outcome.

Criticisms to the closed loop theory were firstly practising movements that differed from the perceptual trace (variable practice) were found to be superior to repeated practice of only the correct movement. Secondly, contrary to the notion that sensory feedback from the limb alone leads to increase in perceptual trace, animals with no sensory feedback from the limbs were capable of skilled movement and learning new actions too. Finally this theory did not differentiate between slow and fast movements, which had different error-detection mechanisms.

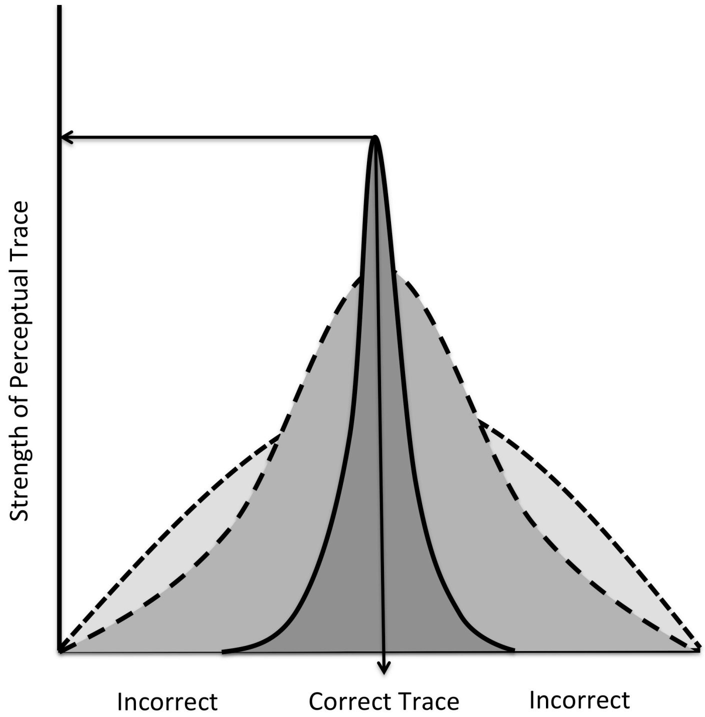

Figure 2.2. Line plot graphical illustrates adaptation of Adam's closed loop theory[40] of practice related increase in perceptual trace strength from early stages (dotted lines with shaded in lighter grey) to late stage (bold line with area shaded in dark grey).

2.2.5.2. Schema Theory

The schema theory developed by Schmidt in 1975 retains several of the strengths of the closed loop theory and replaces its defective concepts[52]. According to the schema theory the ability to perform a movement is dependent on three factors as follows:

1. A generalised motor program (GMP)
2. Recall schema
3. Recognition schema

Movements are first made by selecting a GMP, i.e. a basic pattern of movement with invariant features like timing, then adding other features required to execute the movement in a particular way. After selection of GMP, four types of information are available for brief storage in short-term memory namely:

i. information about initial conditions (e.g. body, object properties)
ii. parameters assigned to GMP
iii. augmented feedback about the movement outcome
iv. sensory consequences of the movement

These sources of information are stored long enough that the performer can abstract two schemas. Recall schema provides parameters to the GMP after assessing the current situational and intentional factors (e.g. for a basketball pass it would include height, speed, distance, spin, etc). Recognition schema is responsible for inherent feedback by movement evaluation. It is thought to be formed by integration of initial conditions, environmental outcomes and sensory consequences. Augmented feedback is critical to schema learning despite the presence of other sensory information. Unlike in the closed loop theory, different movements despite the production of errors are considered beneficial justifying variable practice.

2.3. Types of Practice

The amount of practice is thought to be the single most important variable in motor learning. Contrary to popular notion, protracted practice leads to continued performance improvements observed by Crossman's[53] experiment on cigar rollers with experience of seven years, and ten million cigars rolled. Practice related performance improvements are better identifiable using logarithmic relationships[22].

2.3.1. Deliberate practice

Ericsson et al[54] defined deliberate practice as activities that are specially designed to improve current level of performance. It incorporates activities designed by trainers that requires effort and are not inherently enjoyable. Deliberate practice is not only important to attain expertise but also to maintain it. Deduced from retrospective and longitudinal studies of musicians and athletes, Ericsson[34] estimated to master a skill it would require 10,000 hours over ten years, driven by a determination to improve. This concept has recently received attention from surgical educators in order to enhance MIS skills acquisition by incentivising simulation[17,18,55]. Other factors that contribute towards learning can classified based on physical interaction with the task as 'on-task' and 'off-task'. They are as follows

2.3.2. On-task attributes

1. **Distribution of practice** refers to time scales and frequency of training sessions interspersed with rest sessions. When practice sessions are "massed" i.e. delivered temporally close to one another with either very little or no rest periods between, it is referred to as *'mass practice'*. Conversely, *'distributed practice'* refers to training that occurs over sessions with larger and more frequent rest periods in between. Distributed practice also referred to as interval training has shown to be superior for MIS skills acquisition[56,57], an observation similarly seen in a golf putting task[58].

2. **Variability** of practice is essentially composed of practising a number of task variations. It plays a greater role in development of open skills than closed skills where the environment is stable and does not require adaptation. In addition practice variability is considered beneficial to transference and retention of learnt skill, which are measures of learning.

3. **Practice schedules** are referred to as *'blocked'* when all trials on one task are practiced together uninterrupted by another task. On the contrary, '*random practice'* involves practice of multiple tasks where the same task is rarely practiced consecutively. A study comparing blocked versus random practice[59] unsurprisingly demonstrated that blocked practice led to earlier and greater improvements in performance, but the gap between the two reduced in later practice trials. Surprisingly, random practice led to greater performance at retention, i.e. measure

of learning observed by performance after a prolonged period of no training. This effect coined as *'contextual interference'* was first defined by William Battig[60] as interference generated due to the context in which skills were practiced, produced decrements in performance during practice, but made the learning of these tasks effective. Contextual interference is considered to be also beneficial for transference of skill, however its beneficial effects are not generalizable as they are task and subject dependent.

4. **Part versus Whole Practice:** When tasks are complex and cannot be grasped as a whole, a strategy is required in order to break them down into smaller components. Effectiveness of part practice is measured by the amount of transfer to performance of the whole task. Skill complexity and organization determines whole or part practice effectiveness[61]. Skills complexity refers to the number of components in a task, and organization refers to the relationship between spatial and temporal components. High organization refers to a higher inter-dependency between spatial and temporal relationships for example when landing a plane; the pilot must select appropriate movements in a timely manner to safely land the aircraft. Whole practice is recommended for skills that are high in organization, whilst part practice is recommended for tasks that are highly complex but low in organization. Surgical tasks are complex in nature and yet when distilled are composed of a series of discrete manoeuvres. Dubrowski et al[62] observed part practice in random order was superior to blocked part practice to acquire a complex bone-plating surgical task. This concept is often used in surgical training techniques when the trainee performs part of the operation at earlier stages followed by whole performance at later stages.

5. **Guidance:** is an assistance technique employed to constrain incorrect manoeuvres and direct correct ones. In surgery guidance can be physical for example when the trainer directs the trainees gestures to prevent incorrect ones or sensory for example the scope guide which aids endoscopic performance by providing visual representations of the colonoscope. Guidance in surgery may also serve to mitigate fear of causing inadvertent injuries from incorrect gestures. Although it is useful at

early stages of learning, temporarily boosting performance, prolonged usage can lead to reliance on explicit learning and therefore limits development of an internal model driven by implicit learning, an effect observed by two experiments which examined MIS skills learnt under augmented visual feedback[63,64].

2.3.3. Off-task attributes

1. **Motivation** is an important factor to ensure engagement resulting in greater learning. Methods to improve motivation are by stressing the importance of the task and goal setting.

2. **Verbal information:** Providing initial verbal cues to learners is important for orientation and has the ability to target improvement according to verbalised parameters. For example, Solley et al[65] observed that participants in accordance with verbal instructions were faster or more accurate when one attribute was stressed upon, or demonstrated intermediate performance on either measure when stressed upon the importance of both. Moreover, a lasting effect was observed when participants were additionally trained under common instructions to focus on both speed and accuracy[65]. The benefits of expert instruction for a surgical knot-tying task was detected by superior performance when repeat tested at retention[66].

3. **Perceptual learning:** Exposure to environmental aspects encountered during execution of task is useful as a preparatory technique in order grasp spatio-temporal relationships. Observational learning is a kind of perceptual learning where the task to acquire is demonstrated before practice, e.g. a surgical trainee often views an operative procedure before executing one. However observation of expert performance alone gains relatively little insight in comparison to performance accompanied by knowledge of results i.e. correctness of action.

2.4. Motor Memory, Retention and Transfer

Motor memory is defined as the persistence of the acquired capability for performance. Based on theoretical concepts, memory can be a motor program, reference of correctness or schema. *Forgetting* is the opposite of learning, referring to the loss of capability. At a behavioural level it is evidenced by the relatively permanent losses in performance.

Retention is tested after at a time interval (*retention interval*) after cessation of practice. At a behavioural level, it refers to the persistence or the lack of the performance arising from maintenance of memory or forgetfulness. Test of transfer also tests persistence of acquired capability for performance, but differs in that it involves testing the same task under different conditions. For example, laparoscopic suturing learnt and then subsequently tested under controlled simulated conditions examines 'retention', but when the same laparoscopic task is tested using a different model or under different conditions such as in the real live OR it tests 'transference'[67].

2.4.1. Measures of Retention and Transfer

Identification of factors influential to retention and transfer are critical to motor learning. The most common methods to measures retention are '*absolute'* and *'relative'* retention. As the term suggests, absolute retention is defined as the level of performance on the trial of a retention test. On the other hand, relative retention can be measures as a difference (difference in scores between end of training and beginning of retention), percentage (difference in scores between end of training and beginning of retention divided by difference in scores of beginning and end of training) or savings. The savings score can usually be estimated in relearning trials after a retention interval deduced by the number of trials required to return to the level of proficiency achieved at the end of the original practice.

2.4.2. Simulation and Transfer of skills

Simulation represents an important and safe method of training motor skills. Simulation models provide an organized context to rehearsing procedural skills i.e. sequential motor skills. Similar to training of pilots using virtual cockpits to enhance procedural skills, surgical educators[17,18] have recommended learning of MIS skills must be initially undertaken in simulated conditions, thereby mitigating harm to patients that would have otherwise been caused by errors associated with early stages of motor learning. For example, Gallagher et al[68] demonstrated in two related experiments that simulation training leads to enhanced performance on live surgical tasks in laparoscopic trainees and novices. Simulators take the form of a conventional physical box such as a box trainer or computer simulation such as virtual reality. Physical simulators can be large and lack high fidelity in recreating perceptual-motor attributes of the motor task. Virtual systems can overcome these limitations and are capable of creating environments that replicate sensory motor attributes. Trials comparing surgical skills learnt under physical and virtual simulators were inconclusive[69-71] showing no clear benefit of one over the other.

2.5. Attention underlying motor performance

The subject of attention is a vast research area and is only selectively covered in this section to highlight the important link between attention and motor performance.

2.5.1. Attention as effort

Complex tasks are considered more effortful or mentally demanding. Measuring intensity of effort has been the basis of measuring task complexity. Initial methods used simpler physiological markers namely pupil diameter, skin resistance, heart rate and subjective rating scales followed by later methods that measure neurophysiological signals from eye and brain activity[72].

2.5.2. Attention as a Resource

A common notion is that attention is a finite resource, limiting the ability to absorb and perceive all aspects in the environment. Its relevance is particularly important whilst multi-tasking as each task competes for common attentional resources. Performance may regress in either one or both tasks, representative of different patterns of interference when the combined attentional demand exceeds the available resource. Interference can be *'structural'* when tasks compete for physical capacity i.e. both require hand/eye movement or alternatively *'functional'* when the task competes for central (attentional) capacity. Research into the impact of interference on operator performance has enabled identification of safer solutions in the field of aviation[73]. Whilst operating surgeon too are engrossed with multiple cognitive processes that are necessary for various functions ranging from generation of skilled purposeful movements to decision-making and vigilance. Usage of complex MIS tools and inexperience of the operator combined with aforementioned factors could exceed the attentional capacity threshold leading to critical issues in safe surgery[74].

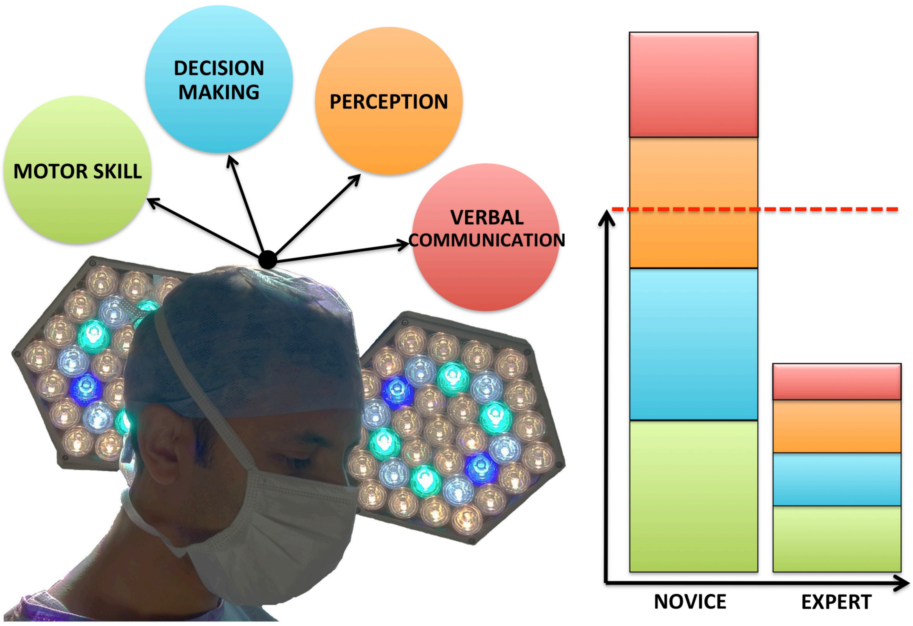

Figure 2.3. Cognitive demands placed on the operator's attentional capacity arise from executing skilled movements, intra-operative decision-making, sensory perception from multiple stimuli generated from the operating environment and verbal communication the OR team. On the right is a novice versus expert hypothetical comparison depicts that the competing cognitive demands exceed the attentional capacity threshold (dotted red line) for a novice but not for an expert surgeon. (Figure adapted from adapted from Gallagher et al[74]).

2.5.3. Selective Attention and Inattention Blindness

Selective allocation of attention is a concept closely linked with interference, can be either '*intentional'* where the individual attends to one whilst ignoring the other or 'incidental' when attention is captured involuntarily in response to an external stimulus. The former occurs by top-down processing and the later driven by bottom-up processing.

Selective attention may be beneficial when trying to focus on a task and ignore distractions. At times this coping mechanism can have adverse outcomes especially to safety in high-risk fields when a critical sensory event is missed. The phenomenon of "inattention blindness" or "change blindness" is best illustrated by an experiment by Simon and Chabris[75]. Participants were required to watch a video and count the number of times three humans in white shirts pass a basketball amongst three other humans dressed in black shirts. The experiment revealed that participants engaged in the counting task selectively ignored humans dressed in black, as a result missing an event lasting several seconds where a man dressed in a costumed gorilla entered into the centre of the video amongst the players, beat its chest and exits.

2.5.4. Attentional focus and Stress

The focus of attention during movement is 'internal' when it centres on an aspect of motor skill such as execution or feedback derived from sensory consequences of the movement. At other times when attention is directed to an expected outcome it is classified as 'external'. Research in sports movement science have identified that the internal focus has beneficial effects on novice golfers but detrimental effects on experts. Conversely external focus for example focusing on the outcome of a golf putt had a beneficial effect on experts but not in novices[76,77]. Degradation of performance by experts is also observed under stressful situations. For example some athletes '*choke'* under pressure, which can be explained by the reinvestment theory[78], which suggest that whilst the skill is learnt to an advanced level, it is disrupted when the performer reverts to a state of conscious movement control akin to internal focus.

2.6. How to measure motor learning?

2.6.1. Performance curves

Performance curves plot performance scores of a group of individuals over successive trials of practice. Dependent on the performance measures recorded, curves may show upward or downward trends. However performance curves cannot be assumed to reflect learning, which essentially is a change in internal state and is not directly observable. Deriving inferences from performance curves is problematic as it represents average group performance but does not account for inter-subject variability or within subject variability. Different subjects show different rates of performance improvement and a given participant's performance varies from one trial to the next. In motor learning tasks where absolute scores exist, performance curves suffer from ceiling and floor effects. Examples of absolute scores are no fewer than zero errors on a trial or no more than zero seconds taken for time to execute a task. Ceiling and floor effects represent corresponding limitations at the top and bottom of the performance scale. As an individual reaches these scalar boundaries, changes in performance levels becomes increasingly insensitive to changes in learning that continue to occur.

2.6.2. Alternative methods (Secondary tasks, Automaticity and Effort)

Despite displaying maximal performance (ceiling effect), further training also known as 'overlearning' may lead to continued learning and greater levels of proficiency not identifiable by simple performance curves. In subjects who display maximal performance, introduction of a secondary task may lead to interference and ultimately degradation in performance of the primary task. On the other hand, overlearning leads to skills becoming more automatic and less susceptible to interference from secondary tasks, thereby displaying stable maximal performance in the primary task and practice related improvements for the secondary task. The principle of dual tasking has increasingly been adopted to delineate expertise or capability between individuals who exhibit identical peak surgical performance[79,80].

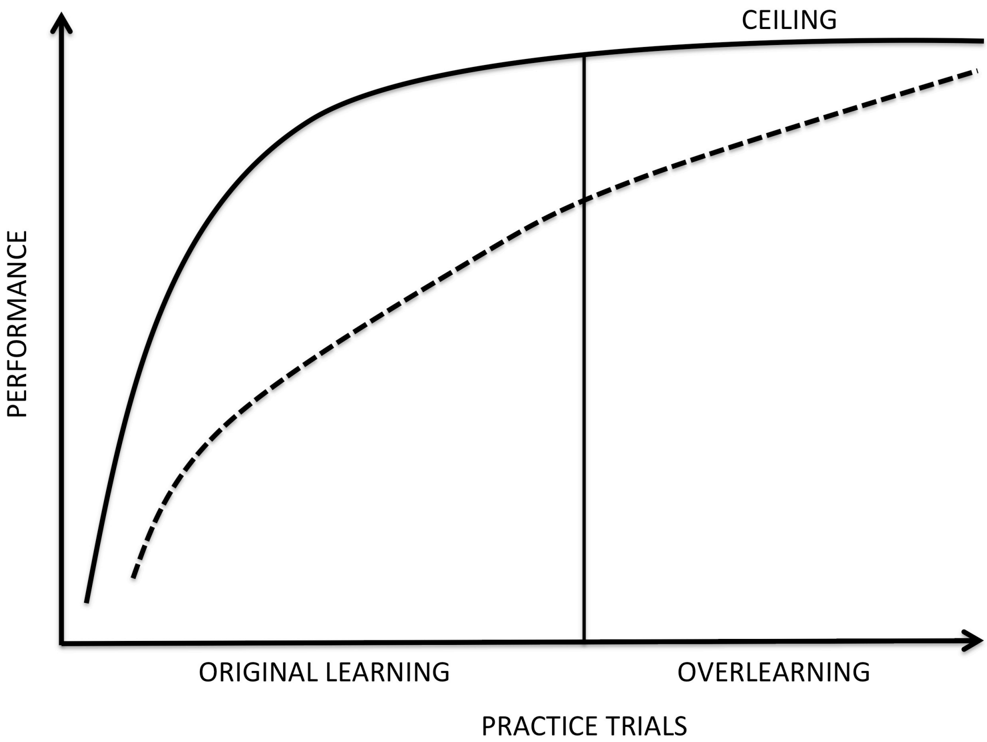

Figure 2.4 Line graph illustrates practice related improvement in performance (y axis) of a primary task (bold black line) until maximal performance in achieved (ceiling effect). Extended practice referred to as overlearning leads to sustained maximal performance of primary task (bold line) and incremental rises in performance of secondary task (dotted line).

2.6.3. Effort and learning

The concept of effort is closely linked to that of automaticity. According to the Fitts' model of learning[42], described earlier (see section 2.2.4.) with practice participants appear to require less physical and mental effort ultimately leading to an expert state of effortless performance. Measures linked with effort namely subjective rating scales[81] and objective physiological markers such as heart rate, pupillary dilatation are used to detect intensity of effort invested. A decrease in effort expenditure on overlearning suggests the presence of continued learning. Surgical skills research has increasingly employed methods to measure amelioration of effort[82,83] along with traditional markers of performance (speed and accuracy).

2.7. Assessment of motor (technical) skills in surgery

Assessments of surgical skills require a sophisticated understanding of surgical research methodology. Towards defining and measuring surgical skills acquisition, Gallagher et al[74] created a framework with standardized definitions and criteria as follows:

1. ***Ability:*** the natural state or condition of being capable
2. ***Skill:*** a developed proficiency
3. ***Task:*** a piece of work to be done
4. ***Procedure:*** a series of steps taken to accomplish an end
5. ***Psychomotor ability*** is defined as one's natural performance with regards to operating in a three—dimensional (3D) space, yet receiving two-dimensional (2D) feedback from a screen e.g. laparoscopic surgery

2.7.1. Principles of surgical skills assessment

Surgical assessment must fulfil validity, reliability and feasibility criteria in order to be adopted[74,84,85]. Validity according to the American Physiological Association is defined as the extent to which a test measures what it was intended to measure[86]. Validity is a broad concept incorporating a number of features as follows:

1. **Face validity** refers whether the model of assessment resembles the task it is based upon. Essentially it a subjective measure of superficial appearance and is crucial for ensuring motivation of stakeholder.
2. **Content validity** refers to extent to which an assessment tool taps into the various aspects of the specific construct in question.
3. **Construct validity** is the degree to which the test measures what it intends to measure and not other variables. In surgical skills assessment this test is often designed to separate individuals based on their level of surgical expertise.
4. **Concurrent validity** measures the degree of correlation of one assessment with another, when both are designed to measure the same thing. In skills assessment it is used to compare the test in consideration against the gold standard.

5. **Discriminate validity** is defined as an evaluation that reflects the degree to which the scores generated by the assessment actually correlate with the factors with which they should.
6. **Predictive validity** measures how well the assessment can predict a relationship between the construct that is being measured and future behaviour.
7. **Reliability** is a measure of precision that describes the ability of an assessment to repeatedly produce consistent and reproducible results.
8. **Test-retest reliability** is a measure obtained by assessing a group of individuals twice at different times. Stability can be measured from the correlation coefficient obtained by correlating the results of the first test with that of the second.
9. **Parallel forms reliability** measures reliability of results obtained from different versions of an assessment tool where both versions interrogate the same skill.
10. **Inter-rater reliability** measures degree of agreement between the results obtained from two or more assessors, calculated by correlation of assessor's scores.
11. **Internal consistency** refers to the degree of correlation between different items on the same test.
12. **Feasibility** refers to the practicality of an assessment tool. Feasibility accounts for availability of resources such as cost and man power e.g. assessors.

2.7.2. Objective methods of assessing technical skills

Assessment of technical skill, usually takes place in the operating room by observation. Often these observations are not based on specific criteria. As a result it suffers from subjectivity and poor inter-rater reliability[84]. Patient outcome metrics such as morbidity and mortality are objective measures of technical skill assumed by its strong relationship[4], yet patient outcomes are also influenced by variance in patient characteristics and intra-operative factors. To overcome these limitations several methods have been developed to measure technical skill as follows:

a. Rating Scales

i. Objective Structured Assessment of Technical Skills (OSATS)

ii. Observational Clinical Human Reliability Assessment (OCHRA)

iii. Global Operative Assessment of Laparoscopic Skills (GOALS)

b. Scoring Systems

i. MISTELS

ii. FLS

c. Dexterity/ Motion tracking systems

i. ICSAD

ii. ADEPT

iii. Optical (Infra-red)

iv. Force measurement

v. Video based tracking

vi. Analysis using language models

d. Analysis of end product on bench model

e. Virtual Reality simulators

f. Eye-tracking

g. Neuroimaging

2.7.2.1 Rating scales

i. Objective Structured Assessment of Technical Skills (OSATS)

The assessment of technical performance by observation according to set criteria, standardizes evaluation and eliminates subjectivity by turning examiners into observers rather than interpreters of behaviour. The objective structured assessment of technical skills (OSATS) introduced by Martin et al[87] in 1997 espoused similar principles to the objective structured clinical examination (OSCE), a widely accepted method of assessment in health sciences. OSATS comprised of six to eight stations where trainees' performance was assessed on live animal or bench models, using task-specific checklists and global rating scales. Expert observation is in real-time and scoring is binary based on correct or incorrect performance e.g. OSATS checklist for small bowel anastomosis was based on 23 criteria. The global rating scale (GRS) involved grading seven components of an operative skill using a five point Likert scale. Grading as per the Likert scale was guided by description of the extreme and midpoints of the scale. Reznick et al[88] demonstrated reliability and construct validity of the OSATS tool. However, the feasibility of OSATS is limited by cost and resource inefficiency. Dearth of expert personnel and accredited centres to administer the examination, and the limited choice of bench top stations are issues impeding the acceptance of this tool. In addition OSATS cannot measure critical elements of skill such as the ability to dissect in the correct plane or optimal retraction[89].

ii. Observational Clinical Human Reliability Assessment (OCHRA)

Borrowing concepts from other high-risk fields that use observation of errors to measure performance, in 1998, Joice et al[90] employed Human Reliability Assessment techniques to measure performance during laparoscopic cholecystectomies. Errors from action or omission of appropriate action, which resulted in a negative consequence by increasing the duration of surgical procedure or required rectification, were classified into ten generic forms[91]. Error modes one to six (see Table 2.2) termed as 'procedural errors' rates the ability of the operator to perform the procedure in the correct sequence and error modes seven through ten termed as 'execution errors' score manipulative ability. Post-operative complications were also accounted to generate an error probability score. The OCHRA tool has been demonstrated to be valid in assessing complex laparoscopic colorectal surgery[92,93]. As the operative procedure is segmented to allow sub-task analysis, the tool has been particularly useful in identifying critical areas where errors most commonly occur[91-93], which serves as useful information to target training. However, it is time exhaustive requiring an expert trained assessor to rate videos of lengthy laparoscopic procedures.

Type	Description
1	Step is not done
2	Step is partially completed
3	Step is repeated
4	Second step is done in addition
5	Second step is done instead of first step
6	Step is done out of sequence
7	Step is done with excessive speed, force, distance, time, rotation, depth
8	Step is done with insufficient speed, force, distance, time, rotation, depth
9	Step is done in wrong orientation, direction, point in space
10	Step is done on/ with the wrong object

Table 2.2 describes classification of errors from 1 to 10 based on the OCHRA tool[91]. Errors 1 to 6 are procedural errors and errors 7 to 10 are execution errors.

iii. Global Operative Assessment of Laparoscopic Skills (GOALS)

In 2005, Vassiliou et al[94], developed Global Operative Assessment of Laparoscopic Skills (GOALS) which is the laparoscopic equivalent for OSATS to address the growing need to assess laparoscopic skills. It is composed of five items (see Table 2.3); performance on each item is rated using a five point Likert scale akin to GRS, where '1' represented the lowest score and '5' the highest. Set descriptors of performance helped anchor extreme and midpoints of the scale. Four out of five items addressed domains concerning laparoscopic surgery namely: a) depth perception; b) bimanual dexterity; c) movement efficiency and d) tissue handling. The fifth and final component assessed autonomy of the operator by measuring degree of dependence on expert instruction. Akin to OSATS tool, task specific checklists were created and scored in binary fashion. Since then a number of studies have demonstrated construct validity of GOALS in assessment of laparoscopic procedures such as laparoscopic cholecystectomy[95], appendicectomy[96], oesophageal fundoplication[97] and hernia repair[98] to name but a few. However, criticisms of the GOALS tool are the lack of expert personnel to rate the laparoscopic videos and it is time consuming, for example a rater requires approximately 30 to 40 minutes to assess a video recording of laparoscopic bypass procedure lasting two hours long[99].

Item	Global Rating Scale 1	2	3	4	5
Depth Perception	Constantly overshoots target, wide movements and slow to correct		Some overshooting/ missing of target, quick to correct		Accurately targeted instrumentation in the correct plane
Bimanual dexterity	Ignores usage of non-dominant hand, poor bimanual co-ordination		Uses both hands, but sub-optimal bimanual co-ordination		Bimanually well co-ordinated movements
Movement efficiency	Inefficient, tentative movements; focus constantly changes or persists without progress		Slow, but planned movements		Confident, efficient, safe. Efficient focus on task
Tissue handling	Coarse movements, injures tissue & adjacent structures, poor grasper control		Handles tissues reasonably well, minor trauma to adjacent tissues		Handles tissues well, appropriate traction & negligible injury to adjacent structures
Autonomy	Unable to complete entire task, even with guidance		Completes task safely with moderate guidance		Completes task independently without prompting

Table 2.3. Items of global rating scales to form the Global Operative Assessment of Laparoscopic Surgery (GOALS) tool are tabulated.

2.7.2.2. Scoring Systems

The need for less laborious, inexpensive, accessible and objective methods to measure laparoscopic performance and simultaneously promote training methods outside the operating room led to the development of the programs such as the McGill Inanimate System for Training and Evaluation of Laparoscopic Skill (MISTELS) and the Fundamentals of Laparoscopic Surgery (FLS). Both assessment tools measure speed and accuracy as surrogate markers of learning or expertise.

i. McGill Inanimate System for Training and Evaluation of Laparoscopic Skill

Fried and colleagues[100] from Montreal, Canada; a well established group in the field of laparoscopic skills education, developed the MISTELS program in order to make laparoscopic skills assessment inexpensive and more accessible. It uses physical simulation in the form of a box-trainer to mimic the set-up of laparoscopic surgery. From the initial seven, five generic laparoscopic tasks were retained and the remaining two were discarded, deemed by inferences that they were not only expensive but did not have additional value. Performance on the five tasks described below was derived by subtracting time taken by the individual on the task and penalties scored on set criteria from a pre-set value[100]. Numerous studies have demonstrated construct validity, inter-rater and retest reliability[100-103].

a) **Pegboard transfer (Task 1):** The operator was required to transfer six pegs, one at a time by grasping it with one hand, transferring to another hand and placing it on a second peg board. Once all pegs were transferred the procedure was performed in reverse sequence. The *cut off* time for this task is 300 seconds and penalties are scored on the basis of pegs dropped.
b) **Pattern cutting (Task 2):** In this task, the operator is required to cut along the boundaries of a circle measuring 5 cms in diameter imprinted on a suspended piece of gauze measuring 10×10 cm^2. The *cut-off* time for this task is 300 seconds and penalties are calculated by deviation of the cut from the drawn circle.
c) **Placement of a ligation loop (Task 3):** Operators must place a pre-tied slipknot on to a foam tubular appendage at a specified site. This task mimics endo-looping tubular structures e.g. appendix. The *cut-off* time for this task is 180 seconds and penalties are calculated by quality of the knot and deviation of the placed knot from the designated site in millimetres.
d) **Intracorporeal knot (Task 4):** This task warrants placing a suture through two pre-marked points on either end of a longitudinal slit created in a Penrose drain. The *cut-off* time for this task is 600 seconds and penalties are calculated by deviation in millimetres of the suture from specified target, gap in millimetres between opposed edges and quality of knot. Amongst the five tasks intra-corporeal knot tying is considered challenging and the author has used this task for the experiments later described.

e) **Extracorporeal knot (Task 5):** This task is similar to the previous task (task 4) in many ways except that the knot is tied via an extracorporeal technique using a knot pusher. The *cut off* time is 420 seconds and penalties are calculated akin to task 4.

ii. Fundamentals of Laparoscopic Surgery (FLS)

In 1997 the Society of American Gastrointestinal Endoscopic Surgery (SAGES) formed a committee responsible for developing a curriculum to train and test fundamental laparoscopic skills[104]. The technical skills component of the program is derived from the earlier described MISTELS and the technical tasks (see Figure 2.5) are adopted from MISTELS curriculum.

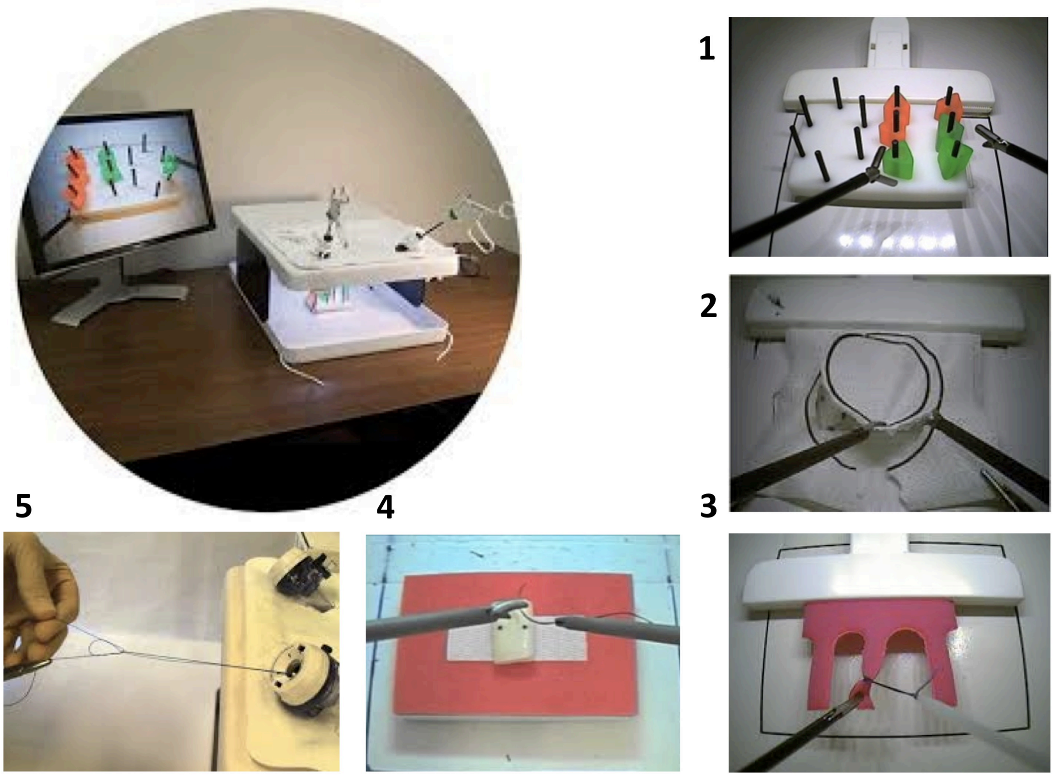

Figure 2.5. Equipment and set-up for tasks (1-5) employed by the Fundamentals of Surgery curriculum are depicted (adapted from[26]). The skills illustrated in clock-wise fashion are peg transfer (task 1), pattern cutting (task 2), loop ligation (task 3), intracorporeal knot tying (task 4) and extracorporeal knot tying (task 5).

By comparing novice and expert performance, cut-off (passing scores) were determined to determine better separation between competent and non-competent performers[100]. In addition FLS technical performance correlated favourably with subjective assessment of performance on GOALS[105] and end of term resident assessments[106]. Endorsed by SAGES and the American College of Surgeons (ACS), FLS has received widespread adoption in surgical residency programs in the United States. Since 2008, passing the FLS skills assessment has become a mandatory component for certification with the American Board of Surgery[107] and over ten thousand surgical residents have been certified in approved centres[26]. In the proceeding experimental chapters, intracorporeal knot tying (Task 4) of FLS and MISTELS curricula was chosen to test expertise and acquisition of skill.

2.7.2.3. Dexterity/ Motion tracking systems

Tracking devices are essentially composed of: i) an object to be tracked (hand or instrument tip), ii) a device to track (hardware) and lastly iii) a system to process measured positional information (hardware and software). Metrics derived from motion analysis has enabled expertise skills classification and are defined in the below table[108,109] (see Table 2.4).

i. Electromagnetic tracking systems

The Imperial College Surgical Assessment Device (ICSAD) device uses commercially available electromagnetic technology (Isotrack II, Polhemus Inc, Vermont, USA) to assess technical performance by tracking an electromagnetic sensor (1 cm^2 in area) attached to the dorsum of the operator's hands and an electromagnetic emitter within the surgical field. It provides basic motion metrics such as time take, number of movements, speed and distance travelled. This device has been used successfully in expertise classification for both open and laparoscopic skills[110-112] and to compare skills acquisition between standard laparoscopy (multi-port) and single incision laparoscopic surgery (SILS)[113]. The advantages of ICSAD are data extraction at a reasonable rate (20 Hz), portability and the lack of line-sight issues associated with optical tracking. However, the sensors are wired, magnetic objects and distance affect signal quality. Beyond the scope of ex-vivo laparoscopic surgery this technology is not feasible as the sensors when worn on the hand or placed on the instruments are obtrusive.

Later developments have led to use of smaller, wireless, inconspicuous electromagnetic sensors (AURORA, Northern Digital Inc, Waterloo, Canada). A study on eighteen participants (nine novices and nine experienced surgeons) demonstrated construct validity on the basis of time, path length and speed measured from the aforementioned electromagnetic sensors attached to the tip of the laparoscopic instruments[114].

Metric (Units)	**Definition**
Time (s)	Total time taken to perform the task
Approaching time (s)	Time taken to reach the target
Search time (%)	Percentage of time spent in the search zone
Idle time (%)	Percentage of time where no movement occurs
Number of movements (n)	Number of movements executed for task completion
Path length (mms)	Length of the curve during performance derived from positional data
Distance efficiency/ movement economy	Measure of economy, estimated by relationship between actual path length and ideal path
Angular area	Area calculated from the distances between the farthest left to right and forward to backward positions of the instrument while performing a task
Volume	Distances calculated between the farthest left to right, forward to backward and in to out positions of the instrument while performing a task
Economy of area	Relationship between the maximum area occupied by the instrument and the total path length
Economy of volume	Relationship between the maximum volume occupied by the instrument and the total path length
Depth perception (mms)	Total distance travelled by the instrument along its axis
Accuracy	Correctness of placing the instrument tip in 3D
Transit profile	Transit trajectory projected onto 2D plane
Deviation (mms)	Deviations on the horizontal or vertical plane from the ideal course of action
Response orientation (radians)	Amount of rotation about the axis of the instrument
Speed (mm/s	Rate of change of the instruments' position
Speed profile	Shape of the speed curve
Acceleration (mm/s^2)	Rate of change of the instruments' speed
Motion smoothness (mm/s^3)	Analysed from the third time-derivative of position, which represents a change in acceleration

Table 2.4. Definitions of metrics assessed on motion tracking devices employed in objective technical skills assessment in surgery (adapted from Sanchez-Margallo et al[108]).

ii. Advanced Dundee Endoscopic Psychomotor Tester [ADEPT]

Cuschieri and colleagues[115] from Dundee, Scotland developed the ADEPT system to test innate manual dexterity and hand-eye co-ordination. It tracks movement of two instruments by a set of potentiometers mounted in a dual-gimbal mechanism. The ADEPT capable of computing execution time, instrument error, path length and task completion demonstrated construct validity[116] and reliability for bimanual laparoscopic tasks[117].

iii. Optical (Infra-red) Motion Tracking

Commercially available optical tracking systems (Polaris, Northern Digital Inc, Waterloo, Canada) are used to assess motion. They employ infrared (IR), and the trackers can be either active when the markers are powered or passive when reflective markers illuminated by projected IR. Despite the advantages of high resolution and wireless technology (passive trackers), the device is relatively expensive and suffers from line of sight issues. As a result more than one limb cannot be simultaneously tracked due to overlapping signals and data can be lost when the signal is blocked by another object. Therefore, unsurprisingly it has been used in collaboration with electromagnetic trackers[118]. The Hiroshima University Endoscopic Surgical Assessment Device (HUESAD)[119] employs optical scale sensors and microencoders to track movement of laparoscopic instrument tips. It was designed to assess the smoothness in the use of laparoscopic instruments. The sensors measure time taken, one distance parameter and two angle parameters, however there has been limited uptake of this technology[120].

iv. Force Measurement Systems

Distortion or complete loss of tactile feedback is a consequence of adopting laparoscopic and robotic surgery techniques respectively. As a result it raises the critical issue of the potency to traumatizing tissue by poor handling of tissue. Measurement of force has demonstrated its potential value in discriminating expertise of skill. In an experiment by Cundy et al[121], sensors placed on a base plate within a Paediatric version of FLS during performance of a simple (peg transfer) and an advanced laparoscopic task (suturing) measured force magnitude, mean and maximal force. Analysis of force generated, better correlated with expertise than standard scoring systems (MISTELS) that account for time and accuracy. When force parameters are used with data from motion and time by Horeman et al[122], it demonstrated an accuracy of up to 100% in identification of laparoscopic skill level.

v. Video Based Tracking

Laparoscopic videos can be used as a source for tool tracking. Some systems rely on artificial markers and others involves introduction of physical markers into the operative field, which raises the issue of biocompatibility for in-vivo scenarios. Artificial markers such as colours on the instrument tips can be tracked using computer vision techniques[108] namely colour segmentation, stereoscopic techniques, or a Continuously Adaptive Mean Shift algorithm[123]. Video-based markers suffer from line of sight problems and low accuracy, yet they have been successfully trialled in *in-vivo* animal laparoscopic procedures[123].

vi. Motion analysis using language models

Limitations of dexterity analysis using above described methods are as follows: 1) they rely on prior expertise classification thus introducing subjective bias, 2) do not account for variation in technique for the same skill and 3) are not capable of automated task segmentation and analysis. Language models can be built by task decomposition of surgical procedures based a hierarchical model (see figure 2.6) and then used to develop a deeper understanding of motion variance between surgeons[124]. Hidden-Markov-Models (HMM) have been used to learn laparoscopic surgical manoeuvres from motion metrics of subjects with varied skill abilities and then predict the level of skill eliminating subjective bias[125] up to an accuracy of 92%[126]. HMMs can also be used for automated segmentation and recognition of tasks with classification accuracy of up to 97%[127]. Linear Discriminate Analysis (LDA) has also been employed successfully, such as in the study by Lin et al[128] who used a LDA to perform automated classification according to skill level and type of task from motions captured during robotic surgery. Alternatively, other language models such as Support Vector Machine (SVM) have been used along with HMM as a hybrid to improve classification of robotic surgery tasks by analysing generated force/torque[129].

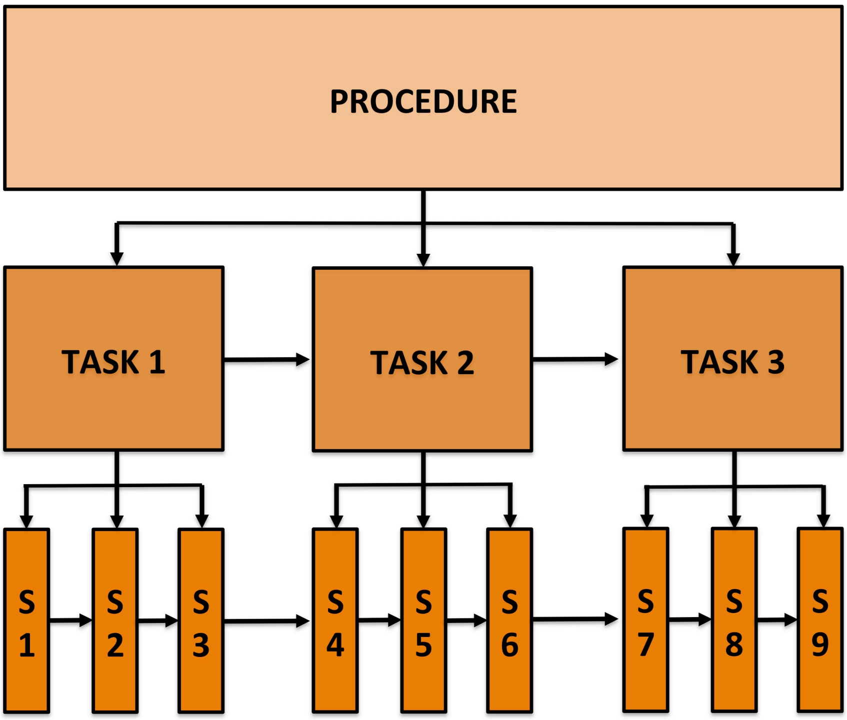

Figure 2.6. Hierarchical decomposition of a surgical task (adapted from Reiley et al[124]). A surgical procedure is composed of tasks and each task is composed of surgical gestures known as 'surgemes' (S).

2.7.2.4. Analysis of the final product

The outcome of a completed surgical task on bench models can be analysed qualitatively[130] or quantitatively. In a study by Datta et al[131], surgical tasks that require anastomosis were assessed for leakage and luminal area created. End product measures correlated with manual dexterity derived from motion metrics (ICSAD). Similarly, laparoscopic knots have been assessed using a tensiometer[132]. In summary, the advantages of combining end product with subjective rating systems such as OSATS are to improve objectivity and accuracy of skills classification[84].

2.7.2.5 Virtual Reality Systems

Virtual Reality (VR) defined by McCoy et al[133] is a collection of technologies that allow efficient interaction with 3D computerised databases in real time by using their natural senses and skills. The advantages of VR systems are that it obviates the need for bench-models, physical simulators and in-vivo assessments. Since the introduction of the MIST-VR (Mentice AB, Göteborg, Sweden), virtual reality systems have become increasingly more sophisticated allowing progression from training and testing of simple generic surgical tasks to validated advanced surgical skills/ procedures[108] (see Table 2.5).

Platform	VR system	Training/Evaluation
Laparoscopic	MIST-VR (Mentice AB, Gothenburg, Sweden)	Simple tasks
	SINERGIA (SINERGIA Consortium, Spain)	Simple tasks
	SIMENDO (DeltaTech, Delft, Netherlands)	Simple and advanced tasks
	LapSim (Surgical Science, Gothenburg, Sweden)	Simple and advanced tasks/ procedures (cholecystectomy, appendicectomy and bariatrics)
	SEP Simulator (SimSurgery AS, Oslo, Norway)	Simple and advanced tasks/ procedures (SILS, cholecystectomy)
	LapVR (Immersion Medical, San Jose, CA, US)	Simple and advanced tasks/ procedures (cholecystectomy and appendicectomy)
	LapMentor (Simbionix Ltd, Beit Golan, Israel)	Simple and advanced tasks/ procedures (cholecystectomy, appendicectomy, incisional hernia repair, bariatrics and colorectal)
	ProMIS (CAE Healthcare, Quebec, Canada)	Simple and advanced augmented reality tasks/ procedures (cholecystectomy, appendicectomy and fundoplication)
Robotic	dVSS (Intuitive Surgical Inc, Sunnyvale, CA, US)	Simple and advanced tasks
	dV-Trainer (Mimic Inc, Seattle, WA, US)	Simple and advanced tasks
	RoSS (Simulated Surgical Systems LLC, San Jose, CA, US	Simple and advanced tasks

Table 2.5. Types and capabilities of laparoscopic and VR simulator systems.

One of the criticisms of earlier generation VR systems was the lack of tactile feedback and fidelity, thus allowing practice of simple generic laparoscopic tasks. However, with technological advancements, new generation VR systems are capable of providing haptic feedback and the improved fidelity enables testing of advances tasks and whole surgical procedures. In addition, VR systems are capable of deducing motion metrics such as path length, economy of movement, and errors. The ProMIS augmented simulator (CAE Healthcare, Quebec, Canada) differs from other VR systems as it combines real time performance with virtual settings by augmented reality. The advantages of assessments by VR are that results are automatically generated providing real-time feedback. The disadvantages are that VR systems still lack the fidelity when replicating operative conditions and the start-up expenses.

2.7.2.6. Eye-tracking

Despite the many advances in measuring surgical expertise using motion metrics, expertise classification is still problematic and remains elusive[124]. Issues raised were the lack of objectivity, most assessments occur in simulated/VR environments and require expert assessors to spend long periods reviewing performance. In addition characteristics of experts such as focused attention during deliberate practice[30,134,135] remain unappreciated.

Eye tracking technologies (see figure 2.7) are capable of measuring eye activity via non-invasive techniques[136]. Eye or gaze behaviour was used more than fifty years ago to study aircraft pilots' interaction with instrumentation in a cockpit during landing[137]. Current eye-tracking technologies are either remote or head mounted, in the latter it is incorporated to wearable glasses. Essentially its components are a light source and camera; the light source illuminates the eye with infrared light and the camera records corneal reflection of infrared light. The technology tracks pupillary position, maps focus of attention and video records the subject's field of gaze[138]. Obtained eye metrics from each eye are as follows:

1. **Blink rate**: measures if the eye is open or closed.
2. **Fixation rate**: measures if point of gaze is moving or fixed.
3. **Vergence**: determined by the horizontal distance between the two eyes at each observation.
4. **Index of cognitive activity (ICA)**: measure of pupil diametric variability. More recently this measure has been linked with effortful cognitive processing when the subject is fatigued or performing a difficult task[139].

Employing these standard eye metrics, gaze behaviour has demonstrated to be different between experts and novices during performance of laparoscopic VR tasks[140-142]. Experts gaze fixated to a greater degree on the target than the instruments, whilst novices fixated almost equally on both target and instruments[140,142]., and / demonstrate tool and target saccadic cross referencing. In addition, experts displayed a more efficient gaze strategy by lesser eye movements between tool and target[142]. Additionally, eye-tracking is a useful tool in detecting non-technical skills such as situational awareness exhibited by a study where expert surgeons whilst performing a simulated surgical procedure were able to detect abnormal vital signs of the simulated patient unlike novices[141].

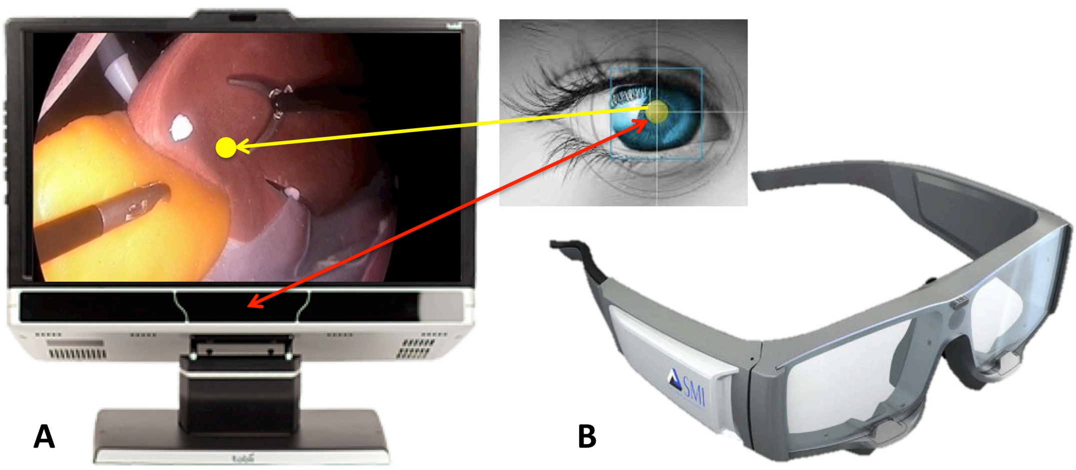

Figure 2.7. Remote eye tracker composed of near infrared micro-projectors and optical sensors. (A) During performance of a laparoscopic surgical task, the intra-operative view is projected on a two dimensional screen. Operators' visual gaze at the point of interest (yellow arrow) is tracked by illuminating infrared light from micro-projectors, which is reflected by cornea of the operator's eye and then registered in real time by image sensors (red arrow). (B) Wearable eye tracker with integrated micro-projectors and sensors.

Data obtained from eye metrics is complex, machine learning techniques have been used to optimise classification in terms of skill or experience[139]. Using LDA and non-linear discriminate analysis, Richstone et al[143], were able to identify experts from non-experts during performance of simulated and live surgical tasks to an accuracy of 92% and 81% respectively. Laparoscopic task segmentation is also feasible using a Parallel Layer Perceptor model to an accuracy of 66% when used alone and can be improved to 75% when combined with video image analysis[144].

2.8 Conclusion

MIS has revolutionized surgical practice, yet its continuous evolution requires surgeons to acquire even more complex technical skills at a time when faced with decreased training opportunities. Motor skills for MIS are uniquely challenging to acquire due to ergonomic challenges posed by instrumentation and the need for sensorimotor adaptation. Identifying the importance to address surgical training, multiple stakeholders have identified the need to improve MIS skills acquisition, resulting in an explosion of research into surgical skills assessment and effects of training.

In this chapter a variety of validated assessment tools employed in assessment of MIS and open surgery are described. Although they are valuable in differentiating between experts versus novices and identifying improvements in performance, they are unable to distinguish between proficient performers and experts. In addition few longitudinal studies track MIS skills acquisition over multiple periods with intensive distributed training regimes. Some tools are hampered additionally by the need of expert assessors to be trained in the use of the assessment tool and dedicate significant amounts of their time. Although they are objective by measuring speed and accuracy of hand/instrument movements, they fail to measure intensity of effort a key feature of motor skill. More importantly these markers measure performance that cannot be assumed to reflect learning, as learning is indirectly observable.

Motor learning occurs internally within the central nervous system by a set of neural processes. Improved understanding in described models of motor learning and identifying the effect of on and off-task factors would be of tremendous value to improving current methods of assessing skill level and enhancing skills acquisition. Multiple psychological motor learning models have described progression from novice to expert. Alternate methods of motor learning assessment have borrowed concepts from these models using intensity of effort or reliance on attention resources as surrogate markers of automaticity, which is associated with expert-status. However effort is often measured subjectively and reliance on attentional resources is tested outwardly by the ability to multi-task.

Functional neuroimaging is capable of providing objective neurophysiological information with regards to correlates for motor learning, attention and investment of effort for motor skills specific to MIS. In the next chapter, specific functions and pathways of neuroanatomical regions in relation to motor learning are firstly discussed. Secondly, properties of neuroimaging techniques are described and justification for selection of the appropriate neuroimaging modality is provided.

Chapter 3

Cortical Correlates of Motor Control & Functional Neuroimaging

3.1. Introduction

The previous chapter established the importance of understanding motor skills learning from a behavioural perspective and the way in which motor learning theory may be applied towards enhancing technical performance in surgery. This chapter shifts the emphasis from a description of motor skills learning theory and focuses on the changes in neurocognition that underpin learning. Motor skills learning described earlier as a difficult to capture internal process, is identifiable by neuroimaging techniques. The human motor system responsible for voluntary movement is composed of several cortical and sub-cortical regions linked to phases of motor learning. Their individual anatomy and specialised function along with their participation in complex pathways connecting cortical and sub-cortical regions to effector muscle groups are initially described in this chapter. Secondly, a variety of functional neuroimaging techniques, their individual strengths and weaknesses towards interrogating specialised functions of brain regions in isolation or integration are also discussed. Thirdly, I focus on describing functional Near Infrared Spectroscopy (fNIRS), providing detailed justification for choosing this neuroimaging modality to interrogate motor skills learning in surgery. Lastly, analytical methods employed to investigate functional brain behaviour in isolated or integrated models are described.

3.2. Neuroanatomy & Neurophysiology of Motor Control

The brain is mainly divided into forebrain, midbrain and hindbrain (see Figure 3.1). The forebrain is mainly divided into the telencephalon and diencephalon. The telencephalon contains the cerebral cortex, basal ganglia, hippocampus and corpus callosum. The diencephalon consists of the thalamus and hypothalamus. The hindbrain located below the midbrain is comprised of cerebellum, pons and medulla. Hindbrain structures and the spinal cord collectively form the brainstem.

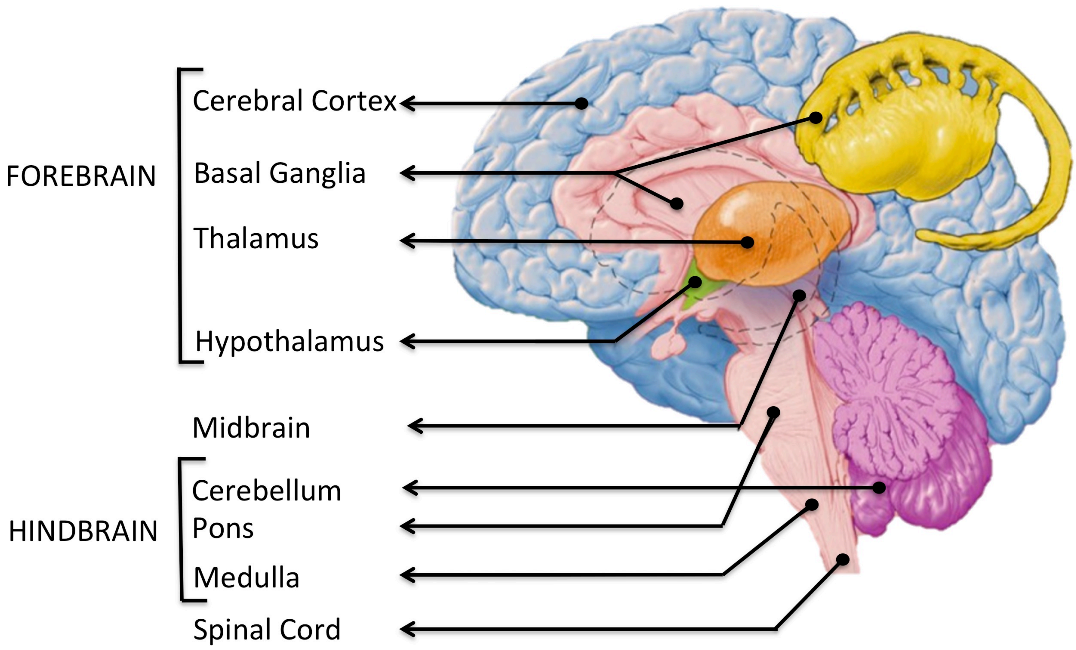

Figure 3.1. Divisions of the brain are diagrammatically represented.

3.2.1. Cerebral Cortex

Several cortical and subcortical regions are distinctly implicated in motor planning, execution and control. At the apex in the cerebral cortex are the primary motor cortex (M1) and secondary motor regions. The M1 located in the frontal lobe (see Figure 3.2) lies anterior to the central sulcus and posterior to the secondary motor areas of Brodmann's Area (BA) 6[145]. Located in BA6 are the Supplementary Motor Area (SMA) medially and the Premotor Cortex (PMC) laterally. Posterior to the central sulcus are the somatosensory regions of the parietal lobe. These regions receive sensory input, which includes limb proprioception (limb position). Dense connections between motor and sensory cortical regions, as a consequence forms a unified functional entity, referred to as the sensorimotor cortex.

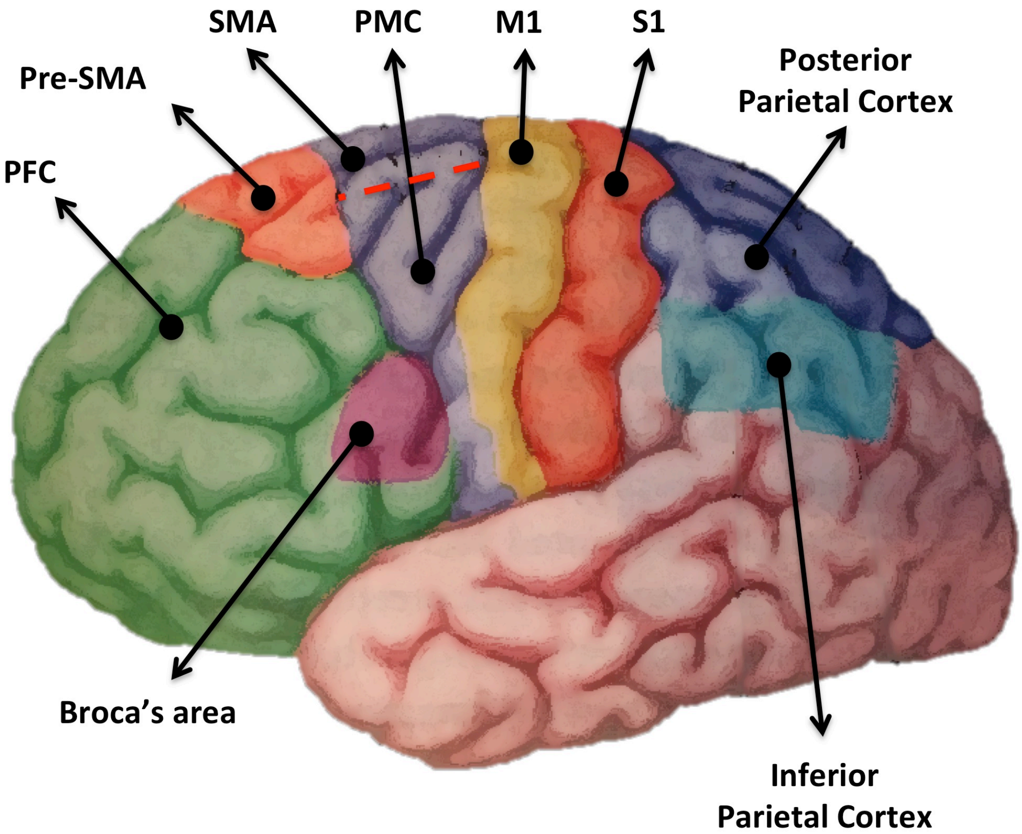

Figure 3.2. Motor regions of the cerebral cortex involved in motor control are diagrammatically represented. Within the frontal lobe are the Prefrontal cortex (PFC), Pre-SMA (Pre-Supplementary Motor Area), SMA, (Supplementary Motor Area), PMC (Premotor Cortex) and M1 (Primary Motor Cortex).

Somatotropic representations of the human body first described by Penfield and colleagues[146] as 'homunculus' are found in motor and sensory regions of the cerebral cortex. These cortical regions have a topographic representation to body surface for both motor and sensory processes. The term 'homunculus' in Latin means 'little man' was borne out of its diagrammatic representation. The motor homunculus located in the primary cortex represents motor function of the opposite side and follows an organized but inverse pattern. For example, the foot is represented at the top of the cerebral hemisphere and tongue movements at the lateral aspect of the cerebral hemisphere (see Figure 3.3). In addition effector representation does not correspond to actual size but is dictated by the level of innervation required for manipulation. For example, muscles of the hand and especially of the thumb have a much larger representation than lower limb muscles (see Figure 3.3).

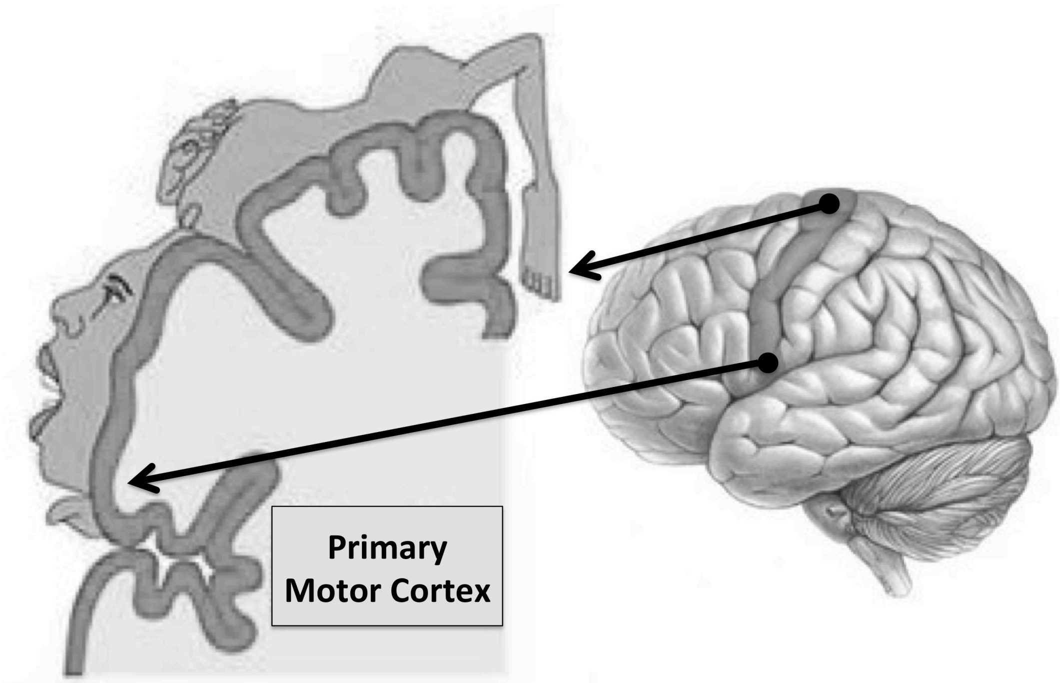

Figure 3.3. Diagrammatic representation of the motor homunculus (adapted from Penfield et al[146]) located in the primary motor cortex (left).

The Prefrontal cortex (PFC) lies anterior to BA6 and can be divided into three smaller areas namely dorsolateral prefrontal cortex (DLPFC), orbitofrontal cortex (OFC), and the anterior cingulate cortex (ACC). Together they play a role in planning, execution and attention for complex motor tasks.

3.2.2. Descending motor pathways

The descending motor pathways are mainly divided into pyramidal and extrapyramidal tract as follows:

i) Pyramidal tract
   a) Corticospinal
   b) Corticobulabar

ii) Extrapyramidal tract
   a) Rubrospinal
   b) Tectospinal
   c) Vestibulospinal
   d) Reticulospinal

3.2.2.1. Pyramidal tracts

Corticospinal tract: The motor regions of the cerebral cortex along with sub-cortical structures namely the basal ganglia, thalamus, cerebellum and brain stem regulate activity of spinal neurons both directly and indirectly. Direct connections are formed by the 'corticospinal' tract also known as the pyramidal tract, responsible for voluntary skilled movement of limbs. As the name suggests, it is a descending pathway arising from M1 and terminates synaptically with either alpha motor neurons or spinal interneurons (see Figure 3.4). The majority of axonal fibres within the corticospinal tract decussate and crossover to the contralateral side at the level of the medullary pyramids to form the lateral corticospinal tract. The remaining fibres amount to approximately 10-20%, and continue to descend on the ipsilateral side to form the ventral/ anterior corticospinal tract.

Corticobulbar tract: The corticobulbar tract is the other pyramidal tract that supplies movement for the head and neck.

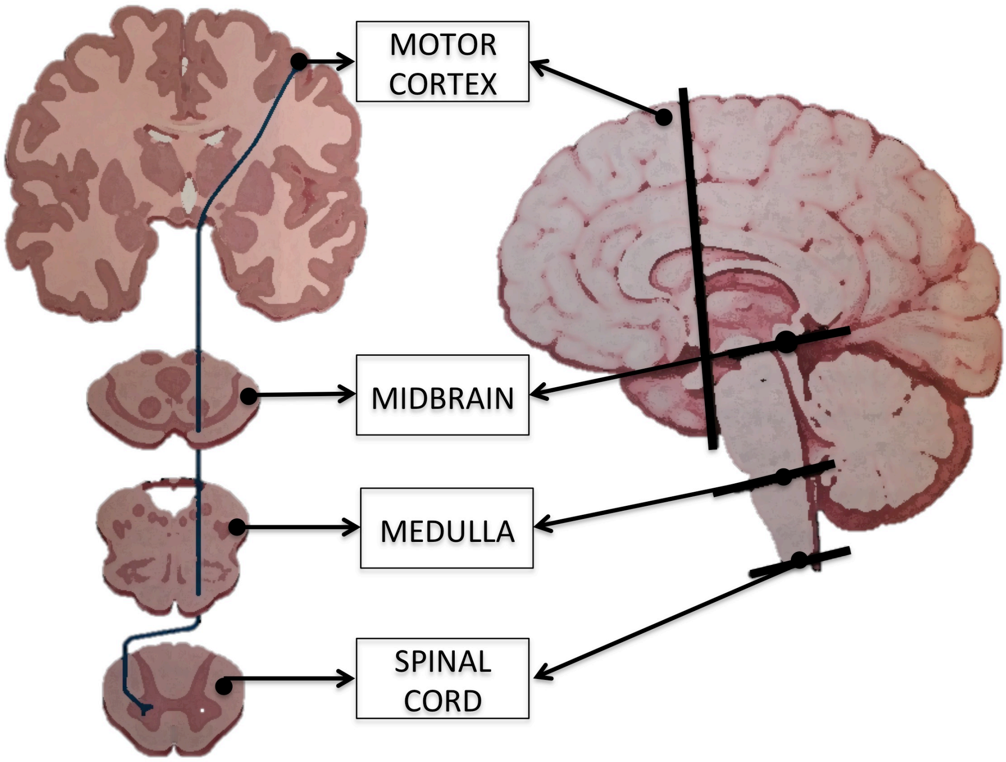

Figure 3.4. Corticospinal tract descends from motor cortex to spinal cord at each level (axially), which corresponds to labelled neural systems on the right.

3.2.2.2 Extrapyramidal tracts

The remaining tracts are collectively referred to as 'extrapyramidal' as they do not traverse through the medullary pyramids. They are the rubrospinal, tectospinal, vestibulospinal and reticulospinal tracts.

1. **Rubrospinal:** Begins from the red nucleus in the midbrain, crosses to the contralateral side near its origin before reaching the medulla. The red nucleus besides receiving input from the same cerebral cortical regions as the corticospinal tract and also from the cerebellum. They too are responsible for movement of distal limb musculature, yet their role maybe smaller due to the presence of fewer axonal fibres than the corticospinal tract.

2. **Tectospinal:** The tectospinal tract arises from the superior colliculus also located in the midbrain and on descent not all fibres decussate. Their role has been implicated in mediating visual reflexes.

3. **Vestibulospinal:** Composed of two pathways, termed 'medial' and 'lateral'. It originates from the vestibular nuclei located in the midbrain, also receives input from the cerebellum and remains ipsilateral during descent. It plays a role in maintaining postural balance as it innervates via lower motor neurones the 'anti-gravity muscles' composed of upper limb flexors and lower limbs extensors.

4. **Reticulospinal:** The reticulospinal tract is also composed of two pathways medial and lateral and remains ipsilateral on descent. The medial reticulospinal tract originates from the pons increases muscle tone and facilitates voluntary movement. The lateral reticulospinal tract arising from the medulla and plays the opposite role to the medial tract by reducing muscle tone.

3.2.3. Basal Ganglia

Basal ganglia as the name suggests is comprised of a distributed set of neural structures buried deep within the cerebrum. The basal ganglia plays a crucial role in motor loops by receiving information from several cortical regions, processing it and relaying it to the motor cortex via the thalamus. It is responsible for initiating and regulating motor commands. In the forebrain it is mainly composed of the caudate nucleus, putamen, nucleus accumbens and globus pallidus. Collectively these structures are referred to as corpus striatum because of the striated appearance of their cell bridges. The caudate and putamen are fused anteriorly, together referred to as the 'neostriatum' and functionally similar. The corpus striatum is the main recipient of signals from various regions of the cerebral cortex especially the frontal lobe and the thalamus. Notably the caudate head receives afferents from the PMC and SMA, whilst the putamen receives afferents from the M1 and S1 regions of the cortex. Akin to the motor homunculus of the M1 lies a corresponding homunculus in the basal ganglia. The remaining structures, which lie below the thalamus are the subthalamic nuclei and the substantia nigra, the latter is sub-divided into the pars compacta and pars reticulata. Along with the globus pallidus interna they are responsible for major outputs via the ventrolateral, ventroanterior and mediodorsal nuclei of the thalamus into the PMC, M1 and PFC.

The basal ganglia, plays a role in balancing motor function via five cortico-striatal loops which can be divided into 'direct' and 'indirect' pathways on the basis of their opposing effects on thalamic nuclei. In sequential order, the direct pathway passes from the cortex through to the striatum, globus pallidus interna, thalamus and end at M1[147,148]. It has an excitatory effect facilitating motor programs in the cortex. The indirect pathway similarly begins at the cortex and ends at M1, albeit passing in order through the globus pallidus extrena, subthalamic nuclei, globus pallidus interna and thalamus. It simultaneously inhibits competing unwanted motor programs, thereby ultimately facilitating desirable movement.

3.2.4. Thalamus

The thalamus located in the diencephalon is comprised of four parts namely hypothalamus, epythalamus, ventral and dorsal thalamus. It forms the central core of the brain surrounded by cerebral hemispheres except at its base (see figure 3.5.). Multinuclear in structure, the thalamus' role in motor control is poorly understood, mainly implicated as a relay for motor and sensory information[149]. Two sets of ventral nuclei and one medial nucleus are associated with motor control. One set comprised of the ventrolateral and ventroposterolateral receive signals from the cerebellum. The other set includes a different ventrolateral and a ventroanterior nuclei receive signals from the basal ganglia. Lastly, the mediodorsal nuclei receives signals from the superior colliculus.

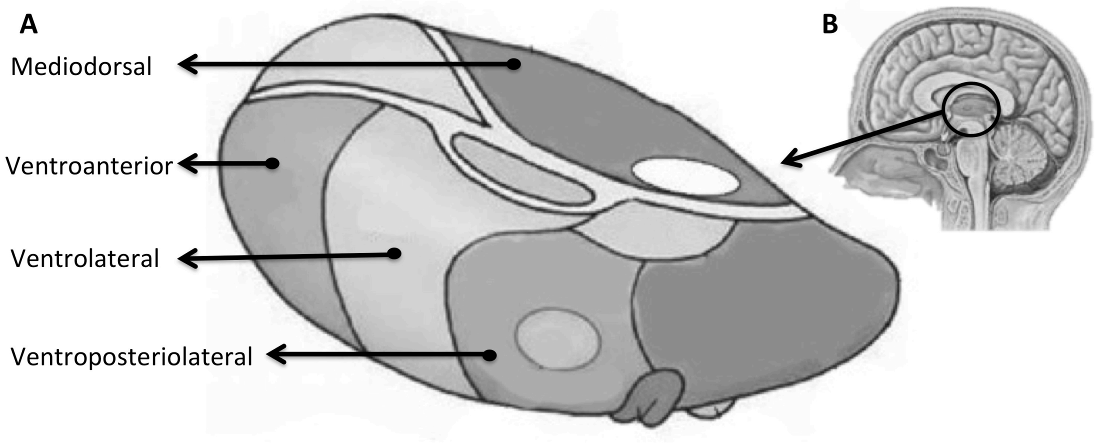

Figure 3.5. Schematic illustration of thalamic nuclei associated with motor control (A) and the location of the thalamus in the brain (B).

All ascending pathways connect sub-cortical structures via the thalamus to the frontal lobe of the cerebral cortex. The function of pathway emanating from the cerebellum is making movements in response to sensory stimulation. Similarly, the ascending pathway from the basal ganglia plays a role in selecting and generating wilful movements[150] in response to external sensory stimuli when alternatives are available[151]. The third pathway arising from the superior colliculus which involves eye movements is implicated in monitoring one's own actions[149].

3.2.5. Cerebellum

The cerebellum also known as the little brain is located in the hindbrain below the occipital and temporal lobes of the cerebral cortex. It forms only 10% of overall brain volume but accounts for approximately half of all the neurons. The cerebellum is divided anatomically into three lobes by two major fissures. The primary fissure separates the anterior lobe from posterior lobe whilst the posteriolateral fissure separates the flocculonodular lobe from the rest (see figure 3.6).

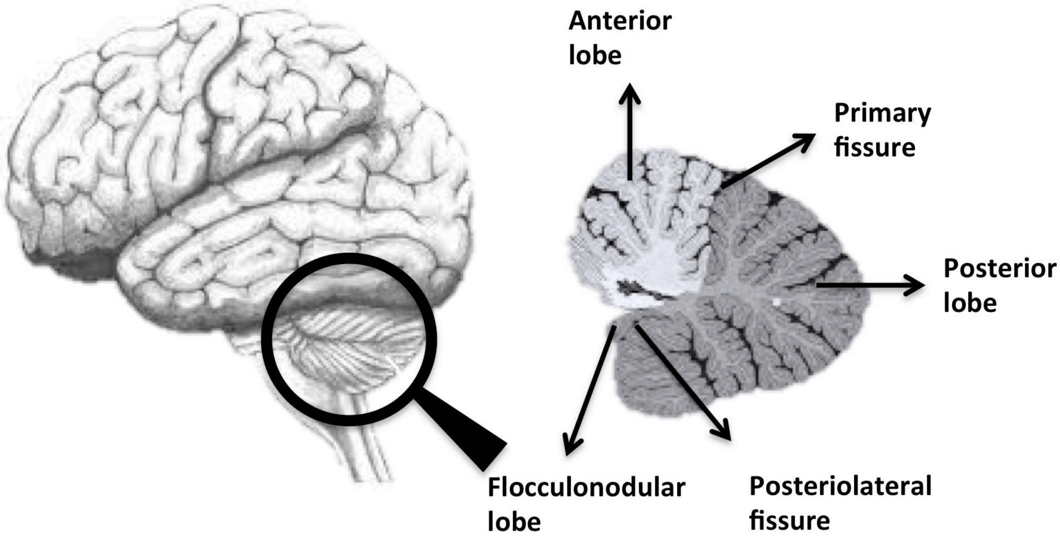

Figure 3.6. Cerebellum located within the magnifying lens is anatomical divided by two fissures (primary and posteriolateral) into three lobes namely the anterior, posterior and flocculonodular lobe.

The cerebellum is further sub-divided into three zones along sagittal planes. These zones from medial to lateral are the *'vermis'* followed by the *'intermediate zone'* and *'lateral hemisphere'*. Three fibre bundles namely the superior, middle and inferior cerebellar peduncles carry input and output signals to and from the cerebellum. The middle and inferior cerebellum mainly carries afferent signals from the pons and medulla respectively, whilst the superior cerebellar peduncle conveys efferent signals.

Anatomical divisions correspond to functional divisions. The 'vestibulocerebellum' is composed of the flocculonodular lobe and its connections with the lateral vestibular nuclei. It is responsible for postural maintenance. The vermis and intermediate zones along with the fastigial and interposed nuclei together forms the 'spinocerebellum'. It receives sensory information (limb proprioception) from the spinocerebellar tract and dorsal column. It sends efferents to the rubrospinal, vestibulospinal and reticulospinal tracts, which play a role in adaptive motor co-ordination based on sensory information received.

The cerebrocerebellum which is composed of the dentate nuclei and lateral hemispheres, connects the cerebrum via the pontine nuclei and thalamus, as the name suggests. It plays a role in planning and timing of movement based on afferent signals received from the parietal cortex via the pontine nucleus. Output signals from the cerebrocerebellum are carried to M1 and PMC via the ventrolateral thalamus.

3.2.6. Organization of human motor system

The descending pathway for the human motor system can be viewed as a hierarchical organization (see Figure 3.7). At its apex are secondary motor regions comprised of the premotor cortex and associative supplementary motor regions located in the cerebral cortex. They are responsible for motor planning based on sensory feedback, experience and current goals. Planning is then translated to action by the motor cortex and the brain stem, the former is also located in the cerebral cortex. The cerebellum and basal ganglia modulate motor planning and generation of movement. At the bottom lies the spinal cord, which transmits all motor output signals distally to limbs.

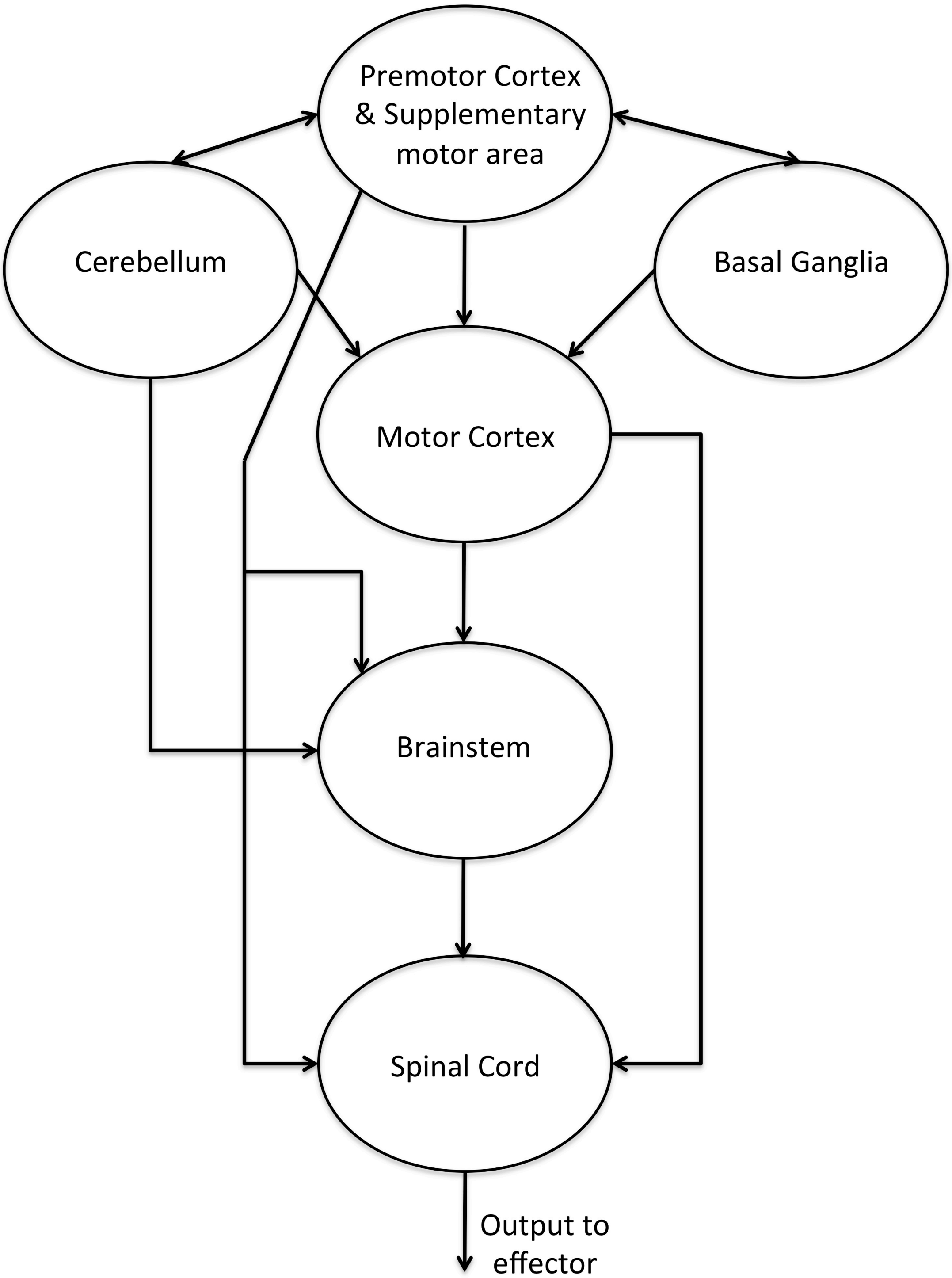

Figure 3.7. Flow chart illustrating hierarchical organization of the descending pathway for the human motor system (picture adapted from Schmidt et al[22]).

3.3. Functional Neuroimaging

Traditional methods of understanding brain function have relied on brain injury or lesional studies. The brain is capable of structural and functional modification in response to external and internal stimuli. This property referred to, as *'neuroplasticity'* is lifelong and not restricted to developmental changes during childhood. Learning or practice of motor tasks has been linked to neuroplasticity at both cortical and sub-cortical levels[25]. Neuroimaging now represents the predominant method in detecting learning related cognitive and behavioural changes[24] manifested by structural and functional changes. Functional neuroimaging, differs from structural neuroimaging as it measures brain activity in response to specific cognitive functions. Brain activity is inferred from metabolic or biochemical information mapped on anatomical locations of the brain allowing localization of neural function[152]. A range of non-invasive methods have been developed since Hans Berger in 1929 first recorded from the human scalp surface an electroencephalogram (EEG)[153]. The neurophysiological basis of non-invasive methods to capture brain activity are dependent on either electrophysiological or haemodynamic principles of energy consumption by the brain.

3.3.1. Neurophysiological basis of Functional Neuroimaging

The underlying principle for brain activity, similar to any other organ is its dependency on energy consumption. Neuronal activity leads to corresponding increase in metabolic activity within the same brain area, with close temporal and spatial correlation[154]. The brain is active even during rest noted by electrical activity from EEG during resting state. This explains the enormous energy consumption by the human brain (approximately 20% of the total human body basal oxygen consumption) when the adult brain constitutes only 2% of the total body mass[155].

Most of brain energy consumption has been attributed to activity in presynaptic terminals or conversion of the neurotransmitter glutamate to glutamine in astrocytes[156]. Glucose is the main energy substrate for the brain where energy in the form of adenosine triphosphate

(ATP) is derived firstly by the glycolyic pathway and later by oxidative phosphorylation[152]. Oxidative metabolism of glucose is supported by brain's oxygen consumption ($CMRO_2$)[155]. When brain metabolism increases it is dependent on the supply of oxygen and glucose, and also on the removal of waste products of metabolism that include lactate and carbon dioxide.

3.3.2. Neurovascular coupling

In the late 19th century, Roy and Sherrington[157] through animal studies suggested that chemical products of cerebral metabolism result in variations to the calibre of cerebral vessels. Thereby this intrinsic mechanism of the brain known as *'neurovascular coupling'*[158,159] regulates the rate of cerebral blood flow (CBF) to meet the energy consumption requirements of the brain[160]. Subsequently, several animal studies[161,162] and more recently human studies[163,164] have correlated brain electrical activity to haemodynamic response. Contrary to the traditional notion that CBF is regulated directly by energy demands of brain tissue, other factors linked to neuronal signalling have been hypothesized. Based on the brain region various neurotransmitters such as glutamate, gamma-aminobutyric acid (GABA) or enzymes such as cyclooxygenase-2 are implicated[156]. To a lesser degree extracellular potassium ion concentration has shown to increase CBF[165].

3.4. Classification of Neuroimaging modalities

Functional neuroimaging techniques as described above measure neuronal activity, yet they defer in how they measure it. Non-invasive methods can be broadly classified as *'direct'* based on electrophysiology techniques or '*indirect'* methods based on haemodynamic techniques.

3.4.1. Direct neuroimaging modalities

Electroencephalography (EEG) measures synchronized synaptic electrical activity of cortical neuronal clusters from the scalp surface using electrodes graphically represented as an encephalogram. Magnetoencephalography (MEG), instead measures changes in magnetic fields in relation to electrical activity.

3.4.1.1. Electroencephalography (EEG)

Richard Caton an English scientist first measured cortical electrical activity of rabbits and monkeys in 1875 using a galvanometer. Applying this principle Hans Berger an Austrian neuropsychiatrist in 1924 first measured encephalograms from humans [166]. EEG mainly measures electrical currents that flow during synaptic excitations of the numerous pyramidal neurons in the cerebral cortex. The underlying physiological mechanism is initiated by an action potential at the pre-synaptic axons to release neurotransmitters. At the synaptic junction neurotransmitters move across the synaptic cleft to bind at the post-synaptic dendritic membrane causing ion channels to open that subsequently leads to the inflow of positive ions. This creates an extracellular voltage, which is more negatively charged in extracellular space around the neural dendrites referred to as the 'sink'. The ions then exit leaving a positively charged extracellular space called the 'source' leading to the creation of an electrical 'dipole'[153]. Individually, a single electrical dipole is too weak to be detectable from the brain surface. However, the aggregate of large synchronized clusters of electrical activity are detectable from the brain surface[167].

Electrodes are placed on the scalp in standardized specific locations using a 10/20 International system[1]. Using this method, adjacent electrodes are placed over the scalp at distances of 10% and 20% of the anterio-posterior or left to right distances of the skull. Positioning is reliant on two bony anatomical landmarks of the skull namely the nasion and the inion (see Figure 3.8). The 10/5 system[168] uses the same principle but allows higher resolution systems to position up to 256 electrodes. Conventional systems require application of electrode gel in order to provide a conductive path between the scalp and electrode. However, application of gel is laborious and the process is untidy making it unacceptable to some participants fro our personal experience. As an alternative, dry electrode EEG systems have been developed recently that do not require application of gel and may improve usability, but have not yet supplanted gel-based systems in terms of quality of captured electrical signals[169].

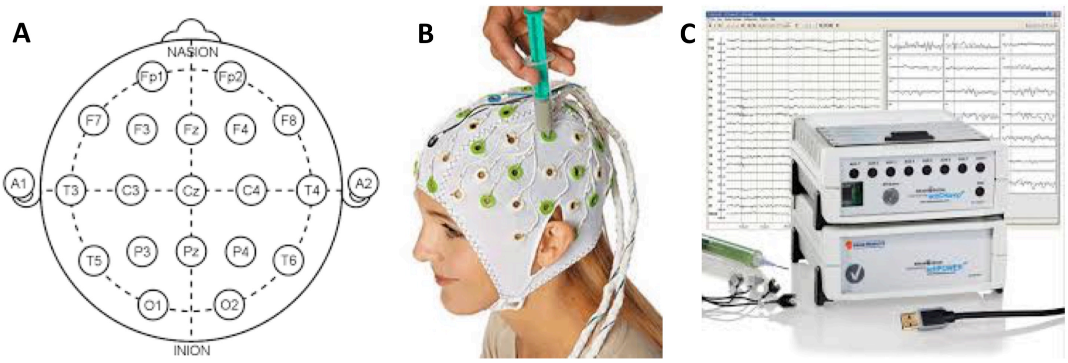

Figure 3.8. Experimental set-up for EEG involves (A) usage of 10-20 system[168] to guide positioning of EEG electrodes. (B) Electro-conducive gel is applied at electrode location and (C) measured signals are amplified using EEG systems (Brain Products GmbH[54]).

Encephalograms vary based on frequency, shape and amplitude of waveform. Commonly observed normal waveforms based on frequency in descending order are as follows:

1. Beta (>13 Hz): observed in alert adults.
2. Alpha (8-13 Hz): seen in adults during a relaxed state.
3. Theta (5-8 Hz): normal when observed during sleep in adults. However in children it is observed during wakefulness.
4. Delta (0.5-4 Hz): are observed during deep sleep.

EEG provides excellent temporal resolution up to milliseconds, is relatively portable, resistant to motion artefacts and inexpensive when compared to other neuroimaging modalities. However, the spatial resolution is poor since electrical activity is measured from synchronized activity of ten thousands of neurons. Pinpointing the anatomical location of the signals origin is challenging as not only is it indirectly derived but cannot differentiate between two spatially adjacent yet functionally different brain regions. Measuring brain activity of deeper brain structures via non-invasive EEG systems remains an insurmountable challenge. In addition, EEG signals may be contaminated by frontal and ocular muscle activity requiring analytic methods for it to be filtered.

Despite the benefits of temporal resolution, cost-effectiveness and portability, EEG has several limitations specific to cognitive monitoring during laparoscopic task performance. It is difficult to control frontal and ocular activity when surgeons are undertaking complex visuospatial tasks. Secondly the poor spatial resolution does not offer the possibility of distinguishing the origin of neural activity from neighbouring multiple motor regions. Lastly the requirement of gel application makes the task unappealing to some participants thereby limiting longitudinal studies.

3.4.1.2. Magnetoencephalography (MEG)

Electric currents from neuronal depolarization of apical dendrites of the cerebral cortex[160] generate small magnetic oscillations. Direction of the magnetic field can be understood by the right-hand rule (see figure 3.9) where the direction of the current is along the axis of the thumb and direction of the magnetic field is orthogonal to the electrical current in the direction of the remaining four curled fingers. Therefore unlike EEG, which detects neuronal electric currents radial and tangential to the skull, MEG cannot detect electric currents perfectly radial to the skull.

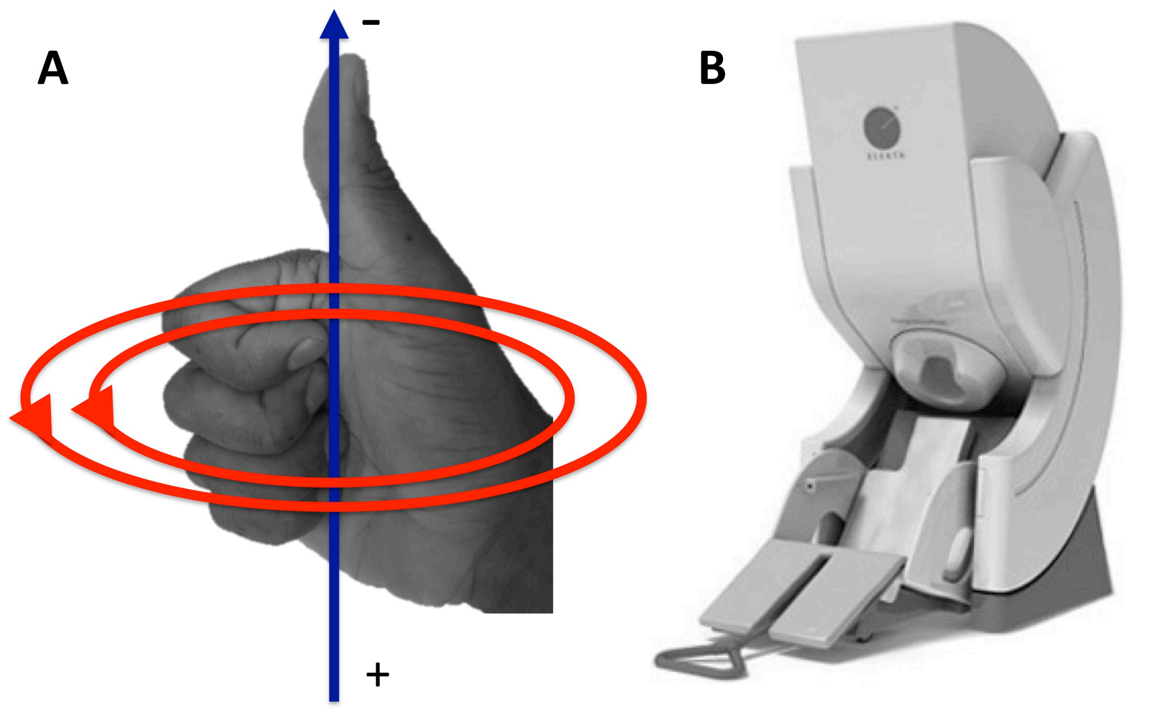

Figure 3.9. Right thumb rule (A) graphically illustrates the vector of electrical current (blue) along the direction of the thumb is orthogonal to vectors of magnetic fields (red) along the curled fingers. (B) MEG equipment requires participants to be in a restricted position where the head is positioned close to the sensors.

Magnetic fields from brain activity are minute in comparison to the earth's magnetic field or the fields produced by an urban environment[170]. Therefore, MEG systems require very sensitive magnetometers in a tank containing liquid helium to enhance superconductivity. In addition, magnetically shielded room are required rendering cost of equipment extremely expensive. The device (see figure 3.9) is large, not easily transportable, restricts the participant to a particular posture and also limits the interaction with ferromagnetic tools. Therefore, an experiment examining human motor behaviour with ferromagnetic tools, for example, using surgical instrumentation would technically not be feasible. This notwithstanding, the advantages of MEG are it provides excellent temporal resolution and better spatial resolution than EEG. Magnetic fields unlike electrical activity do not suffer from 'shunting effects' where conductivity from the brain is dampened by overlying structures such as the skull and cerebrospinal fluid causing distortion.

3.4.2. Indirect neuroimaging modalities

Indirect methods share commonality in that they all measure haemodynamic response to neuronal activation via the phenomena of *'neurovascular coupling'*. Over the years a number of techniques have been discovered which include Positron Emission Tomography (PET), Single-Photon Emission Computed Tomography (SPECT), functional Magnetic Resonance Imaging (fMRI) and functional Near-Infrared Spectroscopy (fNIRS).

3.4.2.1. Positron Emission Tomography (PET)

PET is reliant on intravenous administration of unstable radionuclide tracers labelled with positrons which are positively charged particles than upon encounter with negatively charged electrons undergo annihilation to release two 511 –KeV photons in opposite directions[171] (see Figure 3.10). These photons exiting the body can be detected by a ring of detectors positioned around the subject. Based on the haemodynamic response, brain activated regions that have greater blood supply will lead to greater release of photons. For this purpose ^{15}O labelled water ($H_2{}^{15}O$) which possesses a short half-life (123 seconds) is injected into the blood stream and images are taken approximately every 40 seconds to measure regional changes in cerebral blood flow between images captured[170].

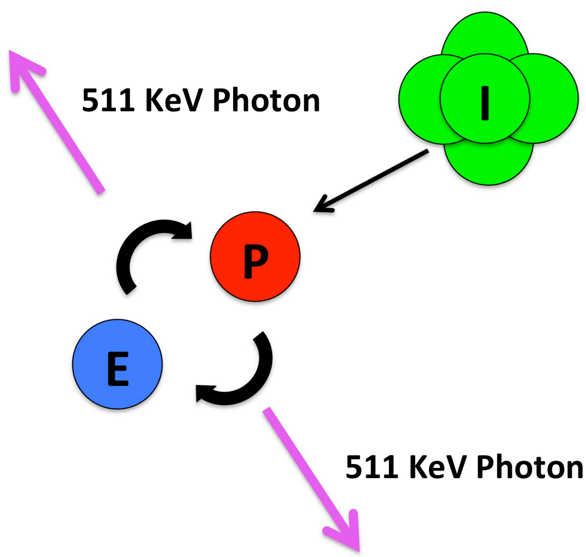

Figure 3.10. Principle underlying PET scan is diagrammatically represented. Radioisotopes (I), release protons (P) which on collision with electrons (E) leads to annihilation resulting in release of photons in opposite directions.

In comparison to other neuroimaging modalities PET has several disadvantages when specifically considering a technique to apply to studying brain function during laparoscopic skills acquisition These include the use of intravenously injected radioactive tracers, poor temporal and spatial resolution, and is relatively expensive. Additionally, the lack of portability, restriction to a lying position renders it unsuitable for the purpose of measuring cognitive behaviour of surgeons during performance of technical tasks.

3.4.2.2. Single-photon emission computed tomography (SPECT)

SPECT technique is similar to PET in the used of radiolabelled tracers and detection of photons. Brain activity is depicted on a colorimetric scale where red depicts greater activity and blue lower activity (see Figure 3.11). However, radioisotopes such as Tc^{99} (technetium) and I^{123} (iodine) with longer half-lives are used to evaluate regional cerebral blood flow. In addition, SPECT does not require cyclotrons as it measures emission of photons alone and unlike PET does not account for simultaneous emission of photons in opposite directions coincident in time[172]. As a result spatial resolution is poorer in SPECT due to the lack of localisation. Despite it being comparatively more flexible by the use of longer half-life radioisotopes and cost-effective than PET scans, the need to use radioactive substances, lack of portability, postural restriction makes it infeasible for interrogating cognitive behaviour of surgeons.

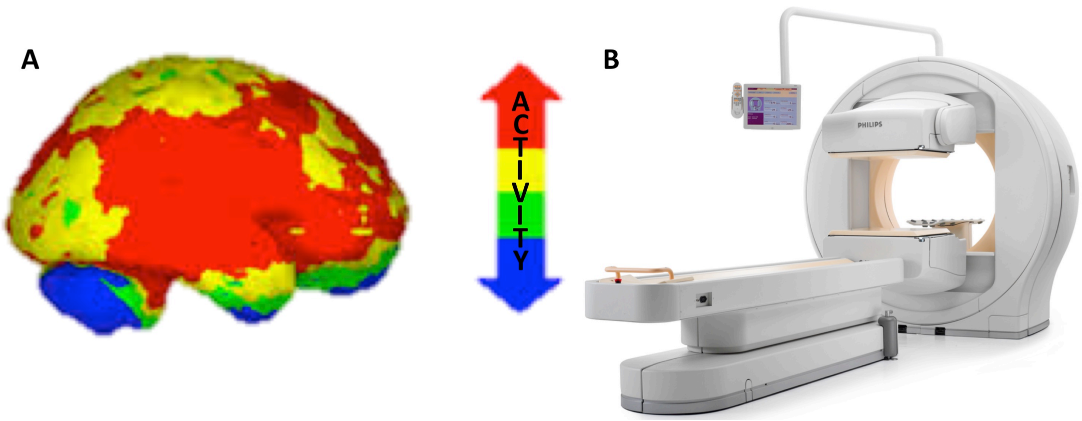

Figure 3.11. Graphical illustration of (A) brain activity using a colorimetric scale employed by PET/SPECT and (B) SPECT device.

3.4.2.3. Functional magnetic resonance imaging (fMRI)

The emergence of fMRI contemporaneously with PET revolutionised the field of functional neuroimaging. It is based on the physics of atomic behaviour in water discovered by Felix Block and Edward Purcell in 1946. Protons when placed in a magnetic field behave like tiny bar magnets aligning parallel to the magnetic field[171]. Deoxyhaemoglobin (HHb) and oxyhaemoglobin (HbO_2) have different magnetic properties whereby the former is paramagnetic and the latter is not. Hence HHb serves as an MRI contrast agent noted in animal studies conducted by Ogawa and colleagues[173] which led to the labelling of it as a *'blood oxygen level dependent contrast'* (BOLD).

HHb concentration is dependent on physiological parameters namely regional cerebral blood flow (rCBF), regional cerebral volume and metabolic oxygen consumption of oxygen[174]. In a region of brain activation, concentration of deoxyhaemoglobin is relatively decreased which can be visualised as high signal intensity when T-2 weighted images are obtained[160].

fMRI has several advantages over other techniques, owing to BOLD contrast it does not require intravenous administration of radioisotopes, exposure to radiation, and provides excellent spatial resolution (3-6 mm) in most applications. However there is a trade-off for spatial resolution with temporal resolution. The temporal sensitivity of fMRI is linked to the lag in haemodynamic response (3 to 6 seconds) relative to the neuronal response. In addition, fMRI is relatively expensive, lacks portability, constrains body posture, is adversely influenced by motion artefacts and restricts the use of ferro-magnetic device. Therefore, despite its many advantages, it is infeasible as tool to measure brain activity in ambulant surgeons when performing the task in an upright position using ferro-magnetic instruments.

3.5. Functional near-infrared spectroscopy (fNIRS)

This method exploits the principles of near-infrared spectroscopy (NIR) to map functional cortical activity[175] indirectly by detecting changes in HbO_2 and HHb. This technique is also known as near infra-red imaging (NIRI), diffuse optical topography (DOT), diffuse optical imaging (DOI) and diffuse optical tomography[176].

3.5.1. Origins and fundamentals of functional near-infrared spectroscopy

Glenn Millikan the son of Nobel prize winning physicist Robert Millikan, first used NIR as an oximeter as a method to measure myoglobin-oxygen levels toward developing a system for oxygen delivery in World War II fighter pilots[177]. About thirty-five years later, in 1977, a former student of Professor Britton Chance, named Frans Jobsis developed an fNIRS prototype and coined the term *'Niroscope'*. Professor Jobsis reported the ability of NIR wavelength range of light (650-1000 nm) to penetrate brain tissue and allow real-time, non-invasive detection of haemoglobin oxygenation[178].

Spectroscopy is founded on the study of light signals. Human tissues are relatively transparent to light within the NIR wavelength a phenomena termed as the '*optical window*', simply demonstrated by shining a torch light on top of a hand in a dark room. When NIR light is shone over the scalp it passes through skin, skull, cerebrospinal fluid and underlying brain tissue. NIR light is either absorbed by light absorbing pigmented compounds named *'chromophores'* or scattered through brain tissue. NIR light is a hundred time more likely to scatter than be absorbed[179]. Attenuation of NIR light due to absorption by the dominant chromophore haemoglobin located in small vessels (<1mm calibre) of the brain is used to estimate changes in chromophore concentration[175]. Haemoglobin's absorption spectra is dependent on the level of oxygenation. The two dominant chromophores for NIR wavelength namely HbO_2 and HHb have different absorption properties and fortuitously are biological surrogate of cortical activation [176]. Total haemoglobin (HbT) is a derived value from the summation of HbO_2 and HHb.

As discussed earlier (section 3.3) brain activation is accompanied by increases in rCBF and regional cerebral cerebral oxygen metabolic rate ($rCMO_2$) where the former exceeds the latter[180]. As a result, NIRS is capable of measuring functional brain activity by detecting a task evoked increase in HbO_2, at times a rise in HbT and a corresponding decrease in HHb (see figure 3.13). Amongst the three Hb species, HbO_2 is the most sensitive indicator of rCBF as it always correspondingly increases with it[181] whereas HbT is not sensitive to small changes in rCBF and changes in HHb concentration are reliant on venous blood oxygenation and volume[182]. The vascular response observed with fNIRS is comparable to the BOLD response measured in fMRI, however fNIRS differs in that it measures both haemoglobin species[183]. More recently from simultaneous dual neuroimaging (fNIRS and fMRI) it has been observed that of the two signals detected using fNIRS, HHb more closely resembles fMRI BOLD[184].

The path length of NIR light between an optode pair of emitter and detector is complex and longer than the geometrical distance between the pair due to the distortion caused by the heterogeneity of tissue it passes through namely scalp, skull and cerebrospinal fluid. It forms a banana shaped profile within grey matter of the cortex (see figure 3.12) characterised by the narrow ends and convexity towards the centre of the brain[175].

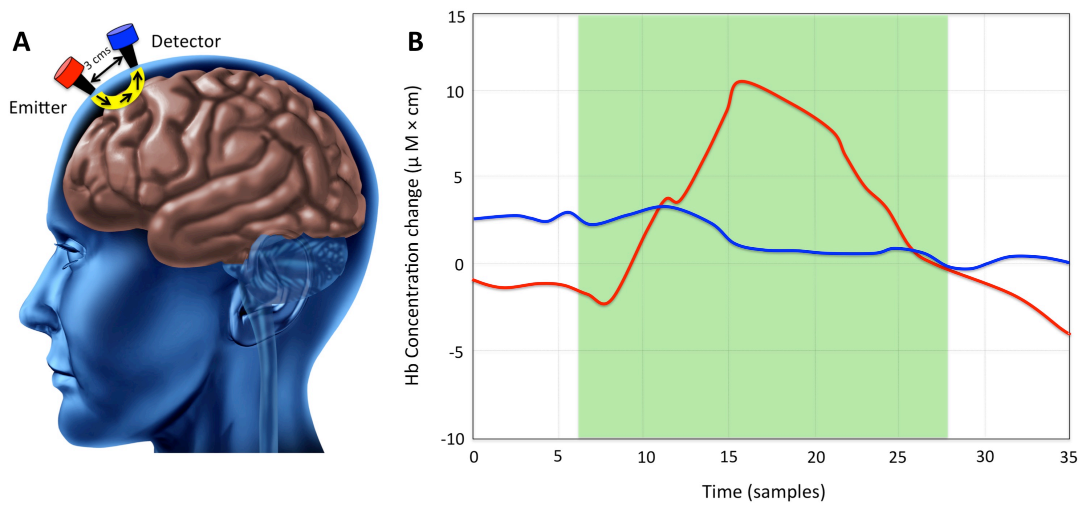

Figure 3.12. Trajectory of NIR light from an emitter (red) to detector (blue) follows a banana shape as graphically illustrated (A). Brain activation from a single fNIRS channel in response to a stimulus is detected by an increase in relative concentration of HbO_2 and a corresponding drop in HHb during the stimulus phase (area shaded in green).

3.5.2. Basic theory of Modified Beer-Lambert Law

In order to quantify changes in concentration of absorbing haemoglobin species, a photon transportation model to measure optical attenuation was developed by Delpy et al[185] called the Modified Beer-Lambert Law (MBLL). It is based on the original Beer-Lambert law, which was valid in only non-scattering media. However in biological media light scatters warrants modification[186]. MMBLL is equated as follows:

$$A = \log_{10} [I_o/I] = B \times c \times d \times \text{DPF} + G$$

In this equation A denotes optical attenuation, I_o the emitted light intensity, I transmitted light intensity, B is the specific extinction co-efficient of the chromophore (μmol^{-1} × cm^{-1}), c is absorbing chromophore concentration (μmol) and d represents distance between emitter and detector optodes. Differential path factor (DPF) is a multiplier and Geometry-dependent factor [G] is an additive to the original law where DPF is the true optical distance of photons, which can be measured unlike G a factor dependent on geometry and represents light intensity loss associated with scattering[187].

3.5.3. fNIRS instrumentation

3.5.3.1. History of fNIRS instrumentation

Since the discovery of NIRS as a brain-mapping tool by Professor Jobsis in 1977 there has been several technological advances in the field (see Figure 3.13). The first single channel continuous wave fNIRS system named NIRO-1000 built in 1989 by Hamamatsu Photonics (Hamamatsu City, Japan) was borne out of a four wavelength fNIRS system invented by Cope and Delpy in 1988[188]. Between year 1994 to 2001, several companies and institutions developed multi channel fNIRS systems. At first they were low-density systems capable of measuring brain activity from 10 channels[189] and later systems were capable of monitoring neuronal activity from up to 128 channels[175]. The use of multi channels enabled analysis of spatial variation from 2D neuroimages termed as 'optical topography' by simultaneous measurement of brain activity from multiple regions, each between a paired optode of light emitter and detector usually separated by 3 cms (see Figure 3.13). If an adequate number of optodes is placed on the head, 3D images of the brain termed as 'optical tomography' can be created, however they are reliant on highly sophisticated image reconstruction algorithms and are feasible to a greater degree in new born infants due to limitations in penetration of NIR light[190].

One of the drawbacks of fNIRS systems were the subjects' freedom of movement was restricted by the length of the optical fibres. Although, the first wireless telemetric system was developed by Hoshi et al[191] in 2001 for clinical studies in children, it was only ten years later that high density multi channel wireless fNIRS systems were developed[175].

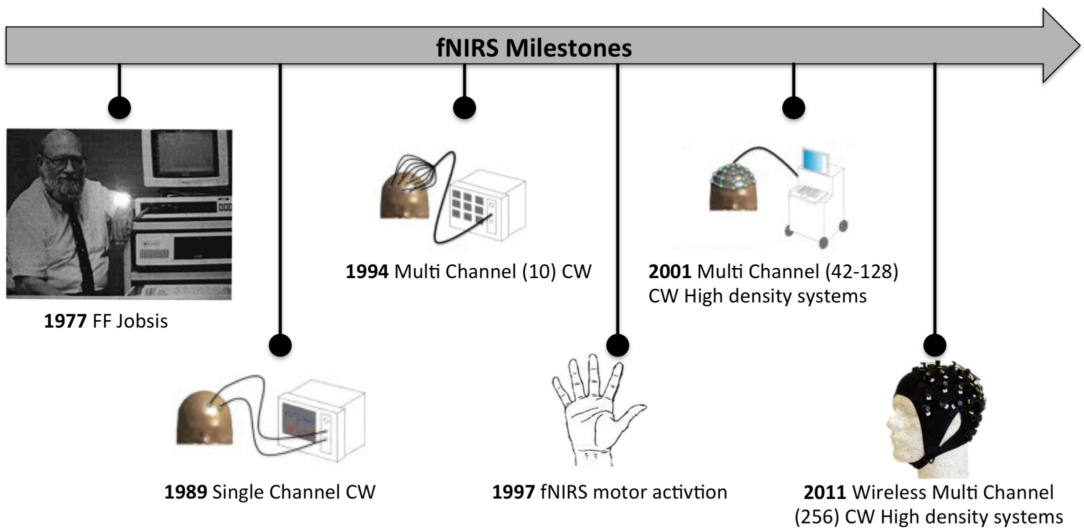

Figure 3.13. Major milestones in development and application of fNIRS instrumentation are graphically illustrated (figure adapted from Ferrari et al[175]).

3.5.3.2. Types of fNIRS instrumentation

Over the years several techniques have been developed to elucidate neuronal activity related haemodynamic response using fNIRS and can be categorized as follows:

1. **Continuous wave**: Continuous wave fNIRS systems emit light at a constant intensity and measures changes in intensity of transmitted light. Employing this method, relative changes in chromophore concentration can be measured but not absolute

values unlike the next frequency and time domain systems. In addition, they are lower in cost and complexity as a result more commercially available.

2. **Frequency domain/resolved**: In this technique, NIR light is emitted constantly, although the intensity of light is modulated in a sinusoidal fashion (see Figure 3.14). Information is then derived from intensity of light transmitted and phase shift which corresponds to time of flight[186].

3. **Time domain/resolved**: Time-domain systems employ a short source of light (in picoseconds) and a fast time-resolved detector to determine photon pathlength[190]. Although this method provides the most information it is very expensive due to requirements of sophisticated technological systems that incorporate high-speed emitters and detectors.

Properties	**Types of fNIRS instrumentation**		
	Continuous wave	Frequency domain	Time domain
Temporal resolution (Hz)	≤100	≤50	≤10
Spatial resolution (cm)	≤1	≤1	≤1
Penetration depth	Shallow	Greater depth	Greater depth
Discrimination between cerebral and extra-cerebral tissue	Feasible	Feasible	Feasible
Instrument size	Small – Large	Large	Large
Cost	Low-High	High	Very high
Instrument stabilization	Not required	Not required	Required
Portability	Easily portable	Feasible	Feasible
Telemetry	Feasible	Not feasible	Not feasible
Measurement parameters HbO_2, HHb & THb	Relative changes	Absolute value	Absolute value
Scattering, absorption coefficient and pathlength measurement	No	Yes	Yes
Tissue HbO_2 saturation measurement %	No	Yes	Yes

Table 3.1. Advantages and disadvantages of fNIRS instrumentation techniques are described based on each property (adapted from Ferrari et al[175]).

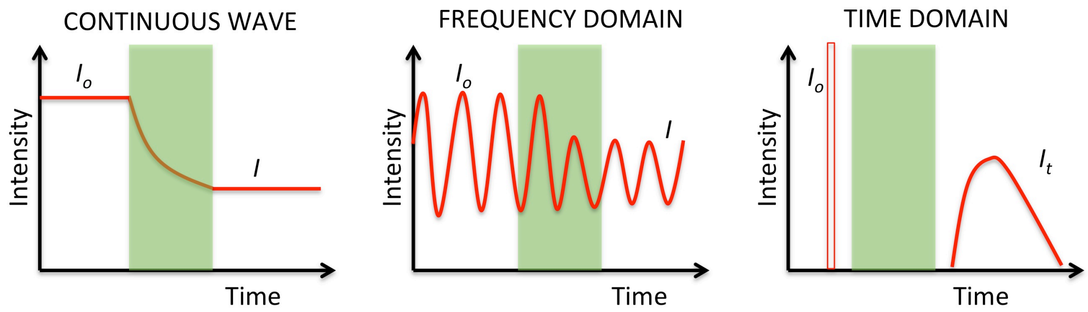

Figure 3.14. Waveforms of three different NIRS techniques namely continuous wave, frequency domain and time domain are graphically illustrated. I_o denotes intensity of incident light and I denotes detected light intensity through brain tissue shaded in green. In time domain technique I_t denotes temporal spread of detected light signal.

3.5.3.3. Advantages of fNIRS

Colier et al in 1997[192] used a single channel fNIRS system to interrogate cortical oxygenation in relation to cyclic compound limb movements. Since then, fNIRS has emerged as a useful method for interrogating neural correlates that underlie simple motor tasks such as finger tapping to complex motor tasks such as apple-peeling[193] or surgical manoeuvres[194].

fNIRS resembles EEG in terms of experimental setting and fMRI in that it measures haemodynamic response[183]. It is less susceptible than EEG to data corruption by movement artefacts[195], offers better localization and spatial resolution of brain activation (see Figure 3.15). Despite poorer spatial resolution than fMRI, the technique provides superior temporal resolution and a more complete capture of brain activation related haemodynamic response as it measures relative changes in both chromophores (HbO_2 and HHb). In addition, it is portable, more economical, does not constrain the experimental environment or restricts subject's movement, which affects fMRI more than fNIRS. Compared to PET and SPECT it avoids the use of ionising radiation making it a safe and repeatable technique for studying surgeons. More importantly, fNIRS does not constrain the environment or movement essential for surgeons to perform complex manoeuvres. In the next chapter the application of fNIRS in monitoring surgeon's cognitions during a variety of tasks are outlined (see Chapter 4). As a result of these features fNIRS is the chosen neuroimaging method for this thesis to investigate neuroplasticity underlying complex motor skills acquisition in surgery.

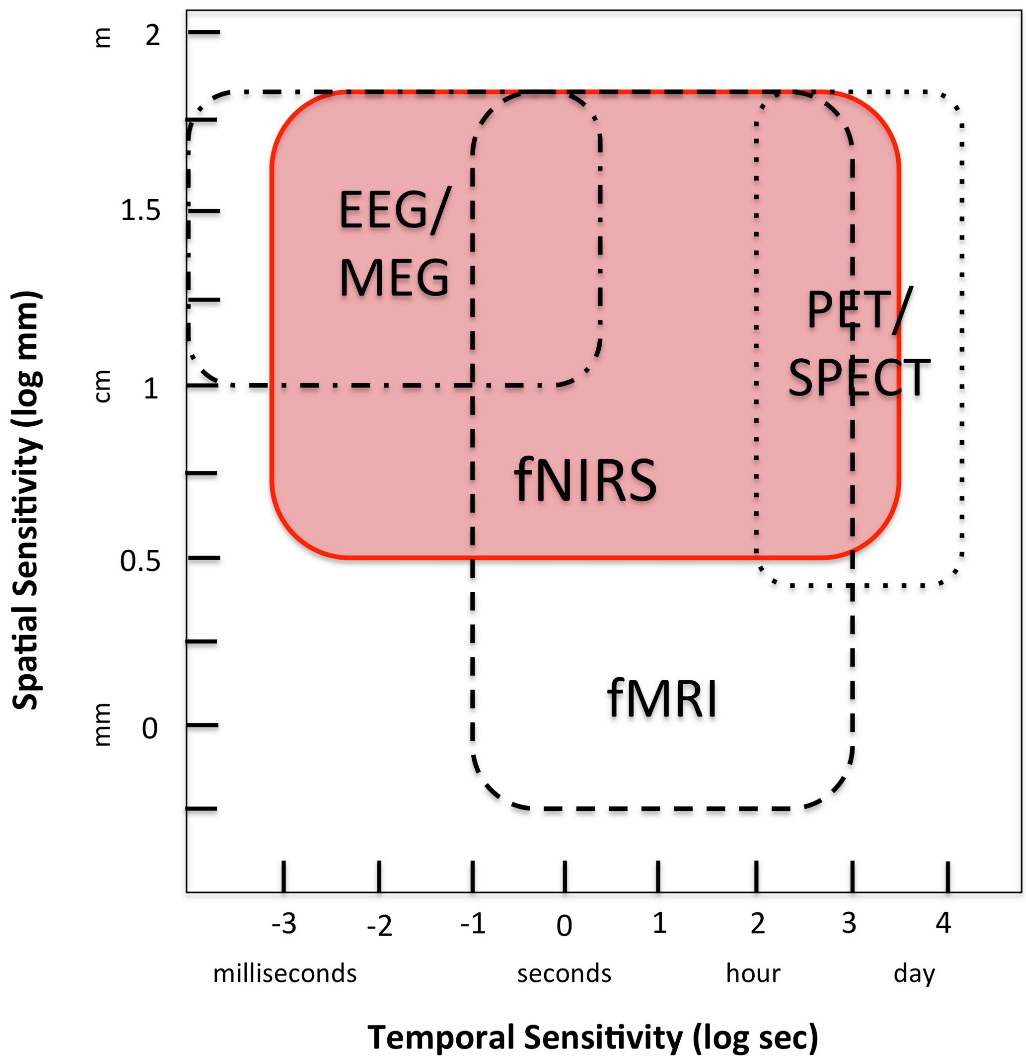

Figure 3.15. Spatial and temporal resolution of NIRS (shaded red) is compared to other neuroimaging modalities (adapted from Strangman et al[176] & Walsh et al[196]).

3.5.3.4. Limitations of fNIRS

Most commercially available fNIRS instruments including the one used in this thesis use continuous wavelength as a light source from the emitting optode. The receiving optodes i.e. detectors cannot quantify attenuation due to unknown light loss from tissue scattering and increased path-length[197]. However changes in attenuation of light can be measured using MBLL allowing calculation of relative changes in concentration of HbO_2 and HHb but not absolute change in haemoglobin content. Therefore experimental design requires a block-design, i.e. comprised of multiple rest (inactive) and short task (active) periods. A number of cortical and sub-cortical regions are responsible for acquiring and executing skilled movements, yet fNIRS is capable of interrogating only relatively superficial cortical structures. A standard inter-optode distance of 3 cm is capable of an imaging depth of only 1.5 cm[176]. Additional factors such as thickness of subject's scalp [198], skin pigmentation and density of hair follicles can affect absorption of NIR light causing interference in signals[199].

fNIRS possesses the same limitations as EEG, MEG and PET in that it can not provide direct anatomical information. An extrinsic frame of reference is required to allow inter-subject and longitudinal analysis. Similar to EEG methodology, anatomical reference is derived by employing the 10/20 International system[1]. High density fNIRS systems employ 10/10 and 10/5 systems for up to 329 scalp landmark positions[200]. However, this method of anatomical reference is more complex in fNIRS as the location of each channel is dependent on the placement of its corresponding paired emitter and detector optode, where as in EEG it is more straightforward with electrode localisation alone. One method is to register fNIRS data to a subject's own structural fMRI image[201]. However, this method is expensive and reliant on access to an fMRI scanner. An alternative method for standalone fNIRS data is to adopt a probabilistic registration method that utilises a standard brain template to approximate the most likely anatomical location underlying each channel[202,203].

3.5.3.5. Hitachi ETG-4000

The Hitachi ETG-4000 (Hitachi Medical Corp., Tokyo, Japan), a commercially available continuous wave fNIRS system released in 2003 is used as the interrogating functional neuroimaging modality for this thesis. It is capable of measuring relative changes in HbO_2 and HHb in up to 52 channels. Optodes are placed in a grid of thermoplastic holders with an inter-optode distance of 3 cm, where laser diodes (emitters) emit light in two wavelengths of 695 nm and 830 nm in order to measure simultaneous changes in both Hb species. Optodes carry signals to and fro a portable mainframe via fibre-optic cables (see Figure 3.16).

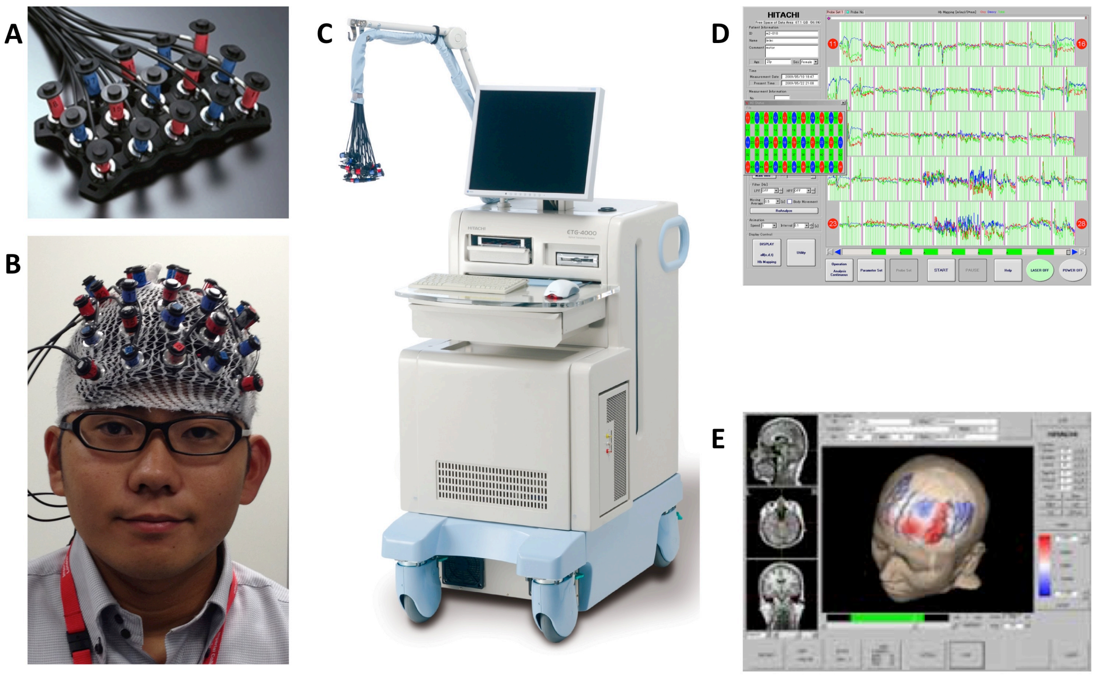

Figure 3.16. Components of the ETG-4000 and real-time data extraction from fNIRS signal are illustrated. (A) Optodes comprised of emitters (red) and detectors (blue) are placed in thermoplastic holders in a grid (3 × 5) with a standard inter-optode distance of 3 cm. (B) Optodes are precisely positioned using the 10/20 International system[1] and held in place with the aid of caps created from elastic netting (Surgifix Inc., San Juan, Puerto Rico) . The main frame of the ETG-4000 (C) is portable by wheels, composed of a processing unit, liquid crystal display, optical fibres, optodes and a 3D digitizer stowed behind the mainframe. Data extracted from detected signals are depicted (D) in waveform for each channel or (E) topographically after creating a 3D mesh from the participant’s anatomical landmarks and channel location.

Data is sampled at 10 Hz and displayed graphically in real-time by waveform or two-dimensional topographic format. Alternatively, three-dimensional topographic visualisation is feasible by registration of optodes using a 3D digitiser (Isotrak, Polhemus, Vermont, USA). This method provides probabilistic registration of each channel. The 3D mesh is initially created by first registering at least four anatomical scalp landmarks namely the nasion, inion, pre-auricular points (right and left) and Cz followed by individual optode position[203].

3.6. Cortical Connectivity

Neuroimaging represents the predominant investigatory technique in behavioural and cognitive science. Mapping functional brain activity can either follow the guiding principle of functional *segregation* or *integration*[24]. Functional segregation i.e. localization of brain activity in relation to task-related function was the initially the predominant choice. However, multiple brain regions are known to underlie a given task function and the concept to attribute function to a specific region in isolation rooted in the field of phrenology was challenged in as early as 1881[204]. Setting aside the conceptual issue the other challenge was to develop a method to measure functional segregation, i.e. the interaction of multiple brain regions that underlie a given task function. Since the 2010, more published papers have adopted connectivity over activation models to present neuroimaging data, suggestive of a shift in the emphasis from functional segregation to integration[28].

Connectivity can be broadly classified as follows on the basis of neuroimaging technique and mathematical modelling for interpretation of derived neuroimaging data (see Figure 3.17).

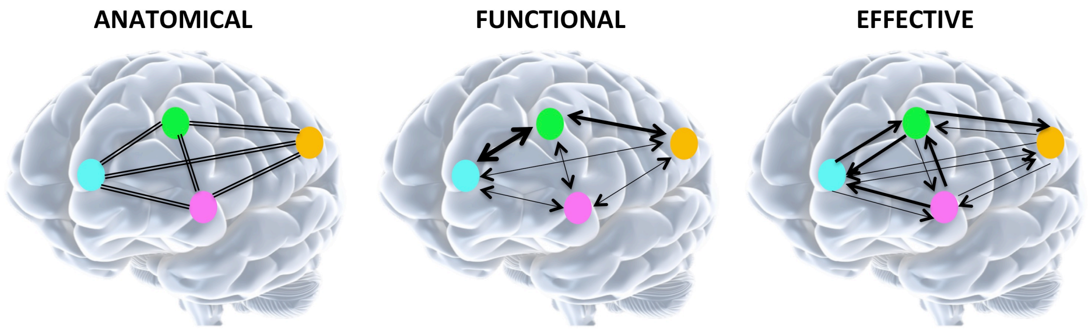

Figure 3.17. Cortical connectivity is based on anatomy (white fibre bundles), functional connectivity (statistical correlations) and effective connectivity (directional effect of one region on another). Flow of connectivity and its strength is depicted by the direction and thickness of black arrows.

3.6.1. Anatomical Connectivity

This requires Diffusion Tensor Imaging (DTI) otherwise known as diffusion weighted magnetic resonance imaging (DWMRI) an imaging method sensitive to diffusion of water within in-vivo tissues, indirectly providing tractographic information on tissue microstructure thus enables characterization of white matter fibre bundle trajectory[205]. Anatomical connectivity has employed to develop new insights into clinical[206] and psychiatric conditions[207]. Recently, Jarbo et al[208] used DTI to identify convergence in anatomical and functional human cortico-striatal pathways that underlie motor learning in healthy individuals.

3.6.2. Functional Connectivity

Functional connectivity is defined as the temporal correlations between spatially remote neurophysiological events[209] derived by analysis of neuroimaging time series. It is based on the principle that brain regions displaying similar brain signals are connected. Gerstein et al[210] noted during analysis of simultaneously recorded action potentials, correlation can arise from multiple reasons one of them being a common afferent input. Statistic correlation of time series signal from one neuronal area with another forms the basis of functional connectivity.

Similar to fMRI and EEG, which uses voxels and electrodes respectively, in the case of signals fNIRS channels are compared to create a correlation matrix. The matrix can be then used to create a network to allow further analysis. Numerous studies[211-217] have used functional connectivity for fNIRS data interpretation of which the vast majority are during resting state[215,218]. Spontaneous low frequency fluctuations in signal generated during the rest state can be used to measure correlation between different brain regions termed as *'resting state functional connectivity'*[219].

The simplest approach is seed-based used in several fNIRS studies on adults[212-214] and children[211,217]. One channel within a region of interest is considered as a seed and the time course of the seed is compared to all other channels[218]. Alternatively, Principal or Independent Component Analysis (ICA) has been used for fNIRS data to measure functional connectivity, where temporal ICA has been found to be more useful due to the relatively high number of temporal samples in comparison to the number of channels[216]. A tool-box with a graphical user interface (GUI) has been created to simplify and standardize methods[215]. RSFC has been used to map fMRI data in relation to motor skills learning[220,221]. More recently Bajaj et al[222] was able to distinguish unimanual from bimanual finger movements using a 52-channel fNIRS system during rest and motor activity by mapping variance in functionally connected motor regions (SMA, PMC and M1).

3.6.3. Effective Connectivity

An integrated system for instance the brain can alternatively be understood by using effective connectivity models. Effective connectivity is defined as the influence one neuronal system exerts over another[223]. The most prevalent methods for effective connectivity analysis are dynamic causal modelling (DCM), structural equation modelling (SEM) and Granger causality (GC)[28]. SEM is a multivariate linear regression tool and GC is a statistical test for analysing if one time series determines another. Amongst the three DCM was designed specifically for interpretation of neuroimaging data first applied to fMRI data[224] and has been the most popular approach. It is based on a multi-input multi-output neuronal model which accounts for causal interactions mediated by unobservable neuronal dynamics[225]. Tak et al[226] introduced the DCM analytic technique to fNIRS data acquired during motor imagery and execution tasks.

3.7. Conclusion

Learning motor skills leads to relative structural and functional changes in multiple cortical and sub-cortical regions based on the property of neuroplasticity. The human motor system is complex and brain regions have specialized functions that play a role at different stages of motor learning. In this chapter, the neurophysiological and neuroanatomical relevance of each brain region is first described. The frontal lobe is composed of multiple regions associated with various stages of motor learning permitting focused examination. The M1 is responsible for motor execution, whilst the SMA and PMC both located anterior to it are implicated with storage of sequential motor skills. The PFC also located in the frontal lobe anterior to the PMC and SMA is responsible for attention, which is crucial towards supporting learning at early stages according to motor learning theories. Therefore if accessible, the frontal lobe provides an ideal location to interrogate disparity in functional behaviour based on motor skill level.

Functional neuroimaging allows direct observation of brain regions implicated during motor task performance. Neuroimaging techniques either measure brain response directly by detection of electrical activity or indirectly by measuring the haemodynamic response to electrical activity based on the principle of neurovascular coupling. The properties of both direct and indirect techniques have been described in this chapter. Although direct techniques are advantageous in terms of superior temporal resolution, poor spatial resolution limits its applicability in isolating activity from multiple brain regions. Selection of the appropriate neuroimaging modality is also dependent on acceptability to the recruited participants, cost-effectiveness to permit repeated testing in longitudinal studies and maintenance of the experimental environment. As a result PET/ SPECT scan is not suitable for the purpose of this thesis due to the requirement of administration of radio-isotopes. Although fMRI provides superior resolution, it severely distorts the experimental environment requiring participants to perform the task in a lying down, claustrophobic position, can be noisy and does not permit usage of ferro-magnetic instruments.

Amongst functional neuroimaging modalities discussed fNIRS provides the greatest compatibility towards examining surgeons during performance of MIS tasks. The justification to utilise fNIRS as the investigative modality for this thesis are its properties of non-invasiveness and cost-effectiveness making it acceptable to participants and allows multiple measurements a requirement for tracking neuroplasticity in a longitudinal study. In comparison to other techniques it is relatively less affected by motion artefacts and is portable, thereby allowing surgeons to perform the task in their natural environment and posture. In addition unlike fMRI it allows usage of ferromagnetic laparoscopic instruments. Cortical activity is measured by the haemodynamic response to electrical activity quantified by relative changes in both HbO_2 and HHb. Primary and secondary motor regions located in the frontal lobe are optically accessible by fNIRS and have previously been investigated in participants performing non-surgical motor tasks such as apple peeling[227].

Brain behaviours of surgeons have been investigated for almost a decade using described neuroimaging technique to assess expertise and track learning for surgical tasks that involve motor skills (see chapter 4). In the next chapter the available evidence is synthesized and critically appraised to identify lacunae in understanding of underlying surgeon’s cognitions associated with motor learning.

Chapter 4

"Inside the Brain of a Surgeon"

A decade's introspection into neurocognitive assessment of surgeons: A systematic review

4.1. Introduction

Manual dexterity is a vital requisite for a surgeon, where the more skilled are observed to display superior operative outcomes[4]. Mastery of complex surgical skills is an acquired ability of a surgeon to consistently execute precise challenging manoeuvres under dynamic conditions. Therefore, development of objective, reliable and valid methods to describe technical competence has value[16] in order to define expertise, track skills learning and scrutinize factors that have the potential to impact on patient outcome. Minimally invasive surgery (MIS) has revolutionised surgical practice yet poses an additional challenge in technical skills acquisition due to its associated longer learning curves[8,9]. Other current challenges to skills acquisition include reduction in working hours[13] resulting in a shift away from the time intensive Halstedian apprenticeship model to a competency-based curriculum[35,228], increased medico legal claims[14], public scrutiny[15] and adaptation to a constant flux of new surgical technology. As a result, this compound effect is manifest as a reduction in training opportunities[16]. Limitations of laparoscopic surgery such as restricted kinematics, paradoxical movements, amplification of tremors at the instrument tips and loss of depth perception were overcome by the introduction of robotic-assisted surgery viewed as the next revolution in MIS. Although robotic-assisted surgery overcame several of the highlighted limitations, it introduced new challenges such as lack of haptic feedback, high cost requires a longer learning curve than traditional open surgical approaches[12]. To overcome the enlisted challenges there is increased impetus to improve surgical training from various stakeholders such as surgical training programs[229], governing bodies[230] at a global and local level[16]. A number of methods over the last decade have been designed to objectively assess surgical skills[84]. Early methods quantified performance by rating scales (OSATS)[87], yet were dependent on the availability of a trained assessor[231]. Later methods employ time and accuracy as end-points[100,104], which have become a mandatory component for surgical residency curriculum in the United States[232], whilst others recorded motion metrics (ICSAD)[112,233] that are now incorporated into sophisticated surgical simulators (LapMentor™, Simbionix, Cleveland, USA).

Despite the introduction of various assessment modalities, they can not explain overlapping surgical performance between surgeons of different grades[19,234,235] and the inability for residents to transfer skills gained by training in the laboratory to the real operative environment[20]. Additionally, individuals with similar motor performance may be mistaken to have achieved equivalent technical competence and yet may differ significantly in terms of workload and attentional demands that portray true differences in their levels of expertise[21,80].

Human factors research has played an important role to optimize operator performance in other high-risk domains for example in the aviation or automotive industry[236-238]. Similar to surgery, it requires skilful motor manipulation of sophisticated technology under variable conditions. Human factors research otherwise known as ergonomics serves to examine human interaction with technology by analysing behavioural, psychological and physiological parameters[239]. Over the last thirty years, there have been seismic advances in surgical technology to improve surgical performance. However, the application of ergonomics within the domain of surgery to evaluate motor skills acquisition and attentional burden remains rudimentary[21,143]. Learning, expertise and attentional capacity are all inter-related concepts that cannot be easily separated for analysis[22]. Focussed attention has been demonstrated to be a key requisite for expert performance[134,240], yet the investigation into cognitive factors underlying performance has prior to this thesis been parsimonious. Appreciating the importance of measuring cognitive demand, subjective questionnaires (NASA Task Load Index)[82,241] and eye tracking[143,242] have been utilised to capture effortful cognitive processing during assessment of surgical skill. However, they both measure cognitive load indirectly, the first a subjective measure and although eye tracking provides objective physiological information on effortful processing, it is incapable of directly investigating neural substrates responsible for motor skills learning.

Motor skill is a concept separate from motor learning and can be defined as the ability to achieve a target end result with maximum certainty, in minimal time and under minimal expenditure of physical and cognitive energy[22]. The widely accepted definition by Schmidt et al[22] defines motor learning as a set of processes associated with practice or experience leading to relatively permanent gains in the capability for skilled performance. Motor learning is hypothesized to follow three stages (see Figure 4.1) commencing from a novice "cognitive" phase to a learner's "associative" phase and ultimately to an expert "autonomous" phase. At the initial cognitive phase, the novice focuses on the elements of the motor task, learns by trial and error and therefore highly sensory feedback dependent. Outwardly, it is observable by the large number of errors and erratic performance. The second, associative phase is where the learner after practice develop the ability to integrate sensory information with appropriate goal directed movement and also link a sequence of movements. Learners at this stage are still dependent on attention whilst developing strategies to improve performance. Performance is characterized by reduced variability and few errors than the cognitive phase. During the final autonomous phase, motor tasks become automatic, which is characterized by consistently high performance with minimal errors and maximal neural efficiency. Outwardly it is portrayed by effortless performance and stability of performance despite multi-tasking. This ability termed as "automaticity"[45,46] is a true representative of expert status achievement and yet difficult to establish[80].

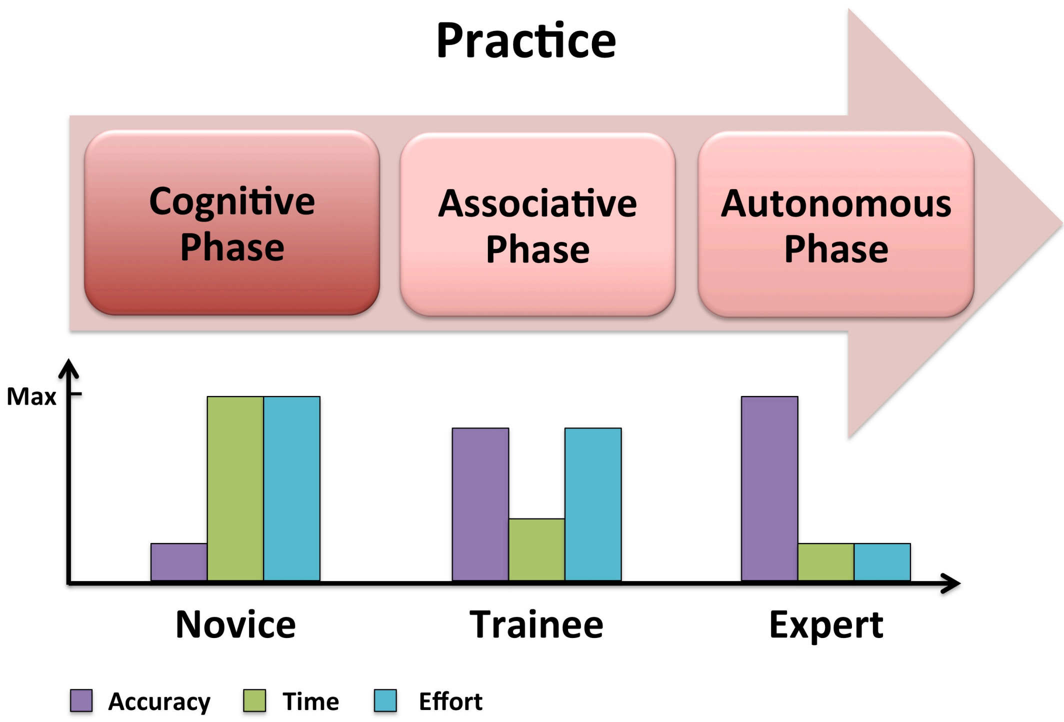

Figure 4.1. Fitts and Posner's[42] three-stage model of learning from novice cognitive phase to expert autonomous phase. Shifts in requirement levels of time, accuracy and effort at each stage are illustrated in relation to practice.

'Neuroergonomics' a convergent field incorporating two disciplines i.e. neuroscience and ergonomics, and describes studies of brain behavior at work[243,244]. It focuses on the underlying brain behavior during physical and mental interaction with tools or technology. They include work, home or leisure activity environments ranging from interaction with appliances at home to use of highly sophisticated technological machinery such as aircrafts and space navigation. The benefits of tapping into this emergent science are vast. It has the unique potential to tap insights into cognitive processes experienced by surgeons for example during skills acquisition or the impact on performance by factors such as fatigue, technological aids, sensory information and task complexity. Information obtained can thus be not only used to facilitate training but also to guide design of safer, user-friendly tools that promote optimal performance.

Advances in the field of functional neuroimaging permit direct observation of brain behavior that underpins motor skills learning[25,245,246]. Available neuroimaging modalities (see Figure 4.2) can be broadly classified into those that directly measure brain electrical activity and those that indirectly measure brain activity by the phenomena of *'neurovascular coupling'* [157]. Simplistically, it's based on the tight coupling between transient neural activity and corresponding rise in cerebral blood flow (CBF)[247].

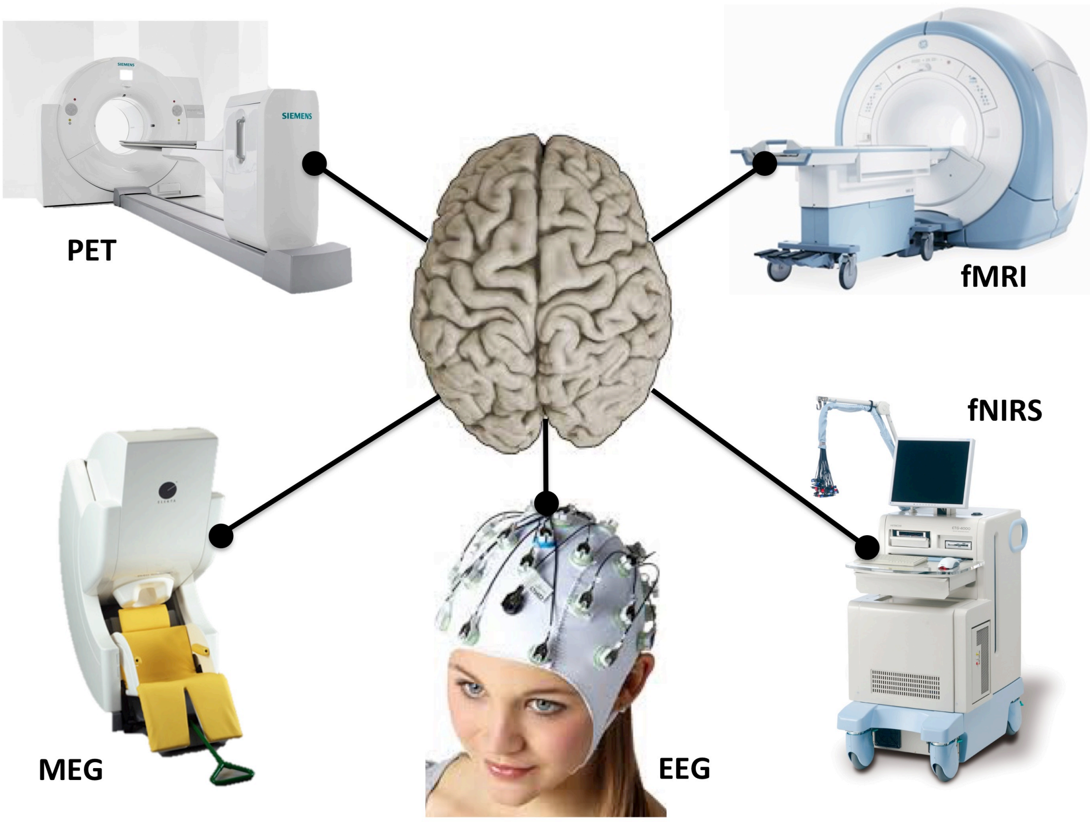

Figure 4.2. Range of functional neuroimaging techniques available to interrogate brain behaviour are illustrated.

Features of the ideal neuroimaging modality would possess portability, practicality, non-invasiveness, cost-effectiveness, resistance to motion-artefacts, and excellent temporal and spatial resolution. Temporal and spatial resolution can be simply described as the corresponding ability to accurately measure the time and location of brain activity. Currently available devices suffer from a trade-off between temporal and spatial resolution[176]. For example direct neuroimaging modalities such as Electroencephalography (EEG) and magnetoencephalography (MEG) provide good temporal resolution but poor spatial resolution. Conversely indirect neuroimaging modalities namely functional magnetic resonance imaging (fMRI), positron emission tomography (PET) and single photon emission computed tomography (SPECT) provide better spatial resolution but poor temporal resolution. Lastly functional near infra-red spectroscopy (fNIRS) an indirect neuroimaging modality may provide the best compromise in terms of moderate temporal and spatial resolution, yet is unable to scrutinize deeper cortical regions[248]. Therefore, selection of the optimal neuroimaging modality is reliant on the surgical task of interest and brain regions under interrogation.

Functional brain behavior can be characterized by activity in distinct cortical regions 'functional segregation' or by the interaction of these distinct regions 'functional integration' [249,250] (see Figure 4.3). Harder to assess, integration can be measured by '*functional connectivity'*, defined as statistical dependencies among remote neurophysiological events [28,209].

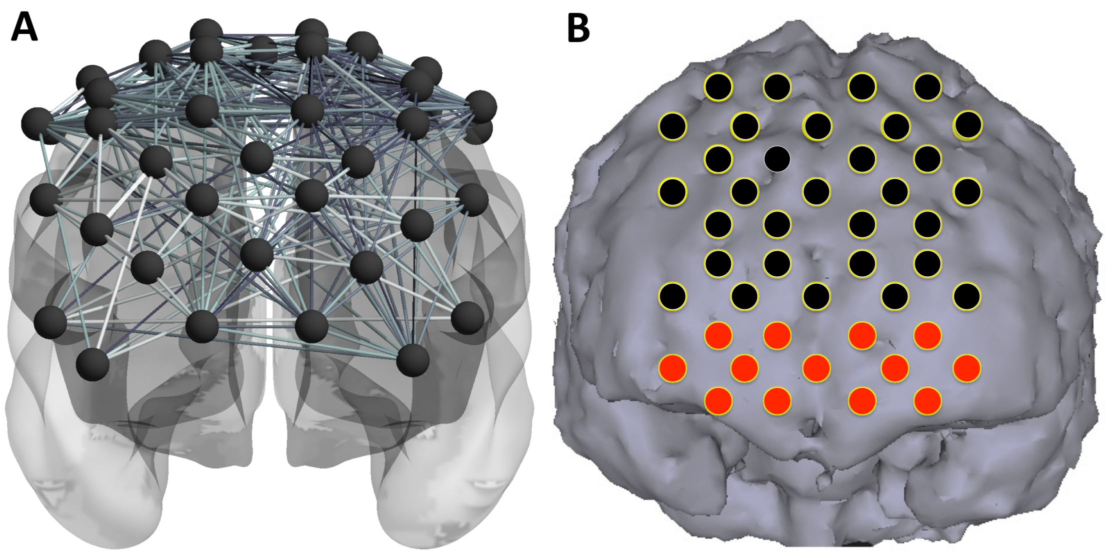

Figure 4.3. Functional brain behavior can be characterized by either (A) interaction of distinct regions 'functional integration' or by (B) activity in distinct cortical regions 'functional segregation'. Active regions are depicted in red.

The aim of this systematic review is to synthesize current literature that employs functional neuroimaging techniques in assessment of various factors that underlie operator performance within the field of surgery. The review incorporates three sub-domains that include exploration into [1] assessment of technical skills [2] acquisition of technical skills and [3] non-technical factors that impact performance.

4.2. Methods

4.2.1. Search methodology

A systematic literature according to PRISMA guidelines[251] was conducted for all articles published prior to December 2015. The search strategy included the following electronic database namely (1) PubMed, (2) EMBASE and (3) OVID MEDLINE. In addition further articles were identified from reference lists of selected articles. A flow chart (see Figure 4.4) displays the sequence of the search strategy. Keywords employed in the search strategy were: ("training" OR "assessment" OR "learning" OR "skills acquisition") AND "surgery" AND ("brain imaging" OR "neuroimaging"). Each article was screened for first author, year of publication, origin, study design, employed neuroimaging modality, brain region interrogated and characteristics of study subjects. In accordance with the guidelines set by the Oxford Centre for Evidence based Medicine [252], the levels of evidence were assigned to each article included in the review.

4.2.2. Inclusion and exclusion criteria

All published experimental or original studies, which utilised functional neuroimaging techniques in interrogation of learning, training methods, assessment or performance of a surgical task were included. Restrictions were not imposed by language. Identified articles were screened for duplication of data, by screening the authors and institution. In certain cases when duplication of data was indeterminable, they were not excluded. Response letters, hypothetical and review letters were excluded.

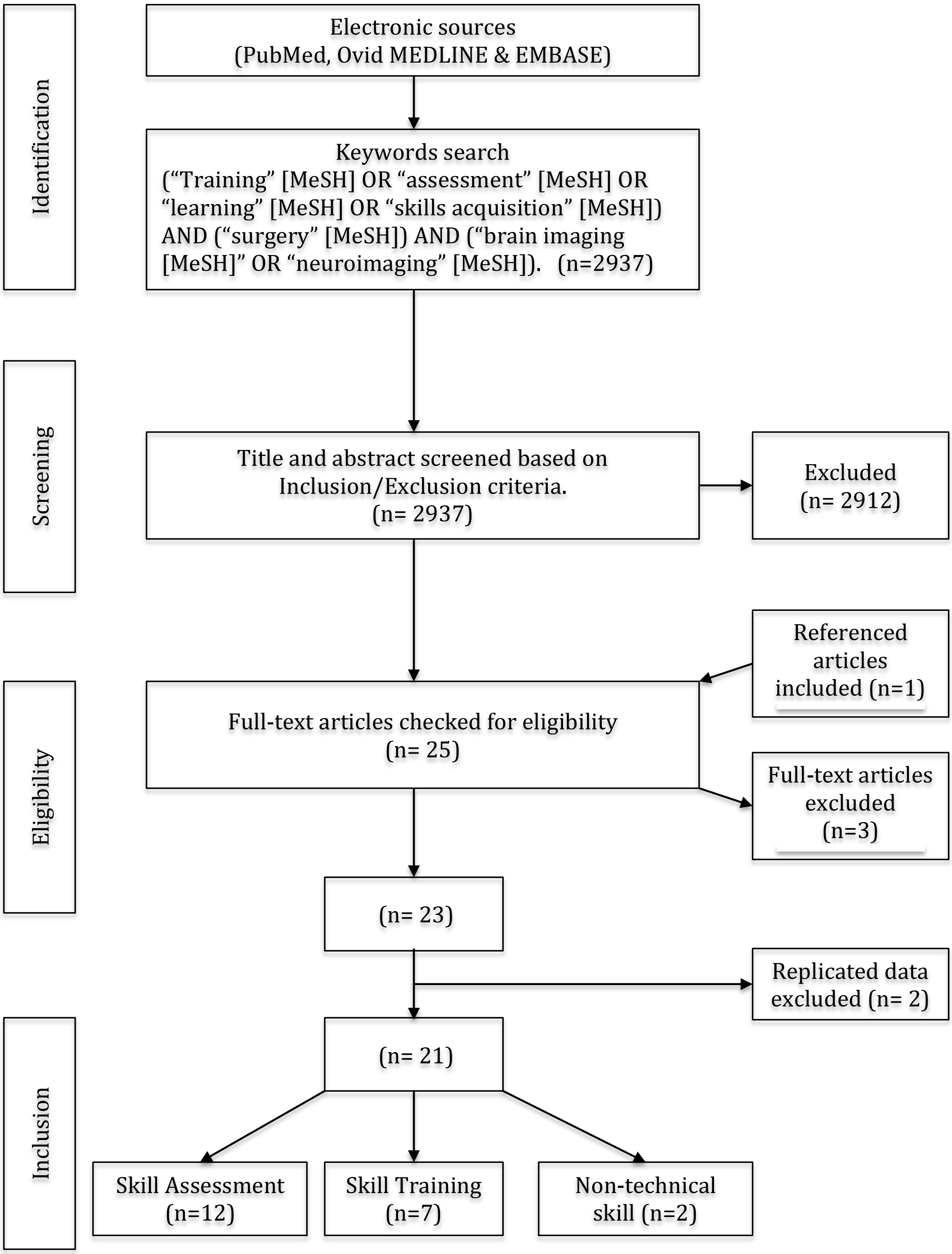

Figure 4.4. PRISMA flow chart illustrating the systematic search methodology.

4.3. Results

A total of 2937 results were identified using the employed search strategy, which was reduced to 25 by the initial screening process (see Figure 4.4). On review of full-text articles three articles were excluded as they lacked original data and an original study identified by hand-searching the reference list of the included full-text articles was added to the review. A further two full-text articles were then excluded due to duplication of original data. A total of 21 eligible articles were included and divided into three separate domains namely utilisation of functional neuroimaging to (1) assess surgical expertise or technology, (2) track learning and variation in training methods, and (3) assess non-technical operative skills such as decision-making, vigilance and mentoring. Three articles were identified as randomised control trials[253-255], one as an exploratory case series and the remaining 17 observational studies were either cross-sectional, cohort or case-control. A total of 388 subjects participated across all included reports resulting to an average (SD) of 18.47 (16.49) subjects per study. A wide variety of non-invasive functional neuroimaging techniques were employed to investigate brain behaviour. Of the reviewed literature fNIRS (n=11, 52%) was found to be the most frequently utilised neuroimaging modality, whilst the remaining utilised EEG (n=5, 24%), fMRI (n=4, 19%), and PET (n=1, 5%) (see Figure 4.5).

Investigation of brain behaviour within the field of surgery to assess expertise, track learning, and evaluate the impact of technology and training methods remain parsimonious ever since the first report in 2007 by Wanzel et al[256]. The number of published articles remained relatively constant ranging from one to four per year. Majority of the published papers by volume were from the United Kingdom (n=10, 48%)[253-255,257-263] whilst the remaining half hailed from Canada (n=3, 14%)[256,264,265], United States (n=3, 14%)[266-268], Japan (n=2, 9%)[29,269], Italy (n=1, 5%)[270], Ireland (n=1, 5%)[271], and China (n=1, 5%)[272].

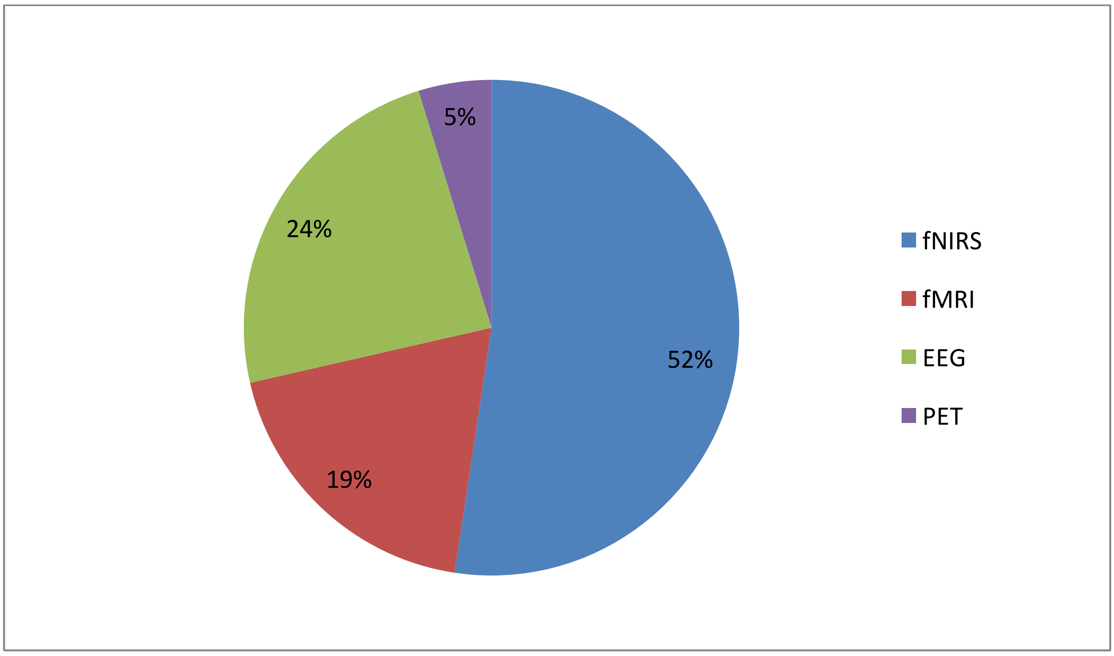

Figure 4.5. Pie chart representation in percentages of the neuroimaging modality employed across all articles included in the review.

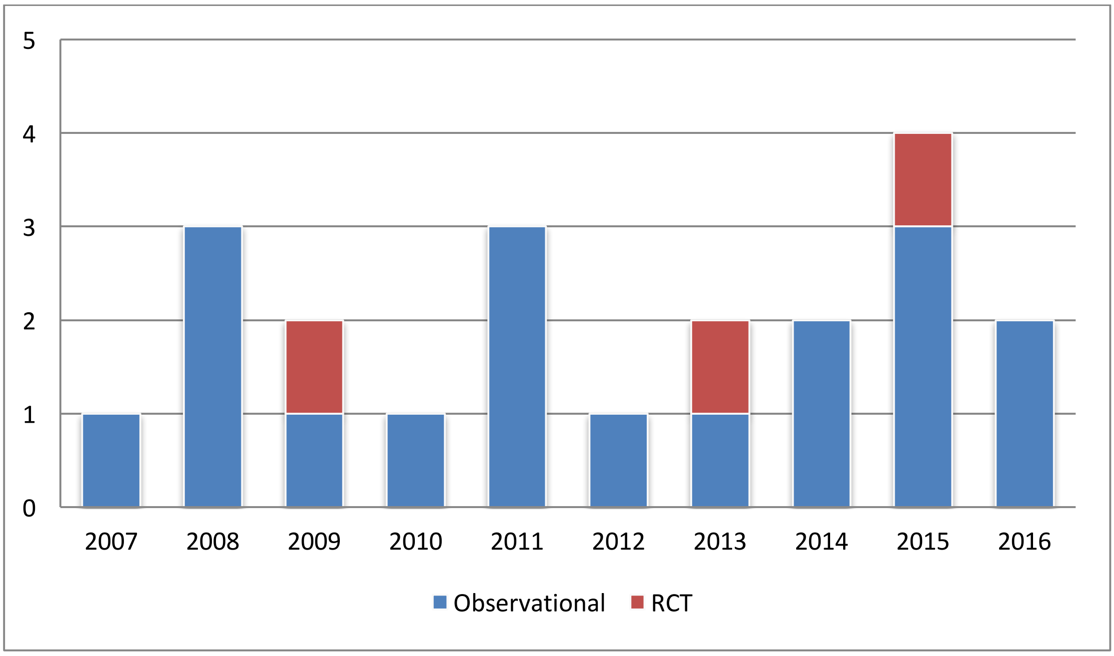

Figure 4.6. Bar chart illustrates number and type of published reports included in this review on an annual basis over the last 10 years.

A wide variety of surgical technical tasks were selected, requiring usage of one or more surgical platforms (see Table 4.1.). Predominantly MIS skills were employed (n=14, 67%)[29,253,254,257,258,262,264-270,272] over traditional open surgical skills (n=5, 28%)[10,255,256,259-261,271], whilst a single study solely examined non-technical skills[263]. Amongst MIS platforms the choice of tasks requiring laparoscopy (n=9) was most popular [29,257,262,264-266,269,270,272]. The remaining required performance of surgical tasks using daVinci robotic surgery (n=3)[267,268,270], virtual robotic surgery (n=2)[253,254] and natural orifice transluminal endoscopic surgery (n=1)[258].

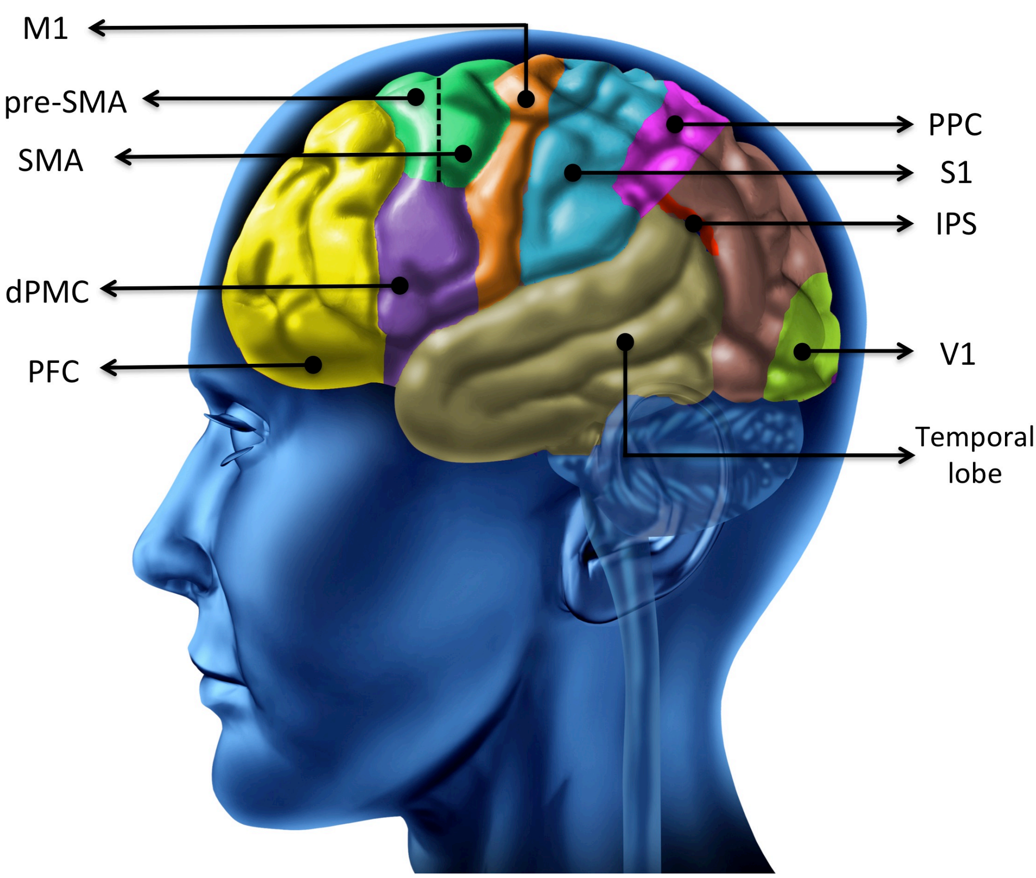

Figure 4.7. Brain regions interrogated by functional neuroimaging are labelled. The prefrontal cortex (PFC), dorsal premotor cortex (dPMC), pre supplementary motor area (pre-SMA), supplementary motor area (SMA), primary motor cortex (M1) are located in the frontal lobe. The posterior parietal cortex (PPC) and somatosensory cortex (S1) are located within the parietal cortex, and the visual cortex (V1) is located in the occipital lobe.

A wide variety of brain regions were investigated ranging from the whole brain to focussed functional brain regions (see figure 4.7; table 4.1). Five out of twenty one (24%) studies, which employed either fMRI or PET investigated whole brain activity[256,264-266,271], while the remaining sixteen of twenty one studies (76%) interrogated regional brain activity[29,253-255,257-263,267-270,272].

4.3.1. Assessment Tool

Functional neuroimaging was utilised as an assessment tool in a total of thirteen studies[29,256,258-261,264-267,269-271]. Its application was towards classification of expertise (n=7)[29,258,259,266,267,271], identification of fatigue (n=1)[261], identification of visuo-spatial ability (n=1) [256] and the effect of task complexity (n=1). Brain behaviour was also assessed to evaluate the effect of different platforms (robotic *vs* laparoscopic surgery)[270] and different angles of visualisation[269] on task performance.

4.3.1.1. Assessment of surgical expertise

Scrutiny of brain behaviour to improve the identification of surgical expertise was investigated in various forms of surgery. Three studies investigated the impact of expertise on brain function during an open surgical skill (hand-knot tying)[259,260,271]. Two studies investigated expertise during performance of laparoscopic skills, which varied from simpler peg transfer tasks [266] to complex intra-corporeal knot tying[29]. One study investigated expertise in multiple tasks using robotic surgery[267] and another in navigation via NOTES[258].

Morris et al[271] assessed competency in an open hand knot-tying task using fMRI to investigate for differential brain activation. Participants (n=9) evenly distributed into three groups of novices (n=3), intermediates (n=3) and experts (n=3) performed a block-designed experiment. The block design comprised of rest periods (motor and cognitive inactivity), surgical task periods and periods when participants were required to imagine executing the task. Two independent assessors, evaluated technical performance on quantitative (number of knots) and qualitative metrics (quality of the knot). In accordance with the level of expertise, experts outperformed intermediates who in turn outperformed novices (p=0.0147), yet this lacked statistical significance on post-hoc paired comparison tests. Experts displayed lower M1 activation (p<0.001) during the surgical task but greater V1 (p<0.05) activation than novices during task imagination. However no demonstrable differences in cortical behaviour were observed between experts and intermediates, perhaps due to the limited number of recruited participants in each group (n=3) or the selection of the task. Limited recruitment in this study could be attributed to the costs involved in usage of fMRI. Intermediates recruited may have achieved proficiency in performance of the open hand knot tying task, thereby displaying similarity in cortical behaviour to experts, which otherwise would have been detectable if the task was harder. Most importantly, the authors do not provide neuroscientific explanations for the differences in expert and novice cortical behaviour. Increased M1 activation in novices could be attributed to increased force of movements generated[273]. However movements involved in task execution were not controlled as participants were asked to perform as many knots as possible within a fixed time period. As a result experts performed almost double the number of knots than novices, which has an impact on M1 activation.

Similarly Leff et al[259,260] in two experimental studies utilised fNIRS to investigate expertise in surgical hand knot tying. Instead of the whole brain being monitored, the focus of investigation was the PFC utilising a 24-channel fNIRS system. In the first experiment[259] a total of 62 right-handed male participants comprised of surgical consultants (n=19), surgical registrars (n=21) and medical students (n=22) participated in a block designed experiment were required to execute five trials of the task which was interspersed with rest periods, i.e. a period of motor inactivity. Technical performance was objectively measured by time taken, path length and number of movements by utilisation of a validated motion-tracking device termed Imperial College Surgical Assessment Device (ICSAD)[274]. Unsurprisingly, consultants and registrars displayed stable technical performance, out-performing novices (p<0.001). Novices recorded within trial improvement from second to third trial (p<0.01). Technical performance on this basic surgical task did not differ significantly between registrars and consultants. PFC brain activation was measured by changes in cortical haemodynamics inferred by a rise in oxyhaemoglobin (HbO_2) and a drop in deoxyhaemoglobin (HHb). Akin to their technical performance both consultants and registrars displayed stability in PFC haemodynamics. Novices displayed initial rise in cortical activation (rise in HbO_2, with concurrent drop in HHb; $p<0.05$) followed by attenuation ($p<0.05$), which was linked to improvements in performance. The novice cohort (n=22) subsequently received 5 sessions of intensive training, which lasted 40 minutes each over the course of a week. The aim of this investigation[260] was to explore training induced longitudinal changes in PFC activity towards the direction of more experienced participants (surgical registrars and consultants]. To facilitate exploration, high dimensional fNIRS data derived from multiple channels and multiple haemoglobin species was embedded using a non-linear dimensionality reduction technique known as 'Isomap'[275] and projected into a lower dimensional topographical space. Differences in cortical excitation were calculable by earth mover's distance analysis which quantifies the dissimilarity between two multi-dimensional distributions[276]. Technical performance was measured by the aforementioned motion-tracking technique utilising ICSAD[112]. The performance of novices improved significantly ($p<0.001$) with training. Embedded data revealed greater clustering commensurate with expertise, i.e. experts and registrars clustered tightly whilst novices showed greater dispersion. However, on receipt of training novices displayed lesser dispersion of cortical activity indexed by progression toward the origin of the embedding, a location densely populated by expert data.

In both these studies[259,260] the PFC a brain region that is responsible for attention, concentration and performance monitoring was interrogated. According to Fitts' model of learning[42], during early stages of motor skills learning, task performance is reliant on attention and working memory in order to select the appropriate response to sensory feedback received, often learning by response to errors[235,277,278]. The PFC at the early stage of learning serves as a "scaffold" to cope with novel task demands, which gives way on well-honed practice[279]. Although novices displayed greater PFC activity than experts and on receipt of training displayed progression towards experts PFC behaviour, demonstrable significant differences between experts and intermediates were not observed, perhaps due to the selection of the motor task. Intermediates recruited to this study may have acquired the ability to proficiently perform open hand knot tying, Selection of a harder motor task may have allowed identification of observable differences in PFC behaviour between experts and intermediates.

Ohuchida et al[29] similarly explored expertise-based neuroplasticity in frontal lobe activity utilising a 22-channel fNIRS system. Instead of simple hand knot tying, a more complex laparoscopic knot-tying paradigm was chosen. Twenty-one participants comprised an unevenly distributed number of experts (n=4), trainees (n=4) and novices (n=13) performed laparoscopic intra-corporeal knot tying within a block-designed experiment. Technical performance analysis was based on the number of knots tied and demonstrated that experts successfully executed more than trainees who in turn executed more than novices (experts >10, trainees >4 and novices <3), yet no statistical comparisons were made. Frontal lobe activation measured by changes in haemoglobin species revealed that experts and all but one novice recorded insignificant activation whilst trainees and one novice did ($p<0.05$). A subset of novices (n=8) further sub-classified into a group with no prior experience (n=4) and another with minimal exposure (n=4), were provided an additional 2 hours of training and then re-assessed with fNIRS. Novices in the former group recorded increased cortical activation ($p<0.05$) whilst novices in the latter group showed reduction in frontal lobe activation ($p<0.05$).

The results of this study[29] suggests that experts and intermediates displayed PFC activity commensurate with motor learning literature, yet novices with no experience at all displayed akin to experts minimal PFC activity and only after receiving additional practice began to display significant PFC activity. This finding of the lack of PFC activity in the very early naive phase warrants further scrutiny, which may suggest a lack of engagement in the task. However, there are several methodological flaws in the design of this study. Subjects were unevenly distributed, the task was performed under temporal demand and most importantly does not provide adequate spatial information on which specific brain regions within the frontal lobe were evaluated[29]. The frontal lobe contains multiple motor regions associated with motor learning and activity in each region is distinctly different based on stage of skill acquired due to their differing functions.

A further study examined whole brain function during laparoscopy[266], investigated correlations between functional brain activation and the degree of laparoscopic experience. However in this study, differential brain activation was captured between a laparoscopic motor and visualisation task. Participants (n=10) comprised of novices (n=5) and experts (n=5), were required to perform a laparoscopic peg transfer task, visualise the same task but with no motor activity and also visualise a segment of a laparoscopic nephrectomy. Both novices and experts were provided with two sessions lasting fifteen minutes each and assessed on each occasion. Technical performance was assessed on time and accuracy based on the MISTELS curricula, a validated method for skills assessment[100]. Brain activation was investigated by PET, which required participants to undertake all tasks in an unnatural lying down position (see Figure 4.8). Expectedly, experts technically out performed novices ($p=0.001$). Both groups improved with practice but novices improved significantly ($p=0.006$) after the first practice session. During execution of the laparoscopic task experts only activated M1 ($p<0.001$) and novices showed greater global activation of the following regions (Lt precentral gyrus, Lt insula, Lt middle frontal gyrus, Rt precuneus and Rt inferior occipital gyrus) ($p<0.001$). Visualisation of laparoscopic surgery evoked greater activation of the occipital region ($p<0.001$) in novices and posterior cerebellum ($p<0.001$) in experts. The results of this study suggest that for execution of a simple laparoscopic task, novices required greater neuronal resources whilst experts only activate the M1 an area responsible for generation of motor execution. However the experimental design is limited by the lack of recruits at an intermediate stage, which in the studies described so far[29,259,260,271] have been difficult to differentiate from experts. Although PET scan provides the potential to examine whole brain activity, the administration of radio-active isotopes may deter greater participation as observed by the limited number of participants in each group (n=5). Lastly PET scan alters the natural environment in which MIS is performed, as it requires participants to perform the task in a lying down posture (see Figure 4.8).

Figure 4.8. Participants performing laparoscopic surgical tasks whilst undergoing functional neuroimaging using A) PET and B) fMRI (Pictures obtained with permission from Duty et al[266] and Bahrami et al[264]).

The impact of NOTES, an emergent surgical technology on differential brain activation was investigated by James et al[258]. Participants (n=29; all right-handed) of varying endoscopic experience (novices=18 & experts=11) were required to perform a simulated NOTES navigation task and 3 independent experts, assessed performance based on time and accuracy. Brain activation within the PFC measured by a 24-channel fNIRS system was defined as a significant rise in HbO_2 coupled to a concurrent significant decrease in HHb. Stress was measured too both physiologically by heart rate, salivary cortisol levels and subjectively by STAI. Participants were provided five minutes for familiarisation prior to task execution in a block-designed experiment of fixed task and rest periods. Experts out performed novices ($p<0.05$) based on good inter-rater reliability (Cronbach's alpha 0.997). Stress measures showed no difference induced by task activity. Greatest task-evoked brain activity was observed in the lateral PFC where experts significantly activated more channels than novices (six channels *vs* three channels). This may be explained by the fact that despite being expert endoscopists, they too were inexperienced in the evaluated task, which required navigation with a flexible scope in the intra-abdominal cavity as opposed to traditional intra-colonic navigation. Secondly, the PFC plays more than one functional role where it is also known to be associated with visuo-spatial working memory[280].

Guru et al[267] investigated expertise-based differences in brain activity for three different daVinci robotic surgical tasks. They were classified based on task complexity into basic skills (ball placement, suture pass and ring peg transfer task), intermediate skills (suturing and knot-tying) and an advanced skill (urethro-vesical anastomosis). Technical performance was measured by a validated video assessment method, which incorporated 6 metrics (i.e. overall time, number of times camera engaged, instruments collided, objects dropped and instruments out of view). Brain activity in the frontal, parietal and occipital lobe were measured by a nine channels EEG system, which estimated levels of cognitive load, mental state and cognitive engagement. Cognitive engagement was further sub-divided into high and low-level engagement, the former suggesting increased attention required and the latter disengagement. For basic skill technical performance, experts and intermediates outperformed novices on all technical metrics ($p<0.013$) however experts outperformed intermediates only on time ($p=0.025$). EEG recordings during basic skill performance displayed higher level of mental state and high-level cognitive engagement in novices than experts ($p<0.001$) and intermediates ($p<0.017$). Intermediates recorded greater degree of cognitive load ($p=0.021$), mental state ($p=0.013$) and high-level engagement ($p<0.001$) in comparison to experts. During performance of intermediate skills, novices recorded a greater number of instrument collisions than experts ($p=0.018$) and intermediates ($p=0.028$). Novices also required longer than experts to complete tasks ($p=0.006$). No significant differences were observed between intermediates and experts. However, on EEG assessment metrics, experts displayed lower mental state, cognitive load and high-level cognitive engagement compared to novices ($p<0.034$) and intermediates ($p<0.028$). Intermediates recorded lower in only high-level engagement ($p=0.018$) than novices. Only experts and intermediates performed the advanced skill task where experts were observed to be faster ($p=0.009$). Intermediates recorded greater cognitive load, mental state and high-level cognitive engagement ($p<0.025$) but lesser low-level cognitive engagement ($p=0.002$) than experts.

Robotic surgery devices are large and require surgeons’ to be seated on a console and their heads docked in a certain position to engage the device. This study highlights the feasibility of cognitive monitoring via EEG during the performance of robotic surgical tasks. However the shortcomings of this study are the small number of unmatched participants in each expertise-group. More pertinently, cognitive metrics are poorly defined and there appears to be considerable overlap between them as a result it is challenging to infer meaningful neuroscientific interpretations from the results.

4.3.1.2. Assessment of other factors on surgeon's brain behaviour

The impact of other factors namely surgical technology, the complexity of the surgical task, fatigue and visuo-spatial ability assessed by functional neuroimaging are described in this section. Baharami et al[264] explored variances in spatial brain activation of the whole brain evoked by complexity of laparoscopic task, albeit in novice subjects only. Laparoscopic tasks in order of complexity were the following: (1) pointing task, (2) right hand peg transfer, (3) left hand peg transfer, (4) bimanual peg transfer and (5) knot-tying. Novice subjects (n=9) practiced all five simulated laparoscopic tasks for ten sessions, each practice session lasting thirty minutes and conducted thrice a week. At the end of training participants performed these tasks whilst cognitively monitored using a 3T fMRI. A non-ferromagnetic box trainer and surgical instruments were used to allow usage with fMRI (see Figure 4.8). Technical performance was measured according to the time taken and accuracy as defined in the validated FLS curricula[104]. Participants expectedly performed better in simpler unimanual laparoscopic tasks, yet no comparative tests were reported. Greater spatial brain activation was observed during execution of complex tasks such as bimanual peg transfer (Lt PM, SMA and SPL) and knot-tying (bilateral PM, SMA, SPL, IPL, MOG, MTG & Lingual gyrus), especially in the SMA than in the easier unimanual pointing task (Lt PM, SMA, M1 and S1). During unimanual task performance, greater activation was recorded when executed right–handed (Lt PM, SMA, M1, S1, IPL, MOG, MTG) compared to when executed left-handed (Rt PM, SMA and SPL).

fMRI, unlike EEG and fNIRS can interrogate deeper brain structures that are involved in motor skill acquisition. Baharami et al[264] demonstrated the feasibility of fMRI to interrogate brain function whilst surgeons performed MIS tasks. Wider recruitment of sensory and motor brain regions underlies performance of complex laparoscopic tasks that necessitates bimanual co-ordination, especially in the SMA. These finding are in accordance with bimanual motor learning literature[25,281]. Identification of neural correlates for bi-manually coordinated surgical skills is valuable, as poor coordination is a prominent feature for trainee's struggles with laparoscopy[282], which has prompted interventions towards engagement of the non-dominant hand[283]. However, the significance of this study[264] is limited by virtue that experts were not studied and fMRI severely distorts the experimental environment by requiring participants to lie in unnatural positions and does not allow usage of ferromagnetic devices (see figure 4.8).

In order to understand the underlying reason for relative ease in robot-assisted bimanual task performance, Bocci et al[270] investigated brain dynamics in expert surgeons comparing robot-assisted versus laparoscopic suturing using brain connectivity analysis. Technical performance was measured by time taken for completion. A 32-channel EEG system measured electrical activity in various regions within the frontal lobe (pre-SMA, SMA & M1), parietal lobe (S1) and occipital lobe (V1). Functional brain connectivity between interrogated brain regions of interest was inferred by coherence within each brain hemisphere (intra-hemispheric) and between brain hemispheres (inter-hemispheric). Intra-hemispheric coherence was measured for the following: (1) M1-S1, (2) M1-SMA, (3) M1-pre-SMA, (4) S1-SMA and (5) S1-pre-SMA. Inter-hemispheric coherence was measured for following: (1) Lt M1-Rt M1, (2) Lt S1-Rt S1, (3) Lt M1-Rt pre-SMA & (4) Rt M1–Lt pre-SMA.

Sixteen right-handed subjects with an average experience of more than seven years of experience in each robotic and laparoscopic surgery performed the suturing task using both platforms in a block-designed experiment. Shorter completion times ($p<0.001$) with robotic assistance are in accordance with current literature[284-287]. Greater connectivity between intra-hemispheric motor (pre-SMA, SMA and M1) and non-motor sensory (S1) regions were observed during laparoscopic performance, whilst greater inter-hemispheric coherence was observed under robot-assistance[270]. These findings are commensurate with studies that have identified extensive inter-hemispheric connections via the corpus callosum are necessary for inter-manual co-ordination[288,289]. However, it is challenging to make meaningful interpretations of this study[270] as participants seemed to have greater laparoscopic than robotic experience. Secondly, a longitudinal study design would have been useful as inter-hemispheric connectivity using a similar investigative technique may have displayed initial rises in connectivity was followed by a subsequent drop[290]. Thirdly, SMA activity for bimanually co-ordinated tasks vary, dependent on discrepancies between movement directionality and magnitude referred to as 'interference'[281]. For example, anti-phase bimanual movements i.e. movements in opposite directions are harder to perform and are observed to have increased SMA activity[291].

A study by Miura et al[269] examined the link between optical axis-to-target view angle (OATVA) in MIS and loco-regional brain activation within the intraparietal sulcus (IPS). Activation in the IPS signifies embodiment[292] a degree of similarity between the body's natural movement and in usage of instruments. A 24-channel fNIRS device measured brain activation whilst predominantly right-handed male novice participants performed the initial aspect of intra-corporeal suturing i.e. inserting the needle to the intended target. Each participant performed five trials of the described task under variable OATVA conditions (15°, 30°, 45°, 60°, 75° and 90°). Performance was not assessed but the majority of participants (4 out of 5) displayed greatest IPS activation when OATVA was 75°. This neuroergonomic study limited by the number of participants and their lack of expertise in MIS helps to determine the most appropriate angles between the endoscopic camera and the instruments.

Evaluation of fatigue induced by night-shift work patterns in a widely researched topic[293-295]. Thus characterisation of brain behaviour to enable evaluation of fatigue over the course of a night shift has value in trying to minimize fatigue-induced errors. Functional neuroimaging has been used to evaluate the impact of fatigue over the course of a night shift on motor and cognitive skills[261]. Surgical registrars (n=7) between 10 PM at night until 8 AM in the morning were serially assessed on a two hourly basis (six sessions) in the performance of a motor (hand-knot tying task) and cognitive task (arithmetic). Brain behaviour in the prefrontal cortex was investigated via a 24-channel fNIRS system measured changes in cortical haemodynamics. Technical performance was measured objectively by motion metrics (ICSAD) and the cognitive task was measured by accuracy of response. Fatigue was also measured introspectively using the Epworth Sleepiness Scale. According to the Epworth Sleep Scale, participants rapidly fatigued between 10PM until 2AM ($p<0.03$), after which no further decline was observed. Significant variance was recorded in technical performance ($p<0.05$) where initial improvement was followed by decline and subsequent stability. Non-significant variation was observed in performance of the arithmetic task. Introspective fatigue corresponded to increases in PFC response when required to perform a cognitive task. The cognitive task evoked significantly greater PFC excitation than the surgical task on HbO_2 and Total haemoglobin (THb) analysis ($p<0.001$), which was presumably because the technical task was better honed. Although it can be inferred that cognitive tasks requires greater attentional resources especially when fatigued, the cognitive task chosen tested arithmetic ability and therefore did not represent a true clinical decision-making task.

Visuo-spatial ability is an important pre-requisite for surgical ability and may have potential value in selection for surgical residency. To establish a link between pattern of cortical activation and visuo-spatial ability in surgeons (n=18), Wanzel et al[256] recorded whole brain activity via a 1.5T MRI system. Visuo-spatial ability was assessed by a block-designed experiment where during task periods, participants were required to perform a mental rotation task (MRT)[296] which comprised of determining whether quasi-3D figures rotated on one or two axes were identical (see Figure 4.9). A previous study by the same group had demonstrated correlation between visuo-spatial ability assessed by MRT and technical performance of a surgical task in this case a Z-plasty procedure[297]. MRT performance assessed by speed and accuracy of response observed inter-subject variability (Score range 17 to 42 out of 50). Greater activation in relation to superior MRT performance was observed in left hemispheric regions of middle temporal, posterior cingulate-precuneus and PM regions ($p<0.01$)[256].

However the findings of this study are severely limited by firstly the lack of correlation between mental rotation task performance by the same participants in the first and second study. Secondly, Z-plasty performance was not re-evaluated in the subsequent study probably due to the incompatibility of fMRI with ferro-magnetic surgical devices making it difficult to infer a relationship with actual surgical ability[298].

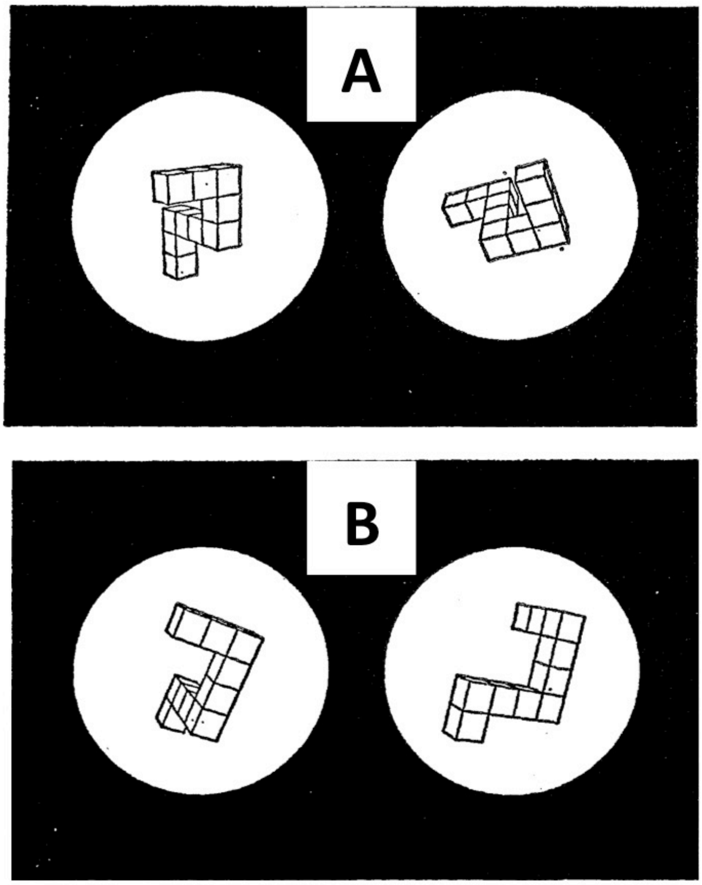

Figure 4.9. A pair of quasi-3D drawing which were (A) identical on mental rotation (B) non-identical despite rotation (Images adapted from Shepard et al [296]).

4.3.2. Brain behavioural assessments as a tool for tracking learning

Eight studies included track learning related changes in brain behaviour by various functional neuroimaging techniques[253-255,257,259,260,266,272]. Of these, three[29,260,266] were discussed in the earlier section (4.3.1.1.) as they each contained multiple experiments comprised of both cross-sectional and longitudinal design. In addition another paper evaluated skills acquisition based on type of neurofeedback provided[255].

4.3.2.1. Tracking motor learning

Neuroplasticity associated with laparoscopy was initially investigated by Leff et al[262]. A 24-channel fNIRS system interrogated cortical activity in the left PFC and right parietal cortex. Novice participants performed a laparoscopic tracking task under different visual feedback conditions (rotated and normal laparoscopic camera views). Cortical activation in aforementioned regions was derived by changes in HbO_2 and HHb. High dimensional fNIRS data was dimensionally reduced by isomapping into a geospace. A Markov chain was then used to detect temporal flow in the created geo space to analyse fronto-parietal neuroplasticity. Markov chain suggests global convergence towards a region of minimal activation, where greatest convergence was seen in the final transition from counter clockwise to normal view conditions. Thus, practice of the laparoscopic navigation task led to attenuation of PFC activity signifying reduction in demands for attentional resources.

Similarly, neuroplasticity underlying complex skills acquisition was serially tracked by Crewther et al[257]. Participants (n=12) were trained to perform laparoscopic intra-corporeal suturing and knot tying over 7 distributed sessions across a fortnight. They were assessed at regular intervals after 2 hours (BASE), 5 hours (MID) and 8 hours (POST) of practice. All participants were also tested for retention four weeks from the last training session (RETEST). Motor performance was assessed by time and accuracy based on the FLS criteria[104]. Stress was measured perceptually by (STAI) and mental work-load (NASA-TLX), and physiologically by heart rate variability, and hormonal markers (salivary and cortisol levels). PFC activity was measured by changes in cortical haemodynamics via fNIRS. Technical performance improved incrementally from BASE to MID and MID to POST & RETEST ($p<0.001$). Perceptual stress (STAI) decreased from MID to POST ($p<0.001$) and workload (NASA-TLX) reduced between BASE & MID to POST & RETEST ($p<0.018$). Non-significant fluctuation in hormonal levels was observed. PFC activation remained stably elevated in all sessions ($p<0.035$) with no lateralisation despite associated performance improvements. These findings highlight that attenuation of PFC activity is task-specific, does not necessarily accompany improvement in performance and automaticity is slow to develop for complex tasks. It also highlights the importance of imaging more than one functional brain region to determine neuroplastic changes associated with skills acquisition.

4.3.2.2. Effect of different training methods on neuroplasticity

Three studies[253,254,272] evaluated the effect of different training regimes on technical surgical skills acquisition of which two were randomised control trials. All three studies explored brain behaviour by using brain connectivity analysis.

The first study by James et al[253] examined learning unassisted (control) to learning under gaze assisted motor channelling guidance (experimental). Participants (n=21) were required to track a moving target, a simulation of ablation on a beating heart. Both groups underwent six sessions of training over eight days and then tested for retention after two months. Participants in the experimental group learnt under conditions where hand movements were constrained to the focus (gaze) of interest to improve accuracy. Control group participants learnt the tracking task unaided. Motor task was assessed by accuracy and time. Cortical activity in the PFC and PPC was analysed by changes in HbO_2 and HHb. A fronto-parietal brain network created to measure network economy. Incremental improvement in technical performance was observed in both groups ($p<0.001$). Initial performance improvements from session 1 to 3 were greater in the control group. However from session 3 to 6, and on retention the experimental group performed better. The experimental group displayed rapider attenuation in cortical activation of PFC and PPC. Brain network analysis recorded a more economical cortical network in the experimental group.

These results suggest that learning under augmentation[253] or modulation[254] of feedback has shown to achieve superior task performance demonstrated by accelerated redistribution of brain activity posteriorly from PFC to the PPC and a more economical cortical network, commensurate to motor learning literature[25,246,299,300]. The PPC plays a role in translation between spatial and motor information during early stages of motor learning and at later stages is involved in storage of acquired skill[246]. In addition this study demonstrates the value of interrogating more than one functional brain region associated with motor skills acquisition.

Similarly, Leff et al [254] utilised a randomised crossover design to compare neuroplasticity underlying learning a collaborative simulated robotic surgical task. Participants (n=20) were required to take virtual biopsies from seven locations in randomised order under either verbal instruction (control) or visual guidance (experimental) from their trainer. Trainer's eye fixation was tracked and telecasted on the participant's screen to provide visual guidance. Motor task performance was assessed by number of biopsies taken and accuracy of movement (path length). Gaze behaviour of participants was tracked. Cortical activation in the occipital cortex was analysed by changes in HbO_2, HHb and THb. Occipito-parietal brain networks were created to measure network econometrics. Under visual guidance, participants in the experimental learning group performed better detected by the number of virtual biopsies taken ($p=0.003$) and accuracy ($p<0.001$). A more focussed gaze behaviour ($p<0.001$), and lesser occipital activation according to HbO_2 ($p<0.001$) and THb ($p<0.001$) analysis were also identified under visual compared to verbal guidance. No significant differences were observed in functional brain network econometrics.

Therefore, collaborative control via the latest dual console daVinci robotic systems has the potential to facilitate training, which otherwise suffers from loss of face to face-to-face communication because both trainee and trainer are perceptually docked onto spatially separate consoles. This finding has been similarly observed when a laparoscopic task learnt under visual guidance provided by the trainer's eye fixation was compared to expert verbal guidance[301]. Trainee gaze behaviour under visual guidance was also more focussed and resulted in attenuation in occipital lobe activation. Therefore, by monitoring eye and brain behaviour, the former study was able to provide a possible explanation for improved performance under visual guidance. The visual cortex (V1) located in the occipital lobe is involved in visual search behaviour. By casting gaze behaviour of the expert onto the trainee's screen, it streamlines visual search allowing location of the intended target by visual saliency. This visual strategy known as a "bottom-up approach" differs from the more effortful "top-down approach" which relies on additional attention and sensory regions namely the PFC and PPC[302]. In addition the finding of streamlined visual search behaviour results in decreased V1 activation commensurate with existing literature[303]. However, cortical network analysis did not display significant differences[254]. This can be explained by the limited coverage of fNIRS channels, which solely focussed on the occipital lobe. Redistribution of cortical activity and greater information on the underlying cortical network would have been feasible if additional regions such as the PFC and PPC implicated in top-down visual search strategies[302] were interrogated.

Zhu et al [272] investigated the impact of learning a laparoscopic task under different conditions on brain efficiency. Over eight training blocks, novice participants (n=18) learnt a bimanual laparoscopic tracking either under no feedback (implicit learning, n=9) or were provided in advance visually the path of the moving object that needed to be tracked (explicit learning, n=9). Participants from both explicit and implicit learning groups were tested for retention, whilst simultaneously cognitively monitored. A 7-channel EEG system measured brain connectivity by T3-Fz (verbal-analytic) and T4-Fz (visuospatial) coherence analysis. Technical proficiency was measured by accuracy of tracking task. Participants in the explicit group technically outperformed those in the implicit group, yet there was no difference at retention (p=0.231). Implicit group participants displayed significantly lower T3-Fz coherence (p=0.027) suggestive of lesser verbal analytic activity. The authors[272] suggest that the lesser verbal analytic activity to support learning in the implicit group infers greater neural efficiency. However information is not provided on whether the participants reach asymptotic levels of performance by the end of training and cognitive assessments was performed only once at the time of retention testing.

4.3.2.3. Effect of neurofeedback on skills acquisition

Self-regulation of brain activity to optimize surgical performance was investigated by Ros et al [255]. Neurofeedback via EEG was used to enable participants to self-regulate brain activity. Ophthalmic surgery trainees (n=20) were randomly allocated into either sensorimotor (SMR, n=10) or alpha-theta (AT, n=10) neurofeedback groups. A subset of participants (n=8) from each group (n=4) served as controls by being assessed three months in advance. Participants from both groups received eight sessions of feedback training lasting thirty minutes each over a period of two to three months. The microsurgical task to be assessed was divided into four sub-tasks namely (1) sideport incision, (2) phaco-wound, (3) capsulorrhexis, and (4) suture. Performance was quantitatively assessed but only using the time taken. Additionally, 2 consultant surgeons provided qualitative analysis by assessing intra-operative videos using a 2-point scale on 54 criteria. Anxiety during performance was measured using the Spielberger Anxiety Inventory. Participants in the SMR group displayed greater overall improvement by a reduction in time taken (p=0.019) and on qualitative analysis (p=0.028), especially during the most complex 'suturing' sub-task (p=0.018). SMR participants were also observed to display increased pause time between sub-tasks, yet it resulted in lower overall task time (R=-0.72; p=0.019). SMR feedback also resulted in decreased trait anxiety (p=0.017). The effects of neural feedback with the aim to promote a relaxed cerebral state conducive to motor learning is interesting and requires further scrutiny. However attendance of neural feed-back sessions on a bi-weekly basis was not strictly compliant due to the working schedules of recruited ophthalmic surgeons. Some participants in the SMR group had longer intervals between neural feedback sessions, which might have had an effect on their slower progression in performance of the surgical task. This highlights the challenges posed to adherence of protocol in longitudinal studies involving surgeons.

4.3.3. Cognitive assessment of intra-operative vigilance and decision-making

The aforementioned cognitive studies focus on interrogation of neural substrates underlying technical skill. Surgeon's cognitive processes intra-operatively are invested in other functions namely intra-operative decision-making and vigilance. For instance, an expert surgeon in a trainer's role may not be performing the technical task but will instead be cognitively invested on observing the task and mapping out intra-operative decisions. Instead of examining cognitive processes in the trainee during training, Hussein et al [268] explored the relationship between objective cognitive metrics of the trainer (EEG) to subjective cognitive metrics of the trainer and trainee. Brain behaviour of a single expert surgeon was assessed using a 24-channel wireless EEG device to measure cognitive processes whilst mentoring a trainee surgeon perform (n=20) urological procedures namely urethro-vesical anastomosis (UVA) and extended lymph node dissection (LND). In addition cognitive demands on the expert surgeon between observing and actual performance were compared. Cognitive processes measured by EEG were distraction, mental workload and mental state. Subjective cognitive assessment of trainee and trainer surgeon was measured by a validated assessment tool namely NASA-TLX. Whilst training LND procedure, trainer's EEG mental workload negatively correlated with his corresponding subjective mental demand (r=-0,74, p=0.05), effort (r=-0.86, p=0.01, r=-0.84, p=0.02) whilst no significant correlation was observed during UVA. No significant correlation was also observed between trainer's EEG and trainee's subjective cognitive load during LND. However, during UVA, trainer's distraction negatively correlated with trainee's subjective mental demand (r=-0.82, p<0.01), physical demand (r=-0.76, p=0.01) and temporal demand (r=-0.69, p=0.03). In addition trainer's mental state correlated with trainee's subjective mental demand (r=0.77, p=0.01), physical demand (r=0.68, p=0.03) and temporal demand (r=0.64, p=0.01). Finally no significant differences were observed when the trainer was observing or performing the task. These results are interesting but limited as only one expert surgeon was cognitively monitored. In addition the results lack an appropriate neuroscientific explanation, especially when mental workload measured on a subjective scale negatively correlated with the objective cognitive metric.

Intra-operative decision-making a key ability of a surgeon to make logical and safe judgements based on dynamic sensory information were assessed by Leff et al[263]. In this study, 22 participants of various grades ranging from novice to experts (novices=10, residents=7, experts=5) were assessed for decision-making quality, confidence and consistency. The task involved selection of the optimal next surgical manoeuvre on 12 videos of simulated laparoscopic cholecystectomies. Video clips were termed "primed" or “unprimed”, based on indication of the next surgical manoeuvre by the presence of a corresponding surgical instrument in the video clip. Cortical activation within the PFC was estimated by changes in HbO_2 and HHb. Unsurprisingly experts and residents were not only more confident than novices ($p<0.001$) but also more consistent (experts $p<0.001$, residents $p<0.05$ and novices $p=0.183$). During decision making for unprimed videos novices significantly activated more channels in the PFC (14/22 channels) than residents or experts (4/22 channels each), which was not the case during primed videos. These results suggest that novices under unfamiliar circumstances displayed a greater amount of dorsolateral PFC activation, an area found to play a critical role in complex value-based decision-making processes[304]. Unlike novices, more experienced participants (residents and experts) displayed a lack of significant activation in the PFC which was attributed to effortless habitual decision-making.

Author	Year	Origin	Modality	Type	Task	Subjects	Brain ROI
Duty[266]	2012	New York, US	PET	Cross-sectional & Cohort	Laparoscopic peg transfer task and visualisation of the same task and part of a laparoscopic nephrectomy.	n=10, (novices=5 & experts=5)	Whole brain
Zhu[272]	2011	Hong Kong, China	EEG, 7 channels	Cohort	Bimanual laparoscopic tracking task. Implicit group learnt it with no feedback. Explicit group were provided visually the path of the moving object that required to be tracked.	n=18, (all novices, explicit learning group=9 & implicit learning group=9)	Frontal and temporal lobe
Ros[255]	2009	London, UK	EEG	RCT	Microsurgical cataract surgery divided into sub-tasks namely (1) sideport incision, (2) phaco-wound, (3) capsulorrhexis and (4) suture.	n=20, (trainee ophthalmic surgeons)	Frontal and parietal lobe.
Guru[267]	2015	Buffalo, NY, USA	EEG 9 channels	Cross-sectional	Robotic surgery tasks (daVinci Surgical System). 3 levels of complexity (1) basic skill (ball placement, suture pass and ring peg transfer task), (2) intermediate (suturing and knot-tying) & (3) advanced (urethro-vesical anastomosis).	n=10, (novices=2, intermediates=5 & experts=3)	Frontal, parietal and occipital lobe.
Bocci [270]	2013	Pisa, Italy	EEG 32 channels	Cross-sectional	MIS intra-corporeal suturing and knot-tying via laparoscopic and robotic platforms (daVinci Surgical System).	n=16, (experienced laparoscopic and robotic surgery)	Frontal (pre-SMA, SMA & M1), parietal (S1) and occipital lobe (V1)
Hussein [268]	2016	Buffalo, NY, USA	EEG 20 channels	Exploratory	To measure cognitive metrics of a surgical trainer whilst training segments of robotic assisted urological procedures.	Single surgical trainer during n=20 urology procedures	Frontal, parietal, temporal and occipital lobe.
Bahrami [264]	2014	Toronto, Canada	fMRI 3T	Cross-sectional	Laparoscopic tasks in order of complexity are (1) pointing task, (2) right hand peg transfer, (3) left hand peg transfer, (4) bimanual peg transfer and (5) knot-tying.	n=9, (all novices, right handed=5 & male=5)	Whole brain
Bahrami [265]	2011	Toronto, Canada	fMRI 3T	Exploratory	Laparoscopic tasks in order of complexity are (1) pointing task, (2) right hand peg transfer, (3) left hand peg transfer, (4) bimanual peg transfer and (5) knot-tying.	n=2, (both right handed, female & novices)	Whole brain
Morris [271]	2015	Dublin, Ireland	fMRI 3T	Cross-sectional	Open hand knot tying task	n=9, (novices=3, intermediates=3 & experts=3; gender unspecified)	Whole brain
Wanzel [256]	2007	Toronto, Canada	fMRI 1.5T	Case-control	Mental rotation task (MRT)	n=18, (all trainees, male=14)	Whole brain
Ohuchida [29]	2009	Fukoaka, Japan	fNIRS 22 channels	Cross-sectional & Cohort	Laparoscopic knot tying	n=21, (experts=4, trainees=4 & novices=13; gender unspecified)	Frontal lobe

Author	Year	Origin	Modality	Type	Task	Subjects	Brain ROI
Leff [259]	2008	London, UK	fNIRS 24 channels	Cross-sectional & Cohort	Open hand knot-tying	n=62, (consultants=19, registrars=21 & students=22; all right handed males)	Frontal lobe (PFC)
Leff [262]	2008	London, UK	fNIRS 24 channels	Cohort	Laparoscopic tracking and grasping task under normal and rotated laparoscopic views.	n=14, (all right-handed, male novices)	Frontal lobe (Lt PFC) & Rt parietal cortex
Leff [259]	2008	London, UK	fNIRS 24 channels	Cross-sectional	Open hand knot-tying	n=62, (consultants=19, registrars=21 & students=22; all right handed males)	Frontal lobe (PFC)
Leff [261]	2010	London, UK	fNIRS 24 channels	Cohort	Open hand knot-tying (motor task) and cognitive task (arithmetic calculation)	n=7, (all right-handed, male surgical registrars)	Frontal lobe (PFC)
James [258]	2011	London, UK	fNIRS 24 channels	Cross-sectional	NOTES navigation task	n=29, (novices=18 & experts=11; all right handed males)	Frontal lobe (PFC)
James [253]	2013	London, UK	fNIRS 24 channels	RCT	Tracking a moving target to replicate simulated ablation on a beating heart	n=21, (control group=11 & experimental group=10, all right handed, males= 15)	Frontal lobe (Lt PFC) & parietal lobe (Rt PPC)
Leff [254]	2015	London, UK	fNIRS 24 channels	Randomised cross over trial	Simulated collaborative surgical task between trainee and trainer. Task created to replicate dual console *daVinci* robotic platform.	n=20, (all right-handed novices, male=19)	Occipital cortex
Crewther [257]	2015	London, UK	fNIRS 22 channels	Cohort	Laparoscopic intra-corporeal suturing and knot-tying	n=12, (all right-handed male novices)	Frontal lobe (PFC)
Miura [269]	2015	Tokyo, Japan	fNIRS 24 channels	Cross-sectional	Simulated virtual reality task of robotic suturing performed under variable optical-axis-to-target-view-angle (OATVA) conditions.	n=5, (all novices, right-handed=4, male=4)	Parietal lobe (IPS)
Leff [263]	2016	London, UK	fNIRS 24 channels	Cross-sectional	Decision making task during visualisation of a simulated laparoscopic cholecystectomy procedure.	n=22, (novices=10, residents=7, experts=5)	Frontal lobe (PFC)

Table 4.1. Summarizes the study demographics, country of origin, brain region investigated, type of functional neuroimaging modality, participant characteristics and surgical task assessed.

Author	Outcomes	Method	Other Results	Brain Results
Duty [266]	Differential brain activation based on expertise during laparoscopic motor and visualisation task. Technical proficiency assessed by time and accuracy.	Multiple PET scans were undertaken for each participant during (1) rest (no stimulus), (2) viewing a peg transfer task, (3) viewing part of a laparoscopic nephrectomy, during performance of a laparoscopic peg transfer task (4) with no prior practice, (5) after 15 minutes of practice and (6) after 30 minutes of practice.	Experts technically out performed novices (p=0.001). Both groups improved with practice but novices improved significantly (p=0.006) after the first practice session.	During execution of the laparoscopic task experts only activated M1 (p<0.001) and novices showed greater global activation of the following regions (Lt precentral gyrus, Lt insula, Lt middle frontal gyrus, Rt precuneus and Rt inferior occipital gyrus) (p<0.001). Visualisation of laparoscopic surgery evoked greater activation of the occipital region (p<0.001) in novices and posterior cerebellum (p<0.001) in experts.
Zhu [272]	EEG measured T3-Fz (verbal-analytic) and T4-Fz (visuospatial) coherence Technical proficiency was measured by accuracy of tracking task.	Performance was tracked across 8 training blocks of 3 trials each amounting to a total of 24 trials for each participant. Participants from both explicit and implicit learning groups were tested for retention, whilst simultaneously cognitively monitored.	Participants in the explicit group technically outperformed those in the implicit group, yet there was no difference at retention (p=0.231).	Implicit group participants displayed significantly lower T3-Fz coherence (p=0.027) a measure of underlying verbal-analytic activity, thereby suggesting more neurally efficient brain behaviour.
Ros [255]	Microsurgical performance assessed quantitatively by time taken. Additionally 2 consultant surgeons provided qualitative analysis by assessing intra-operative videos using a 2-point scale on 54 criteria. Anxiety was measured using the Spielberger Anxiety Inventory.	Ophthalmic surgery trainees (n=20) were randomly assigned into one of two neurofeedback groups: SMR (n=10) and AT (n=10). A subgroup of 4 participants in each group was tested 3 months in advance to serve as controls. Participants from both groups received over a period of 2-3 months, 8 × 30 minute sessions of feedback training.	Participants who received SMR feedback displayed greater overall improvement by a reduction in time taken (p=0.019) and qualitative analysis (p=0.028), especially during the most complex 'suturing' sub-task (p=0.018). SMR participants were observed to display increased pause time between sub-tasks, yet it resulted in lower overall task time (R=-0.72; p=0.019). SMR feedback also resulted in decreased trait anxiety (p=0.017)	Participants were cognitively monitored only during neurofeedback sessions and not during performance of the microsurgical task.
Guru [267]	Technical performance was measured on a validated video assessment method, which incorporated 6 metrics (overall time, number of times camera engaged, instruments collided, objects dropped and instruments out of field). EEG measured levels of cognitive load, mental state and engagement which was further sub-divided into high and low level engagement.	Novices performed only basic and intermediate skills, whilst intermediates and experts performed all three levels of skill. All participants were cognitively monitored during the surgical task assessment.	Basic skill: Experts and intermediates outperformed novices on all technical metrics (p<0.013), experts outperformed intermediates only on time (p=0.025). Intermediate skill: Novices recorded greater number of instrument collisions than experts (p=0.018) and intermediates (p=0.028), and also required longer than experts (p=0.006). No significant differences were observed between intermediates and experts. Advanced skill: Experts completed the task faster than intermediates (p=0.009).	Basic skill: Novices displayed higher level of mental state and high-level cognitive engagement than experts (p<0.001) and intermediates (p<0.017). Intermediates recorded greater degree of cognitive load (p=0.021), mental state ((p=0.013) and high-level engagement (p<0.001). in comparison to experts. Intermediate skill: Experts displayed lower mental state, cognitive load and high level cognitive engagement compared to novices (p<0.034) and intermediates (p<0.028). Intermediates compared to novices recorded lower in only high-level engagement (p=0.018). Advanced skill: Intermediates recorded greater cognitive load, mental state and high-level cognitive engagement (p<0.025) but lesser low-level cognitive engagement (p=0.002) than experts.

Author	Outcomes	Method	Other results	Brain results
Bocci [270]	Technical performance was measured by time required for task completion. Intra-hemispheric coherence was measured for the following: (1) M1-S1, (2) M1-SMA, (3) M1-pre-SMA, (4) S1-SMA & (5) S1-pre-SMA. Inter-hemispheric coherence was measured for (1) Lt M1-Rt M1, (2) Lt S1-Rt S1, (3) Lt M1-Rt pre-SMA & (4) Rt M1–Lt pre-SMA.	16 subjects performed the robotic and laparoscopic suturing task whilst EEG recorded brain activity in 5 regions of interest namely (1) M1, (2) S1, (3) V1, (4) SMA and (5) pre-SMA. Brain activity was also recorded during periods of motor inactivity (rest).	Time taken for robotic suturing was shorter time than laparoscopic suturing (p<0.001).	Greater intra-hemispheric coherence was observed during laparoscopy in comparison to robotic suturing for the following regions SMA-M1, SMA-S1, M1-S1 and S1-preSMA (p<0.001). Conversely, greater inter-hemispheric coherence was observed during robotic surgery for the following regions Lt M1-Rt M1, Rt-Lt S1, Rt M1–Lt pre-SMA and Lt M1–Rt pre-SMA (p<0.001)
Hussein [268]	EEG measured distraction, mental workload and mental state for the surgical trainer. NASA-TLX measured subjective cognitive load of both trainer and trainee. They composed of 6 sub-scales namely (mental demand, physical demand, temporal demand, frustration, effort and frustration)	Cognitive behaviour of a single expert surgeon was assessed objectively (EEG) and subjectively (NASA-TLX) during training of segments of 20 robotic assisted operations. Mentored operative segments included extended lymph node dissection (eLND) and urethra-vesical anastomosis (UVA). In addition cognitive behaviour of the surgeon was compared when observing and performing the aforementioned tasks.	Performance not assessed. Correlation of subjective cognitive load (NASA-TLX) with objective data (EEG) is reported in the next column.	During eLND mentors, trainer's EEG mental workload negatively correlated with his corresponding subjective mental demand (r=-0,74, p=0.05), effort (r=-0.86, p=0.01, r=-0.84, p=0.02). During UVA, trainer's distraction negatively correlated with trainee's subjective mental demand (r=-0.82, p<0.01), physical demand (r=-0.76, p=0.01) and temporal demand (r=-0.69, p=0.03). In addition trainer's mental state correlated with trainee's subjective mental demand (r=0.77, p=0.01), physical demand (r=0.68, p=0.03) and temporal demand (r=0.64, p=0.01) No significant differences were observed when the expert trainer surgeon was observing a trainee perform the task to when performing it.
Bahrami [264]	Technical performance was measured by time and accuracy. Spatial brain activation was measured for the whole brain.	9 subjects practiced all 5 simulated laparoscopic tasks for ten sessions. Each practice session lasted 30 minutes, which were conducted thrice a week. At the end of training participants performed these tasks whilst cognitively monitored using a non-ferromagnetic, fMRI compatible box trainer & surgical instruments.	Participants expectedly performed better in less complex unimanual laparoscopic tasks (no comparative test used).	Greater spatial brain activation was observed during execution of complex tasks such as bimanual peg transfer (Lt PM, SMA and SPL) and knot-tying (bilateral PM, SMA, SPL, IPL, MOG, MTG & Lingual gyrus), especially in the SMA than in the easier unimanual pointing task (Lt PM, SMA, M1 and S1) During unimanual task performance, greater activation was recorded when executed right–handed (Lt PM, SMA, M1, S1, IPL, MOG, MTG) compared to when executed left-handed (Rt PM, SMA and SPL).
Bahrami [265]	Technical performance was measured by time and accuracy. Spatial brain activation was measured for the whole brain.	1 hour training & familiarisation of fMRI. Subjects were cognitively monitored during performance of 5 laparoscopic tasks using a customized, fMRI compatible box trainer & surgical instruments.	Participants expectedly performed better in less complex unimanual laparoscopic tasks (no comparative test used).	Extent of brain activation increased with task complexity. Bilateral SMA activation was observed the most complex knot-tying task.

Author	Outcomes	Method	Other results	Brain results
Morris [271]	2 independent assessors rated technical performance quantitatively (number of knots) & qualitatively (square or slip knots).	Subjects performed a block-designed experiment of rest (no motor activity), executing the surgical task, imagining the surgical task and finger-tapping.	Experts outperformed intermediates who in turn outperformed novices, yet lacked statistical significance.	Significantly lower M1 activation ($p<0.001$) was observed in experts compared to novices during the surgical task. Conversely, experts displayed greater V1 ($p<0.05$) activation than novices for the imagining task.
Wanzel [256]	Mental rotation test (MRT) assessed by speed and accuracy. Differential brain activation (Experimental task – Control task)	Block designed experiment composed of MRT of quasi-3D figures and a control task (visual recognition). A previous study had demonstrated correlation between MRT performance and surgical task performance (Z-plasty)[297]	Inter-subject variability in MRT scores (Range: 17/50 to 42/50)	Differentially higher activation was observed in bilateral parietal regions, IFG, MFG and Lt PM ($p<0.01$). Greater activation in relation to MRT performance was observed in left hemispheric regions of middle temporal, posterior cingulate-precuneus and PM regions ($p<0.01$).
Ohuchida [29]	Technical performance was assessed by number of knots tied. Cortical activation was assessed in the frontal lobe by task related changes in oxyhaemoglobin (HbO_2), deoxyhaemoglobin (HHb) and total haemoglobin (THb).	21 participants were cognitive monitored via a block-designed experiment consisting of fixed rest phases (20 seconds) and task phases (60 seconds of knot tying). A subset of novices with no prior (n=4) and minimal exposure (n=4) were provided an additional 2 hours of practice and were then re-assessed via the same block-designed experiment.	Experts outperformed trainees who out performed novices, yet no statistical comparison was made.	Experts and all but one novices showed insignificant frontal lobe activation. Trainees and one novice with minimal experience displayed significant activation ($p<0.05$) on a number of channels derived by HbO_2 and THb analysis. Additional training for novices with no prior exposure displayed increased cortical activation ($p<0.05$) whilst for novices with minimal prior exposure showed reduction in frontal lobe activation ($p<0.05$) on HbO_2 analysis.
Leff [260]	Technical performance assessed by time, path length and number of movements via a motion tracking device (ICSAD)[112]. High dimensional fNIRS data derived from multi-channel, multi-haemoglobin species (HbO_2, HHb and THb) was embedded into a lower dimensional topographical position using isomapping techniques. EMD calculated differences in cortical activation.	62 participants of varied expertise were provided 15 minutes for task familiarisation. Novices were provided an hour of training. Block designed experiment composed of unfixed task periods and fixed rest periods. Novices in addition received 5 × 40 minute sessions of training provided over a week and underwent similar repeat testing.	Consultants and registrars showed no differences in technical performance yet both significantly out-performed novices ($p<0.001$). Training improved novices' performance significantly ($p<0.001$).	Greater clustering in the isomap was commensurate with expertise. Experts and registrars clustered tightly whilst novices showed greatest dispersion. On receipt of training novices showed lesser dispersion.
Leff [262]	Cortical activation of Lt PFC & Rt parietal cortex was derived by changes in HbO_2 & HHb. fNIRS data was dimensionally reduced by isomapping into a geospace. Markov chain was used to detect temporal flow in geo space to analyse fronto-parietal haemodynamic behaviour.	14 novice participants performed a block-designed experiment of fixed rest and motor task periods. Laparoscopic tasks were performed first under normal view, followed by 90-degree clockwise rotation, then 90-degree counter-clockwise rotation and finally under normal view conditions again.	Technical performance not presented.	Markov chain suggests global convergence towards a region of minimal activation. Greatest convergence was seen in the final transition from counter clockwise to normal view conditions.

Author	Outcomes	Method	Other results	Brain results
Leff [259]	Technical performance assessed by time, path length and number of movements using a motion tracking device (ICSAD)[112]. Cortical haemodynamics of PFC measured by changes in HbO_2 and HHb.	62 participants of varied expertise were provided 15 minutes for task familiarisation, Novices were provided an hour of training. Block designed experiment composed of 5 repetitions of self-paced task and fixed rest periods.	Consultants and registrars displayed stable technical performance, out-performing novices ($p<0.001$). Novices recorded within trial improvement (2^{nd} – 3^{rd}; $p<0.01$).	Consultants and registrars displayed stability in PFC haemodynamics. Novices displayed initial rise in cortical activation (rise in HbO_2, with concurrent drop in HHb; $p<0.05$) followed by attenuation ($p<0.05$) related to improvements in performance.
Leff [261]	Technical performance assessed by time, path length and number of movements using a motion tracking device (ICSAD)[112]. Cognitive task assessed by accuracy of response. Introspective Epworth Sleepiness scale for measurement of fatigue. Cortical haemodynamics of PFC measured by changes in HbO_2, THb and HHb.	7 participants over the course of a night from 10 PM until 8 AM were assessed at 2 hourly intervals (total of 6 sessions). During the assessment, participants were required to perform 6 trials of the surgical task and one trial of the cognitive task in a block designed experiment comprised of fixed rest and self-paced task periods.	According to the Epworth Sleep Scale, participants rapidly fatigued between 10PM until 2AM ($p<0.03$), after which no further decline was observed. Significant variance was recorded in technical performance ($p<0.05$) where initial improvement was followed by decline and subsequent stability. Non-significant variation was observed in performance of the arithmetic task.	Cognitive task evoked significantly greater PFC excitation than the surgical task on both HbO_2 and THb analysis ($p<0.001$). Cognitive task evoked cortical responses increased over the course of the night on both HbO_2 and THb analysis (p<0.01
James [258]	3 independent experts, assessed performance. PFC activation was assessed by both significant rise in HbO_2 and significant drop in HHb. Stress was assessed by HR, salivary cortisol and STAI.	29 participants provided 5 minutes for familiarity. Block designed experiment of fixed task and rest periods. Participants had to navigate between two set targets as many times as possible during the task period.	Experts out performed novices ($p<0.05$) based on good inter-rater reliability (Cronbach's alpha 0.997). No significant fluctuation in HRV was observed in either group supported by no significant rise in salivary cortisol levels,	Greatest activity was observed in LPFC where experts significantly activated more channels than novices (6 channels versus 3 channels).
James [253]	Motor task was assessed by accuracy and time. Cortical activity in the PFC and PPC was analysed by changes in HbO_2 and HHb. Fronto-parietal brain network created to measure network economy	21 participants were blinded and randomised into a control group (unassisted learning, n=11) and experimental group (gaze-contingent motor channelling guidance, n=10). Both groups underwent 6 sessions of training over 8 days and tested for retention 2 months later.	Improvement in performance was observed in both groups ($p<0.001$). Initial performance improvements from session 1 to 3 were greater in the control group. However from session 3 to 6, and on retention the experimental group performed better.	Experimental group displayed more rapid attenuation in cortical activation of PFC and PPC. Similarly experimental group displayed a more economical cortical network
Leff [254]	Motor task was assessed by number of biopsies taken and accuracy of movement (path length). Gaze behaviour of participants were tracked. Cortical activation in the occipital cortex was analysed by changes in HbO_2, HHb & THb. Occipital brain network was created to measure network econometrics.	20 participants were required to take virtual biopsies from 7 locations in randomised order under either verbal instruction (control) or visual guidance (experimental) from their trainer. Trainer's eye fixation was tracked and displayed on the participant's screen to provide visual guidance. All participants performed the task under both conditions in randomised order in a block-designed experiment.	Participants under visual guidance (experimental) performed better in number of virtual biopsies taken ($p=0.003$) and accuracy ($p<0.001$). Gaze behaviour was more focussed ($p<0.001$) in the experimental group.	Experimental group participants displayed lesser occipital activation according to HbO_2 ($p<0.001$) and THb ($p<0.001$) analysis. No significant differences were observed in functional brain network econometrics.

Author	Outcomes	Method	Other results	Brain results
Crewther[257]	Motor task was assessed for time and accuracy. Stress was measured perceptually by (STAI) and mental work load (NASA-TLX), and physiologically by heart rate variability, and hormonal markers (salivary and cortisol levels). Cortical activation in the PFC was analysed by changes in HbO_2.	12 participants were trained over 7 sessions across 2 weeks totalling to 8 hours of practice. Testing occurred after 2 hours (BASE), 5 hours (MID) and 8 hours (POST) of practice. 4 weeks later retention was tested similarly (RETEST). Assessments involved a block-designed experiment of self paced task and rest periods.	Technical performance improved from BASE to MID and MID to POST & RETEST ($p<0.001$). Perceptual stress (STAI) decreased from MID to POST ($p<0.001$) and workload (NASA-TLX) reduced between BASE & MID to POST & RETEST ($p<0.018$). Non-significant fluctuation in hormonal levels was noted.	PFC activation remained elevated in all sessions ($p<0.035$) with no lateralisation.
Miura [269]	Task performance was not assessed. Cortical activation in the intra-parietal sulcus was analysed by changes in HbO_2.	5 novice participants performed 5 trials of a robotic suturing sub-task (insertion of needle) under 5 variant OATVA conditions (15°, 30°, 45°, 60°, 75° & 90°).	Performance not assessed	IPS activation was greatest at the point when the needle was inserted. Brain activation was most significantly elevated at an angle of 75° OATVA.
Leff [263]	Confidence of decision measures by a 6 point subjective scale. Decision making was checked for consistency and in accordance to the script. Spielberger STAI subjectively measured stress. Cortical activation in the PFC was analysed by changes in HbO_2 and HHb.	22 participants were briefed and assessed for their knowledge of the operative task. 12 video clips of a simulated laparoscopic cholecystectomy procedure were randomly played at terminated at the juncture where an intra-operative decision is required. 5/12 videos were primed i.e. the next step was declared by the presence of an instrument in the video and the remaining 7/12 were unprimed (undeclared).	Experts and residents were more confident than novices ($p<0.001$). Experts and residents decision making was more consistent ($p<0.001$ and $p<0.05$ respectively) than novices ($p=0.183$). Experts also agreed to a greater extent than residents and novices with the script of the videos (experts=90% vs residents 78.3% vs novices 53.3%). Stress levels between groups were insignificant ($p=0.574$)	During decision making for unprimed videos novices significantly activated more channels in the PFC (14/22 channels) than residents or experts (4/22 channels each)

Table 4.2. Summarizes outcomes, methods, results of brain behaviour and other metrics of performance.

4.4 Discussion

This systematic review synthesizes current evidence for the application of functional neuroimaging in objective assessment of technical and non-technical skills, and track changes in the brain associated with longitudinal skill acquisition. It also demonstrates the feasibility of a variety of functional neuroimaging techniques to interrogate surgeons' brain behaviour during performance of technical and non-technical tasks.

Of the reviewed studies, most employed fNIRS as the investigative modality, understandably due to its advantages of non-invasiveness, safety, portability, relative resistance to motion artefacts, cost-effectiveness and allows the surgeon to maintain his natural posture. The advantages of PET and fMRI are that it captures whole brain activity as a result it helps in locating neural correlates for a particular technical task, however the disadvantages are the need to perform tasks in a lying down posture, incremental costs, prohibition of ferro-magnetic instruments for fMRI and the administration of radioactive substances for PET. Although EEG offers good temporal resolution its disadvantages are poor spatial information. fNIRS in comparison to EEG also allows scrutiny of multiple motor regions, which play a role at various stages of skills learning. Therefore for the reasons stipulated above fNIRS was the chosen investigative modality for the purposes of this thesis.

The vast majority of studies included in this review examined technical skills over non-technical skills of which MIS skills were more often scrutinized than open skills, understandably due to the scale of the problem faced in MIS skills acquisition. Cortical behaviour modulates in response to learning a property referred to as neuroplasticity and these changes are captured by mapping loco-regional brain activity or by inter-regional brain activity otherwise known as connectivity[305]. Several studies have established strong links between neuroplasticity and motor practice or learning[25,277,306]. In accordance with motor learning literature, novices required greater amount of neuronal resources than more experienced individuals. This was observed by the greater amount of activation of multiple regions in the frontal, parietal and occipital lobe for laparoscopic[266] and open surgical tasks[271].

However despite almost a decade's investigation into brain behaviour of surgeons, several gaps in our understanding of motor skills acquisition exist. The PFC a measure of attention allowed differentiation between experts and novices for simpler motor tasks[259,260] (open hand knot tying) but did not characterise differences between experts and intermediates[259,260,271] presumably because the skill was proficiently acquired by intermediates.

One would expect a harder task to help characterization between different groups of expertise, but James et al[258] in a NOTES task demonstrated the opposite effect. More interestingly Ohuchida et al[29] during performance of a challenging laparoscopic task (laparoscopic suturing) detected experts and intermediates displayed PFC activity commensurate with motor learning literature, yet novices with no experience at all displayed similar to experts minimal PFC activity. Only after receiving practice novices began to display significant PFC activity. The lack of PFC activity detected during the very early naive phase in performance of a complex bimanual laparoscopic task warrants further scrutiny, which may suggest disengagement in the task. However, several methodological flaws were identified in the design of this study namely the uneven distribution of subjects in each expertise group, the task was performed under temporal demand and most importantly does not provide adequate spatial information on which specific brain regions within the frontal lobe were evaluated[29]. Spatial information whilst interrogating a motor sequence task is vital as the frontal lobe is composed of functionally diverse brain regions namely the PFC, SMA, PMC and M1, which are involved at various stages of motor learning[25,245,246,307].

In order to circumvent gaps in knowledge of expertise development it is important to go beyond simple expert versus novice comparisons and instead track progression with skills acquisition[308,309]. Few studies included in this review have employed a longitudinal design to track skills learning understandably due to the difficulty in ensuring adherence to a strict training protocol. Reznick et al[18] advocated that surgical skills acquisition must follow motor learning models where trainees with practice progress from the novice “cognitive” phase to expert “autonomous” phase[310]. Experts display the characteristic of automaticity, which essentially is the ability to perform a skill with little effort and minimal demands on attention[46]. Rather than testing automaticity indirectly by measuring the ability to multi-task, Leff et al[235] proposed that learning associated reduction in attentional demands could be investigated by monitoring the PFC, a brain region that is responsible for attention, concentration and performance monitoring. During early stages of motor skills learning, task performance is reliant on attention and working memory in order to select the appropriate response to sensory feedback received, often learning by response to errors[235,277,278]. The PFC is one of several regions implicated to serve as a “scaffold” to cope with novel task demands, which gives way on well-honed practice[279].

Several studies of human bimanual motor sequence learning, have confirmed that the PFC is essential during early learning stages and with practice it attenuates trending towards automaticity[311-313]. This is further validated by studies in the aviation and motoring industry where practice or expertise in a task has inversely correlated with magnitude of PFC activity[314,315]. Hence the results of several studies[29,253,259,260,266] included in this review, which display attenuation of PFC activity in experts are inline with motor learning literature. However as this effect was not observed in all cross-sectional studies[29,258], it would be of added benefit to incorporate other regions associated with motor learning to investigate for associated changes and reorganisation of cortical activity.

During early motor learning, changes are driven by an ‘associative loop’ comprised of frontal, parietal and pre-motor regions of the cortex. Later changes during advanced stages are driven by a ‘motor loop’ comprised of posterior and deeper regions of the brain which include motor areas of the cortex, putamen and cerebellum[245,307]. The PMC, SMA and M1, all located in the frontal lobe play a crucial role in both early and late stages of motor learning[25,306]. The PMC is responsible for movement selection according to visual cues[316] and storage of sequential motor skills[306]. The SMA not only plays a vital role in encoding sequential motor skills but also in bimanual co-ordination[281,288,311]. Lastly the M1 plays a role at all stages of learning but activity varies according to temporal scales. From longitudinal studies M1 activity initially decreases during early learning due to 'habituation' followed by increase in activation at later stages due to 'specialisation'[317,318].

A complex MIS task such as laparoscopic suturing and knot tying serves as useful task to employ for interrogation of motor learning. Firstly it is challenging to acquire serving as a barometer to perform complex MIS operations warranting several investigations into assessing and learning the skill[27,29,80,82,100,114,121,228,270,319,320]. Secondly the motor skill involved has elements of learning sequences of movements and under multi-sensory adaptation. Therefore for reasons stipulated, this task has been chosen for the purpose of this thesis.

In summary, this collective body of literature represents important steps to identify neural correlates of skills learning and surgical skills acquisition. On the basis of current evidence, the benefits of functional neuroimaging represent an emergent stream of research with abundant potential towards enhancing performance in surgery. All but one study included in this review were undertaken in a simulated environment and the one study that did not, monitored a trainer's cognitive engagement during mentoring of trainees[268]. Although it is appreciably important to progress towards cognitive monitoring in a live OR environment, a proper understanding of the neural correlates underlying MIS skills acquisition within a controlled environment must be first realised. Noticeably, exclusion criteria in several experiments[254,256,264,267,269,270] were not sufficiently stringent. Participants were neither of homogenous gender nor handedness, both are recognised factors that can cause variable brain activation patterns during motor performance[199,321]. Elimination of left handed and female participants might not be the solution, considering that approximately 10% of surgeons are left handed[322] and women are increasingly representative of the surgical workforce, however studies must undertake sub-analysis accounting for these factors.

4.5. Research Hypothesis

The findings of this systematic enables formulation of the following series of research hypotheses with regards to mapping functional neuroplasticity of complex bimanual MIS skills learning. Based on motor learning literature they are defined as follows:

- Cortical activity in the frontal lobe will vary based on expertise and practice.
- Novices in the early attention demanding cognitive phase will display profoundly greater PFC activity than trainees and experts. In addition as the task is novel they will recruit secondary motor regions responsible for encoding learnt sequential movements to a lesser degree.
- Experts who have reached an autonomous phase of motor skills learning in comparison to intermediates and novices will display minimal PFC activity and greater SMA activity a region associated with late stages of skills acquisition.
- Intermediates will display the greatest PMC activity a region associated with initial stages of skills encoding than both experts and novices. In addition they will display greater PFC activity than experts but lesser SMA activity than novices.
- Novices on receipt of extended practice, will display frontal lobe neuroplasticity consisting of i) attenuation in PFC activity, ii) an initial rise followed by a decrease in PMC activity and iii) a sequential rise in SMA activity. In addition M1 activity with practice will show an initial decrease in activity due to habituation followed by an increase in activity due to specialisation.
- Lastly when novices who have attained proficiency are tested for retention after an extended interval of no practice, will display regression by greater attentional needs manifested by greater PFC activity than at the final phase of their training period.

Chapter 5

Frontal lobe neuroplasticity associated with expertise in Laparoscopic Intracorporeal Suturing.

Contents from this chapter have been published in:

Shetty K, Leff DR, Orihuela-Espina F, Yang GZ, Darzi A. Persistent Attentional Demands Despite Laparoscopic Skills Acquisition. *JAMA Surgery* 2016 Jul 1; 151(7): 682-684.

5.1. Introduction

Laparoscopic intracorporeal suturing and knot tying (LICS) is an advanced surgical manoeuvre that is technically challenging to perform owing to ergonomic constraints. Proficiency in LICS is an essential component of the Fundamentals of Laparoscopic Surgery curriculum (FLS)[263], which all emerging General surgery trainees must achieve in order to be certified by the American Board of Surgeons[323]. Critically, LICS serves as a barometer to perform advanced laparoscopic procedures, as the inability to achieve stable performance in this skill is likely to constrain progress in performing more complex minimally invasive surgery (MIS)[27]. Furthermore, complex MIS skills necessitate protracted practice and yet working-time regulations deplete live training opportunities[13] leaving less time to acquire these skills. In this regard, technologies that capture a trainee's capability to learn, reliably track the stage of motor learning and evaluate the phase of task internalisation are likely to enhance both objective assessment and surgical selection.

Current assessments of LICS performance including dexterity analysis[320,324], blinded video ratings[320] and end product assessment[27] have improved objectivity in assessment of technical skills but fail to evaluate the internal model of motor skills learning which occurs at the level of the central nervous system[25]. Technical skills training should emulate "cognitive models" that describe the way in which skills are acquired and internalised, in the hope of tracking progression from the early cognitive phases through to associative and finally automated stages of performance[18,42], which may explain why certain trainees progress rapidly whilst others struggle. However, to date, the focus of technical skills assessment in surgery has centred on behavioural manifestations of learning *i.e.* motor outputs and not necessarily cognitions or surgeon brain function; arguably leading to oversimplified interpretations regarding progress and performance. For example, two individuals with similar motor performance may be mistaken to have made equivalent technical progress and yet may differ significantly in terms of workload and attentional demands that portray underlying differences in their levels of expertise[80]. Similarly, at times of stress or fatigue significant changes in attention may occur, potentially limiting capacity for safety critical judgements despite seemingly stable technical performance[261].

Monitoring demands on surgeons' attention is clearly as important as evaluating their technical skills and is more likely to illuminate differences in capabilities to respond to unanticipated events, comprehend instructions and make operative decisions. For example, expert surgeons through deliberate practice not only refine technical skills but also enhance their capacity for dual-task performance (automation). Conversely, novices attending solely to the technical aspect of the procedure are less likely to have additional brain resources to cope with extra demands[21]. Such differences in the allocation of attentional resources are evidenced through disparity in the magnitude of recruitment (known as *"activation")* of executive control and motor centres in the brain. Brain activation and inter-regional brain communication (known as *"connectivity")* are dynamic and change in response to practice, learning or expertise[25,246,277]. This phenomenon termed *"neuroplasticity"* occurs at several levels within the central nervous system and can be inferred by changes in maps (spatiotemporal resolution) of brain activation[245]. Specifically, early stages of motor skills learning rely on executive control regions such as the prefrontal cortex (PFC). However, with practice, PFC activation typically attenuates[25,227,245,246,259,279,325] and recruitment shifts to the premotor cortex (PMC), supplementary motor area (SMA), striatum and cerebellum[25,227,245,246,325] as illustrated in Figure 5.1.

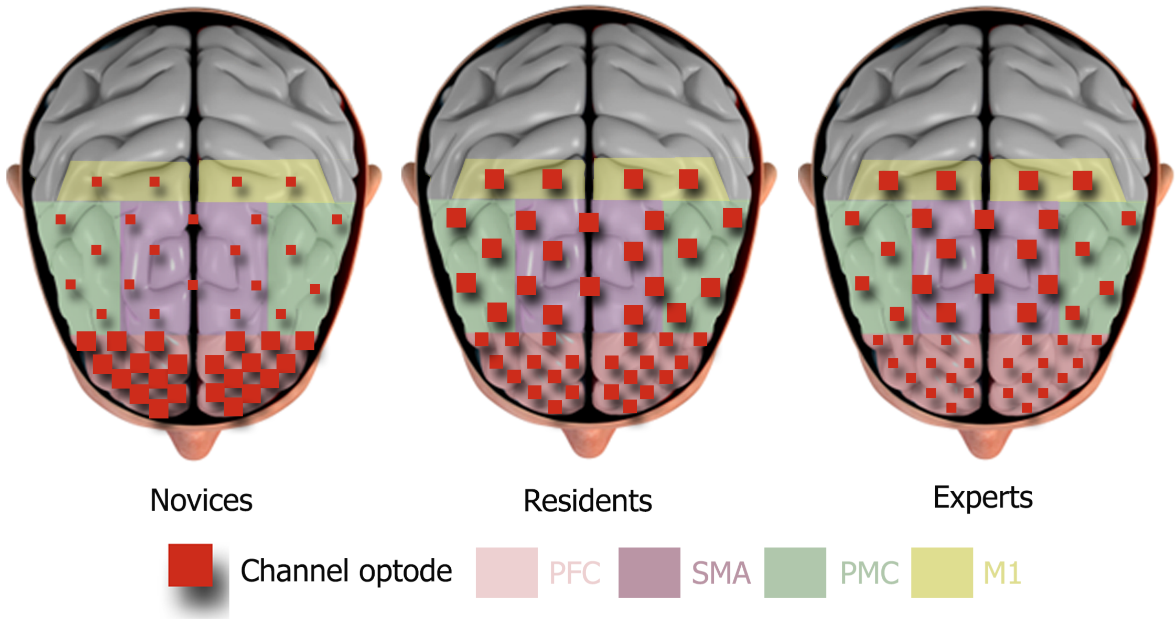

Figure 5.1. Hypothetical regional activation maps based on predicted learning associated neuroplastic changes of cortical substrates located in the frontal lobe. Schematic illustration depicts the relative importance (denoted by size of red squares) of the PFC (prefrontal cortex), PMC (dorsal premotor cortex), SMA (supplementary motor area) and M1 (primary motor cortex) according to expertise.

Advances in functional neuroimaging technology now permit non-invasive *mapping* of dynamic changes in brain activation associated with technical skills training in realistic environments[277]. Investigators have recently capitalised on these developments to map learning-related changes in surgeon brain behaviour during open[259,260] and laparoscopic skills acquisition[29,265,266] using a range of neuroimaging technologies (see Chapter 4). These studies highlight the complexity of brain activations during MIS[29,265,266], confirm the relative importance of the PFC during novel tasks demands[259] and are in line with learning-related attenuation in executive control[25,227,245,246,259,260,277,279,325]. However, for MIS skills in particular published studies suffer methodological flaws and are incongruous with motor learning theory[29]. Neuroimaging studies of surgical brain behaviour have focused predominantly on activation and not inter-regional communication, interrogate the PFC but not other relevant brain regions[246], and have failed to adequately deconstruct complex MIS skills to facilitate delineation of subtask response[29]. In particular, further research is required to clarify an apparent prefrontal redundancy amongst novices during LICS[29] which is contrary to the theory that the PFC acts as a scaffold to support novel tasks demands[279] and fails to align with the observation that novices significantly activate the PFC during open surgical knot-tying[259,326].

The theory that underpins this thesis is that there may be a preliminary stage in motor learning where individuals fail to engage the PFC when naive to an extremely complex motor task, hence the apparent similarity between novices and experts in the extent of PFC recruitment during LICS (see Figure 5.2). Moreover, we anticipate that differences between expert surgeons and novices are more likely to manifest as dissimilarities in the recruitment of secondary motor areas (e.g. PMC, SMA) and frontal brain connectivity that have not been sufficiently evaluated to date.

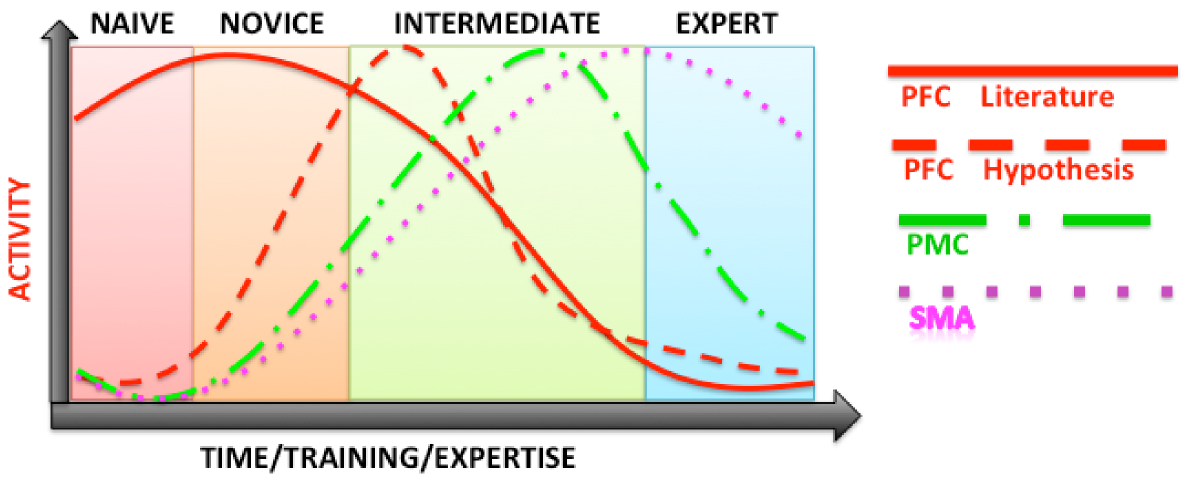

Figure 5.2. Motor learning hypotheses for complex bimanual skills in laparoscopic surgery. Attenuation of Prefrontal cortical (PFC) activation (bold red) with expertise in-line with motor learning literature in contrast to alternate tested PFC activation (dashed red line) for a complex motor task. Similarly, predicted neuroplastic changes within the Premotor Cortex (PMC - green) and Supplementary Motor Area (SMA - magenta) are illustrated.

Indeed, recent data suggests that repeated deliberate skills practice leads to adaptation or modulation in the strength of functional connectivity[299]. Theoretically, with more time on the task, the brain network is moulded in such a way to form neighbourhoods (modularity) with increased local connectedness whilst retaining fewer long-range connections to minimise the overall network characteristic path-length (see Figure 5.3). These network properties commonly characterise complex biological networks[327] and are collectively referred to as *"small world"* topology. A network displaying small world topology has distinct advantages over a *random network* or a *lattice network* in its capacity for information transfer[328]. Repeated task practice has been shown to increase the small world topology of functional brain networks[329]. Therefore, it is conceivable that through years of training, the brain of the expert surgeon is wired to maximise efficiency, minimise redundancy and optimise small world properties. In contrast, the brain network of a novice surgeon is more likely to be a characterised by a haphazard arrangement of spurious and random connections.

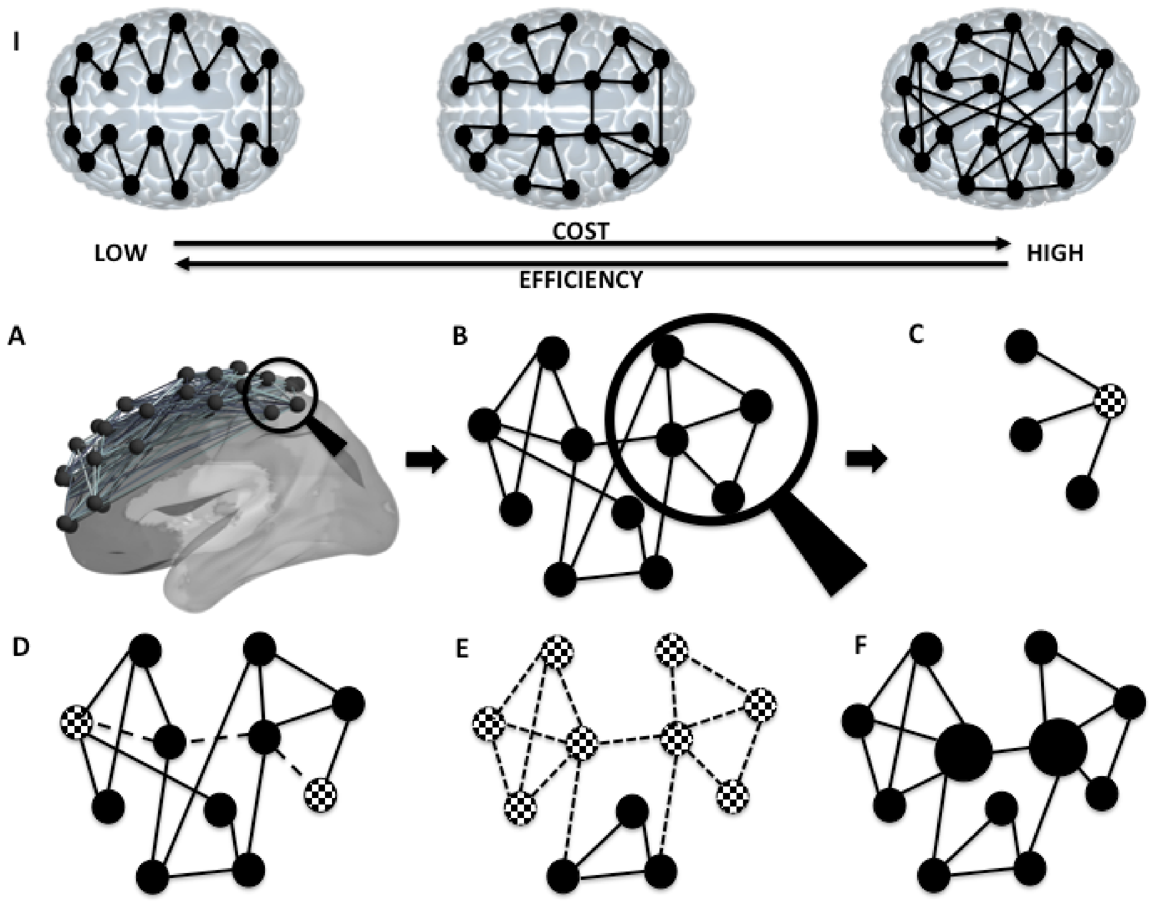

Figure 5.3. Schematic illustration describing brain network architectures and graph theoretical properties. Network architectures (I) the brain network on the left is of lattice topology, which is of low cost but does not favour fast processing due to the lack of integration. The figure on the right is a costly network that has greater number of edges of which many are random and of great distance. The network in the centre is moderate in cost and efficiency as it has numerous local connections in nodes close together and few long-range connections depicting a network property called *"small worldness"*, which is optimized for information transfer. This arrangement also results in a smaller mean characteristic path length. Under the focused portion of the functionally connected brain [A] a segment of the network [B] consists of nodes (dots) connected by edges (lines). The degree of the node [C] of interest (shaded) is 3, derived by the number of connected edges to other nodes. Shortest path length between the shaded nodes is 3, denoted by the least number of edges (dashed lines) connecting them [D]. Modularity represents the clusters of locally connected nodes forming neighborhoods [E]. Hubs are the largest nodes in the network attributed by the highest number of edges [F].

In this observational study I seek to address the limitations of previous studies by separately monitoring the cortical response to each LICS sub-task (e.g. needle drive insertion, double-throw, etc), evaluating both PFC as well as secondary motor regions and analyse expertise-dependent disparity in activation, network connectivity and small worldness. The hypotheses are as follows: (a) novices will not activate the PFC after limited training on LICS and any differences in PFC recruitment between novices and experts shall be statistically insignificant; (b) activations in secondary motor brain regions will be significantly greater in trainees and experts than novices and (c) lower network costs and greater 'small world' network topology will characterise expert laparoscopists.

LICS can be sub-categorised into *'iso-directional'* phases such as needle insertion and *'anti-directional'* phases such as knot-tying phases, which may place greater demands on executive control processes since anti-phase co-ordination patterns are less stable, harder to perform and require greater attention and practice [291,330], In this regard, LICS sub-phases place idiosyncratic demands on bimanual co-ordination, are uniquely influenced by directional interference and therefore offer a paradigm for further exposing differences in brain response to *'anti-directional'* versus *'iso-directional'* co-ordination patterns, and allows identification of task phases that may enable better discrimination between learners. Finally, (d) based on directional interference, '*anti-directional'* knot-tying tasks are hypothesised to burden the PFC and SMA more than '*iso-directional'* needle insertion tasks.

5.2. Methods

5.2.1. Subjects

Following Local Research Ethics Committee (project number: 05/Q0403/142) approval, 32 right-handed male subjects were recruited from the National Health Service and Imperial College London. Subjects were consented and screened for neuropsychiatric illnesses and for left-handedness. Participants abstained from consuming caffeine or alcohol for 24 hours to avoid any spurious effects on cerebral haemodynamics[199]. Based on previous operative experience subjects were classified as novices (n=12), trainees (n=11) and experts (n=9).

Novices comprised of medical students (mean age ± S.D. = 22.4 ± 1.6 years) with no laparoscopic experience. Trainees (PGY 3-8), (mean age ± S.D. = 33.8 ± 2.8 years) were categorized based on their laparoscopic experience of less than 50 procedures that required LICS. Experts, (mean age ± S.D. = 42.7 ± 3.6 years) comprised entirely of bariatric and upper gastrointestinal surgeons who were board certified and had performed over 100 procedures that required LICS.

5.2.2. Setting and Laparoscopic Intra-corporeal Suturing Task

The experiment was conducted in a controlled laboratory setting devoid of distracting stimuli (see Figure 5.4). Subjects performed three sets of LICS in a box trainer [iSurgicals, UK] on a penrose drain which had pre-marked entry and exit points on either side of a laceration as used in the FLS curriculum[263]. LICS was performed with laparoscopic needle holders (model 26173KC; Karl Storz GmBH& Co, Tuttlingen, Germany) using a standardized length of 15 cms of 2/0 Polysorb suture material (Ethicon Ltd, Sommerville, NJ).

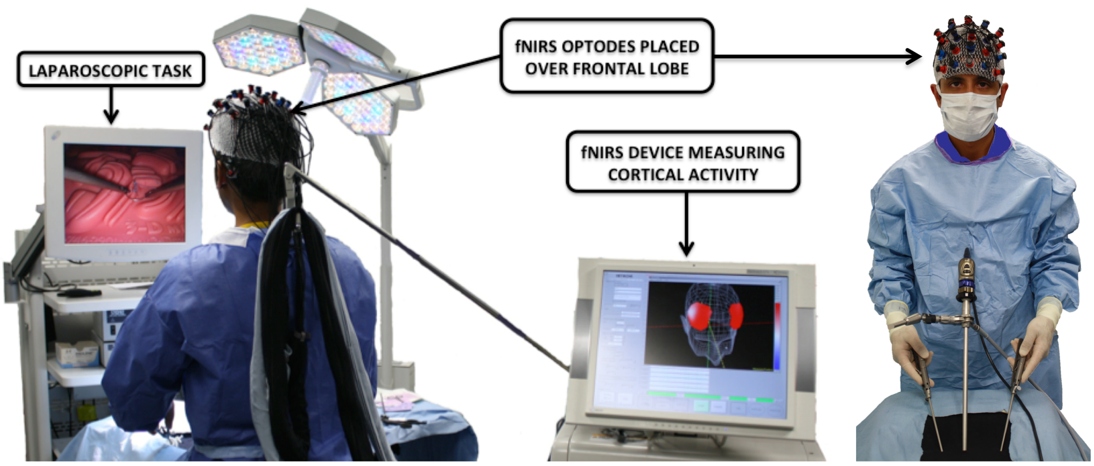

Figure 5.4. Experimental task set-up displays participant performing laparoscopic intracorporeal suturing and knot tying (LICS) whilst simultaneously OT (ETG-4000 Hitachi Medical Corp.) monitored.

LICS is a serial motor skill requiring complex *'iso-directional'* and *'anti-directional'* movement sequences to achieve stitching and knot tying within a confined compartment. The operator uses long rigid instruments under multi-sensory adaptation to perform a motor task in a confined workspace (see Figure 5.4). Here, LICS was deconstructed into three sub-tasks (see Figure 5.5) namely needle drive insertion (NDI), double throw knot (DTK) and a pair of single throw knots (STK) [319,331]. Task deconstruction is necessary for a complex serial skill acquisition since it facilitates part-practice [52], and allows for the identification of rate limiting sub-tasks across LS acquisition. Methodologically, task deconstruction also aimed to reduce the variance in task durations (i.e. between novice and experienced operators) and facilitate assessment of the evoked brain response by each sub-task.

In the first sub-task namely NDI (see Figure 5.5 Panels A-B) the operator's dominant hand (right) adopts a manipulative role to grasp the needle in the optimal position and angulation, and insert it through the edges of the tissue. Movements of the non-dominant hand (left) although '*iso-directional*', play a supportive and postural role. DTK involves at first synchronised precise bimanual rotatory *'anti-phase'* movements to enable twice looping the long end of suture around the needle-holder. The short end of suture is then grasped and the knot tied through opposing forces on the suture, which apposes the wound edges (see Figure 5.5 Panels C-D). The final sub-task STK requires formation of a single loop, tying a knot and repeating this sequence by inverting the hand gestures (see Figure 3 Panels E-F). Unlike NDI, during DTK and STK both hands play a manipulative role requiring greater synchronization of *'anti-phase'* motions. However, DTK is noticeably more exigent because the property of the suture makes it prone to uncoiling, meaning that looping the suture twice mandates a greater degree of stable bimanual movement control.

In this study LICS differed from the task 4 described in the FLS curriculum[1] only in that the needle needed to be grasped in the right orientation before tissue insertion as opposed to being pre-mounted. The justification for modifying it was two-fold; firstly grasping it in the right orientation is an integral element of the skill rated using checklists[319] and secondly it would obviate any interaction with the participant during the inter-trial period when a fresh set is provided after one trial of task is completed. All novices first viewed an instructional video (5 minutes) followed by a training session (2 hour) involving one to one LICS training. To ensure between-subject consistency in the type and delivery of training, the same instructor delivered each training session. Upon training completion, novices were assessed to ensure that LICS could successfully be performed within 600 seconds before being able to undertake the study. The maximum time restriction was based on the notion that any additional time would result in a negative FLS score. Prior to study commencement, trainees and experts were provided with time for task familiarisation (15 minutes). A block design experiment (see Figure 5.5) was conducted comprising of three blocks each incorporating episodes of baseline 'rest' episodes (30s), 'task' (variable temporal duration) and 'post-task' rest (40s). During episodes of rest, subjects were instructed to remain still and view the centre of a 2D video-monitor. During task episodes, subjects were instructed to perform a given sub-component of LICS on a Penrose drain as accurately and swiftly as possible, paying equal attention to both criteria.

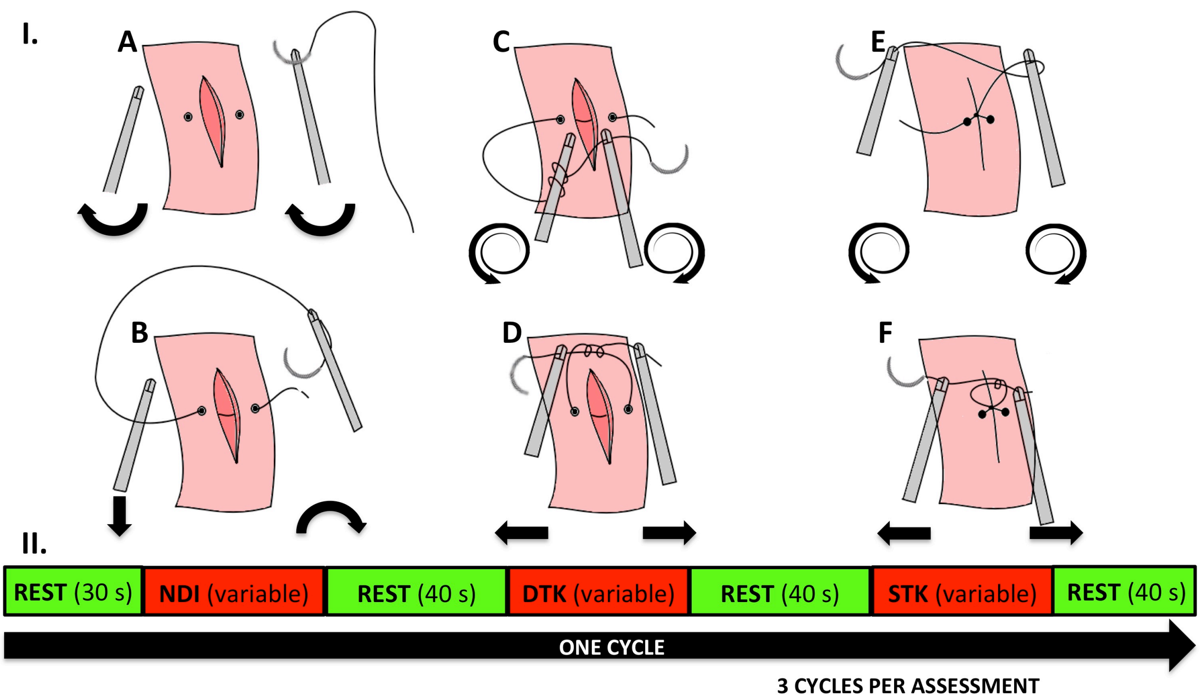

Figure 5.5. Displays one cycle of laparoscopic intra-corporeal suturing (LICS). LICS sub-tasks were performed in a block-design experimental paradigm comprised of rest periods (green) which alternated with sub-tasks (red) namely needle drive insertion (NDI), double throw knot (DTK) and a pair of single throw knots (STK).

5.2.3. Task Performance and Mental Workload

All experiments were recorded to allow two independent assessors undertake retrospective 'blinded' video ratings of performance. Task completion times obtained from the video were combined with end-product analysis of sutured Penrose drains to generate a LICS objective performance score for each subject using a validated formula [Score = 600 – (time in seconds) - (penalties $\times$10)][27]. Therefore, performance assessment accounted for both time on the task and inaccuracies (penalties on the quality of the knot, opposition of the edges and distance of the stitch from the pre-marked entry and exit points on the Penrose drain). Upon study completion all participants were requested to complete a NASA Task Load Index (NASA TLX)[332] questionnaire for each LICS sub-task. NASA-TLX measures subjective workload and has been found to correlate with measures of visual attention and mental stress during MIS[83]. Between-group comparisons of workload and performance data were analysed using the Kruskall-Wallis test ($p<0.05$). Post-hoc comparisons between any given two experience groups were conducted using the Mann-Whitney U test (Bonferroni adjustment = $p<0.017$) for non-parametric data. Within-group comparisons of workload were performed first using a Friedman test ($p<0.05$) and post-hoc paired analyses were performed using a Wilcoxon Sign Rank test (Bonferroni adjustment $p<0.017$) to determine which sub-phase of LICS was most burdensome. Statistical analysis was performed using IBM SPSS v20; SPSS Inc., Chicago, IL, USA.

5.2.4. Cortical Haemodynamics and Cortical Activation

Recruitment of brain tissue for function is conveniently tightly coupled to changes in cortical haemodynamics. 'Neurovascular coupling' describes the phenomenon that brain activation results in a change in cerebral blood flow leading to an increase of oxygenated haemoglobin (HbO_2) and a decrease in deoxygenated haemoglobin (HHb) which is washed out of an activated area faster than oxygen consumption[157]. Optical imaging techniques such as OT facilitate detection of local haemodynamic changes indicating the magnitude of brain activation simultaneously at multiple different brain regions (referred to as "channels"). As a functional neuroimaging modality, OT is non-ionising, portable, relatively resistant to motion artefacts, and allows the use of ferromagnetic devices within the study field making it suitable for interrogating surgeons. In OT, each channel represents a volume of brain tissue between a Near Infrared (NIR) emitter and detector where brain activation is measured by modulation in cortical haemodynamics. A channel was considered "activated" if a task-evoked increase in HbO_2 was coupled to a decrease in HHb of which either reached statistical significance (WSR p<0.05), commensurate with the physiology of neurovascular coupling[157].

5.2.5. Optical Topography

Cortical activity was recorded throughout using a commercially available 44-channel OT system (ETG-4000; Hitachi Medical Corp, Tokyo, Japan). In this OT experiment, a series of detectors and emitters emitting NIR Light at 690 and 830 nm were placed 3 cm apart in thermoplastic holders in a configuration of two 3×5 arrays over the frontal lobe as depicted (see Figure 5.6).

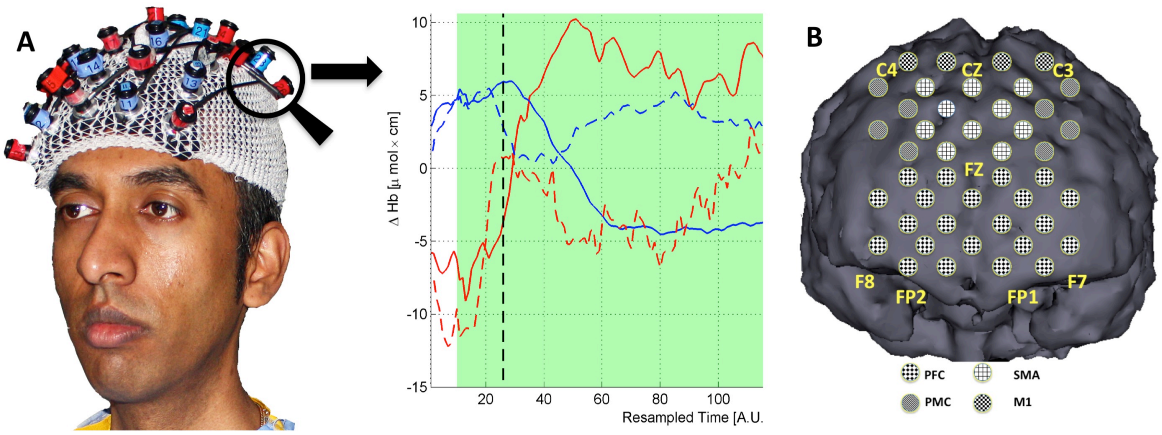

Figure 5.6. Participant with OT optodes (A) positioned over the brain regions of interest. Optodes are either emitters (red) or detectors (blue) arranged in a grid like fashion and between them lie *channels*, which represent a loci of cortical tissue assessed. In a given channel (under the lens) cortical activity is recorded. As depicted in the signal recording a rise in HbO_2 (bold red line) and concomitant fall of HHb (bold blue line) signify activation. The converse represented by dotted lines does not portray activation. (B) 44 channels positioned according to the 10-10 International System. 22 channels were located over the Prefrontal Cortex (PFC) and an additional 22 channels (relabeled as 23-44) were located over the dorsal Premotor Cortex (PMC), Supplementary Motor Area (SMA) and Primary Motor Cortex (M1).

Positioning of emitters and detectors were guided by adopting the International 10-10 system to ensure that channels were positioned accurately over the brain regions of interest by measuring and placing scalp markers at ten positions in relation to standard reference points. Registration of emitter / detector positions to underlying cortical channel position was achieved by obtaining 3D locations of all emitter / detector positions and five fixed reference points about the subjects' scalp namely nasion, inion, bilateral external auditory meati and the vertex using a 3D digitizer (Isotrak, Polhemus, Vermont, US)[333]. 3D positional data of a representative case was projected onto an individual magnetic resonance image to visualize the estimated cortical structures underlying each position using 3D Composite Display Unit (Hitachi Medical Co., Japan). The cortical structures identified (see Figure 5.6 and 5.7) underlying each channel were PFC (22 channels), right and left PMC (8 channels), SMA (10 channels) and M1 (4 channels).

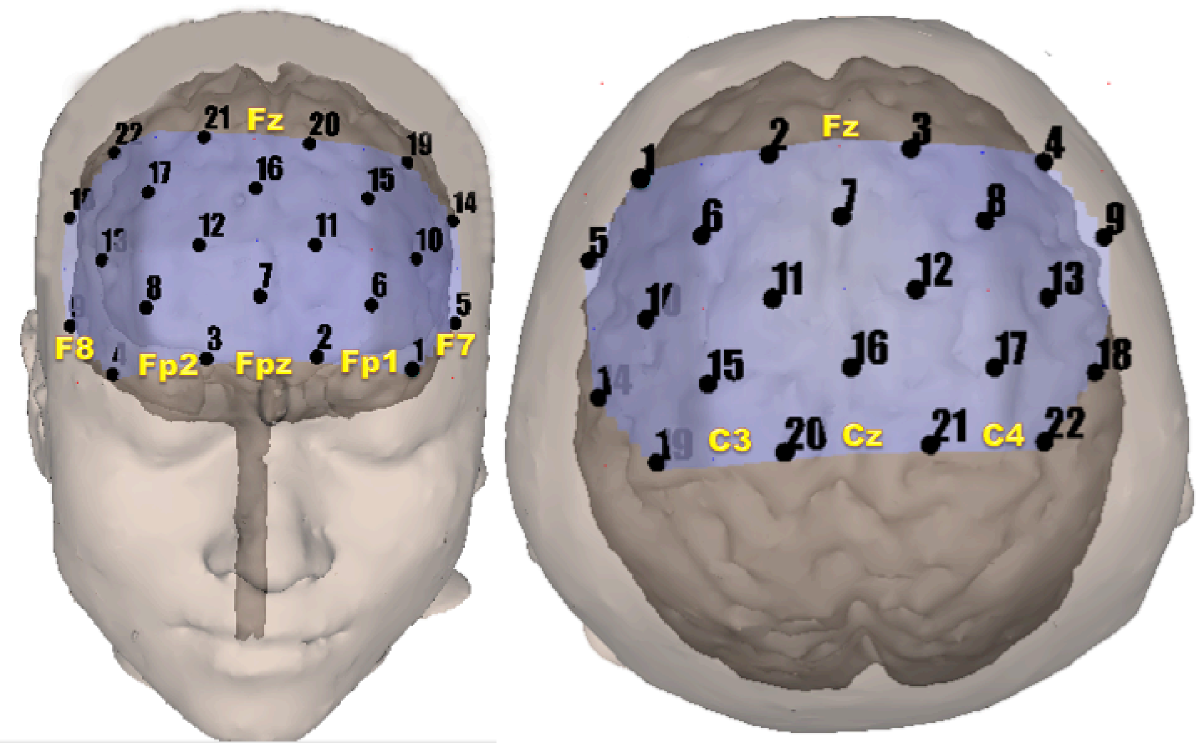

Figure 5.7. Channel locations (black dots and numbers) were obtained by transferring topographical data to a 3D cortical surface of a T1-weighted MRI image. Relative International 10-10 cranial markers are illustrated in yellow. (Left) The first set of 22 channels (channel number 1-22) was positioned over the Prefrontal Cortex. (Right) Second set of 22 channels (relabelled as channel number 23-44) are positioned over the dorsal Premotor Cortex (PMC), Supplementary Motor Area (SMA) and Primary Motor Cortex (M1).

5.2.6 Pre-processing, Filtering, Data Integrity of Cortical activity

Optical data was processed utilising a bespoke Matlab-based software package named Imperial College Neuroimage Analysis (ICNA)[334]. Relative changes in optical data were converted to changes in haemoglobin data by applying the modified Beer-Lambert law (MBLL). Converted haemoglobin data was decimated from 10Hz to 1Hz to reduce physiological noise and linearly detrended to eliminate system drift before undergoing integrity checks to eliminate unreliable data caused by abrupt optode displacement or saturation of the detectors[199]. For a given OT channel, data was resampled to overcome between-subject variations in LICS subtask durations such that the resampled baseline rest (10s), LICS task (55s) and post task rest episodes (10s) were consistent across subjects. The sampling for LICS task was determined by calculating the average time to peak for all task blocks. Rest data used for resampling comprised only 15s of data prior to task onset, thereby allowing for temporal delay in the recovery of cortical haemodynamics following task offset. Average rest was calculated by averaging each block of resampled rest data. Temporal windows of data representing a given LICS subtask for analysis comprised of 55s of data commencing 6s after task onset, allowing for the inertia in cortical haemodynamics following stimulation. For a given LICS subtask, average task data was calculated by averaging blocks of these task windows.

Two complementary analytical approaches were adopted to evaluate and compare task-evoked activation(s). The first approach enabled the detection of the spatial location of channel activations within each experience group. Specifically, for a given channel, LICS subcomponent and haemoglobin species, rest Hb data was compared to task Hb data. A channel was considered “activated” if an increase in HbO_2 was coupled to a decrease in HHb of which either Hb species reached statistical significance (WSR $p<0.05$).

The second approach facilitated between-group comparisons in the magnitude of cortical activation. A new variable $\triangle$Hb was generated by subtracting rest from task data (i.e. $\triangle Hb = HbTask - HbRest$) for every given channel of data and Hb species. The magnitude of the increase in $\triangle HbO_2$ and decrease in $\triangle HHb$ reflects the magnitude of channel activation. For a given region of interest (ROI), the ROI$\triangle$Hb represents an average of the $\triangle$Hb across all channels in that ROI for each sub-task of LICS. In order to compare the cortical activation of the overall LICS task, a variable "*combined LICS*" was computed for both Hb species that was the average $\triangle$Hb of all sub-tasks. Between-group comparison in ROI$\triangle$Hb was conducted using the Kruskall-Wallis (KW) test ($p<0.05$). For a given ROI, if the difference in $\triangle$Hb between groups was statistically significantly, the Mann-Whitney U (MWU) test was employed to compare ROI$\triangle$Hb between any two experience groups. To account for total potential number of comparisons (×72), Bonferroni correction was applied, adjusting for threshold of significance to $p<0.0007$.

Finally in order to eliminate confounding effects of practice and learning, only data from experts was used to compare cortical haemodynamic changes ($\triangle HbO_2$ and $\triangle HHb$) across the PFC, PMC, SMA and M1, between sub-tasks based on differential interference. Specifically, for a given ROI, ΔHb data acquired during *'anti-phase'* co-ordination tasks namely DTK and STK was compared to data obtained during the *'in-phase'* NDI task using the MWU test for significance ($p<0.05$).

5.2.7. Stress

Cerebral haemodynamics are known to be influenced by stress induced changes in sympathetic autonomic activity[232]. Hence, heart rate (HR) was continuously monitored throughout the experiment using a portable electrocardiogram (ECG) fastened as a belt over the chest of each participant (Bioharness v2.3.0.5; Zephyr Technology Limited, Annapolis, MD). HR data obtained during rest and task periods were separated accordingly. R-waves of the ECG were detected using the software package ICNA and confirmed on meticulous visual inspection. Mean HR and heart rate variability (HRV) was computed by calculating the R-R wave intervals and its standard deviation (SD_{RR}). Stress can be inferred by an increased HR and decrease in SD_{RR}[335]. Between-group comparisons were conducted using the Kruskall-

Wallis (KW) test ($p<0.05$) and if significant followed by a Mann-Whitney U (MWU) test ($p<0.05$) for paired group comparisons.

5.2.8. Graph construction

Figure 5.8 (steps I to V) illustrates the workflow of graph construction. First, expertise-based group averaged haemodynamic data was obtained for each LICS sub-task. An *association matrix* was generated by cross correlation of the group averaged haemodynamic time course (step III)[253]. In the association matrix, each channel can be considered as a 'node' in the network while the functional associations derived from the bidirectional cross correlation between every two nodes form the edges as graphically represented in Figure 5.3. The association matrix was subsequently pruned to form an *adjacency matrix* by applying a threshold based on the strength of inter-nodal associations as defined by the Floyd-Warshall (FW) algorithm [336], which calculates all paths between any two given nodes and chooses the shortest, most traversed path that maximises flow of information transfer between those two nodes in terms of efficiency[336]. This is performed by accounting for the two outputs provided by the algorithm, which are a weight (the sum of shortest distance between two nodes) and a sequence of edges between nodes to achieve it. This approach yields a total of nine network graphs, each representing a sub-task of LICS and an operator-experience group. From each of these graphs, data describing the network's topology was extracted in terms of node metrics (degree), edge metrics (number of edges, path length) and combined node–edge metrics (i.e. modularity and small worldness)[337] as illustrated in Figure 5.3. These network metrics were used to compare neural efficiency in frontal brain behaviour according to surgical expertise.

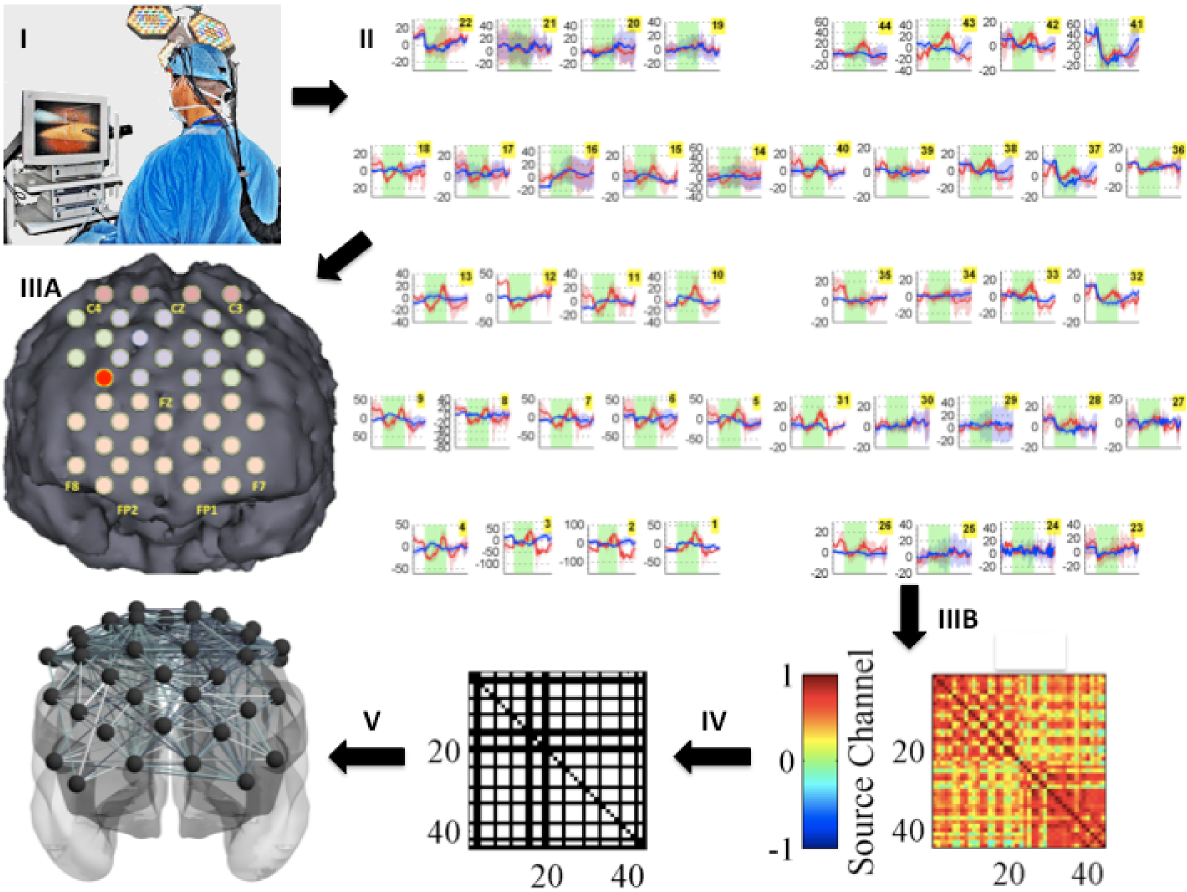

Figure 5.8. Work flow for cortical network graph generation. *Step I*: recording of cortical activity; *Step II*: group-averaged Hb data for all 44 channels; *Step IIIA:* mapping of cortical activation; *Step IIIB:* compute an association matrix which represents the bidirectional cross correlation (edges) between every two channels (nodes), the colour scale represents positive correlations (red) and negative correlations (blue); *Step IV:* weaker correlation are pruned from the association matrix to form the final adjacency matrix, and finally *Step V* displays the topographical representation of the adjacency matrix (correlations = edges and channels =nodes) to form the cortical network.

5.2.9. Network Small-World Topology

Network small world topology, a signature property observed in real world complex networks[327] is characterized by high local clustering and optimised long-range connections facilitating short mean characteristic path lengths in network circumnavigation (see Figure 5.3). A measure of small world topology known as the small-world index (SWI)[338] was computed by taking the average value of the ratio between the quotient of the clustering coefficients and characteristic path lengths of the derived network against 100 random networks, each of which preserve the same degree distribution as the derived network while maintaining connectedness[328,337]. A network with a SWI of greater than one is considered to be small world[338]. Between-group comparison in SWI was performed using the Kruskall-Wallis test ($p<0.05$), and post-hoc comparison between any two groups of expertise was performed using the Mann-Whitney U test ($p<0.05$).

5.3. Results

5.3.1. Participants, Exclusions and Data Integrity

A total of 30 of 32 participants successfully completed three LICS procedures with concomitant frontal cortical brain monitoring. Two subjects were excluded from the study. One novice could not complete a double-throw knot within the stipulated time allocation leading to abandonment of the experiment. Data for one expert was excluded as a fire alarm triggered during the experiment leading to premature cancellation of the experiment. Following data integrity checks[199] a number of channels were excluded due to artefacts or noise [i.e. 181 out of a total of 1364 (13.2 %)].

5.3.2 Objective Assessment of Technical Skill

For each subject, median FLS scores were calculated across three LICS procedures and were then used to quantify expertise-related differences in technical skill (median FLS scores ± IQR: Experts = 487 ± 53, trainees = 400 ± 90, novices = 334 ± 76, KW: $p<0.001$). As anticipated (see figure 5.9), experts outperformed trainees (MWU: $p<0.001$) and trainees outperformed novices (MWU: $p<0.001$).

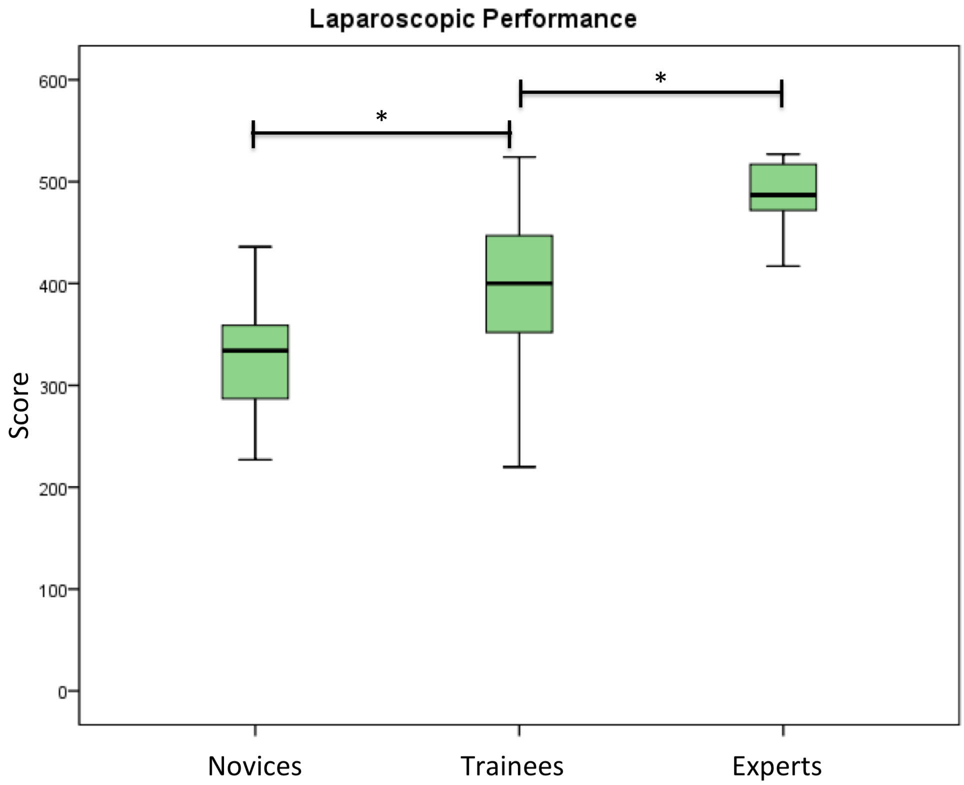

Figure 5.9. Box and whisker plots illustrate expertise related differences in technical performance of LICS evaluated as per FLS scoring criteria (KW p<0.001). Experts outperformed trainees who in turn outperformed novices (*=paired group comparison MWU p<0.001).

5.3.3. Subjective Assessment of Cognitive Workload

Significant differences in cognitive burden were observed at the level of expertise (median NASA TLX scores ± IQR = experts: 29.16 ± 24.8, trainees: 54.66 ± 31.33, novices 63.33 ± 36.50, KW: p<0.001). Post-hoc analysis revealed that experts found LICS less cognitively challenging than both trainees (MWU p<0.001) and novices (MWU p<0.001). Although novices appeared to find LICS more burdensome than trainees, differences in workload indices did not reach statistical significance (MWU p=0.128). As illustrated in Table 1**,** upon LICS segmentation novices found needle drive insertion and double-throw knots significantly more burdensome than experts (NASA TLX scores median ± IQR: novices NDI = 64.33 ± 36.65, experts NDI = 32.16 ± 24.08, MWU p=0.026; Novices DKT = 72.33 ± 26.33 versus experts DKT = 38.33 ± 22.83, MWU p=0.016). Mental workload did not vary between LICS sub tasks in trainees and experts, whereas novices found the double-throw knot significantly more challenging than either needle drive insertion and single-throw knots (median NASA

TLX scores ± IQR: NDI = 64.33 ± 36.65 versus DKT = 72.33 ± 26.33, WSR p<0.05; SKT = 52.33 ± 27.67 versus DKT = 72.33 ± 26.33, WSR p<0.01).

NASA TLX	**Experts [E]** (Median ± IQR)	**Trainees [T]** (Median ± IQR)	**Novices [N]** (Median ± IQR)	***p values*** (Group comparison)
LICS Overall	29.16 ± 24.83	54.66 ± 31.33	63.33 ± 36.50	KW **p<0.001** E vs T **p<0.001** E vs N **p<0.001** T vs N p=0.128
[NDI] Needle Drive Insertion	32.16 ± 24.08	51.66 ± 30.67	64.33 ± 36.65	KW **p=0.029** E vs T p=0.069 E vs N **p=0.026** T vs N p=0.101
[DTK] Double Throw Knot	38.33 ± 22.83	56.33 ± 35.00	72.33 ± 26.33	KW **p=0.031** E vs T p=0.075 E vs N **p=0.016** T vs N p=0.217
[STK] Single Throw Knots	20.83 ± 47.25	51.33 ± 34.00	52.33 ± 27.67	KW p=0.200 E vs T p=0.062 E vs N p=0.238 T vs N p=0.847
Friedman **p=0.04** NDI vs DKT **p=0.006** NDI vs SKT p=0.271 DKT vs SKT **p=0.002**	Friedman **p=0.135** NDI vs DKT p=0.161 NDI vs SKT p=0.575 DKT vs SKT p=0.123	Friedman p=0.441 NDI vs DKT p=0.213 NDI vs SKT p=0.477 DKT vs SKT p=0.286	Friedman **p=0.012** NDI vs DKT **p=0.021** NDI vs SKT p=0.062 DKT vs SKT **p=0.010**	***p values*** (Sub-task comparison)

Table 5.1. Expertise based group comparison of subjective workload (NASA TLX). The lowest row represents sub-task comparisons within each group.

5.3.4. Heart Rate Variability and Stress

No significant difference was observed between groups in mean HR during LICS (mean HR ± S.D: experts = 82.14 ± 9.68, trainees = 84.64 ± 11.79, novices = 81.43 ± 11.69, ANOVA p =0.48).

5.3.5. Cortical Activation Maps

Group averaged cortical activation maps for each LICS sub task are depicted in Figure 5.10. The extent and spatial location of frontal activation(s) was observed to vary according to surgical expertise and LICS subtask, as follows:

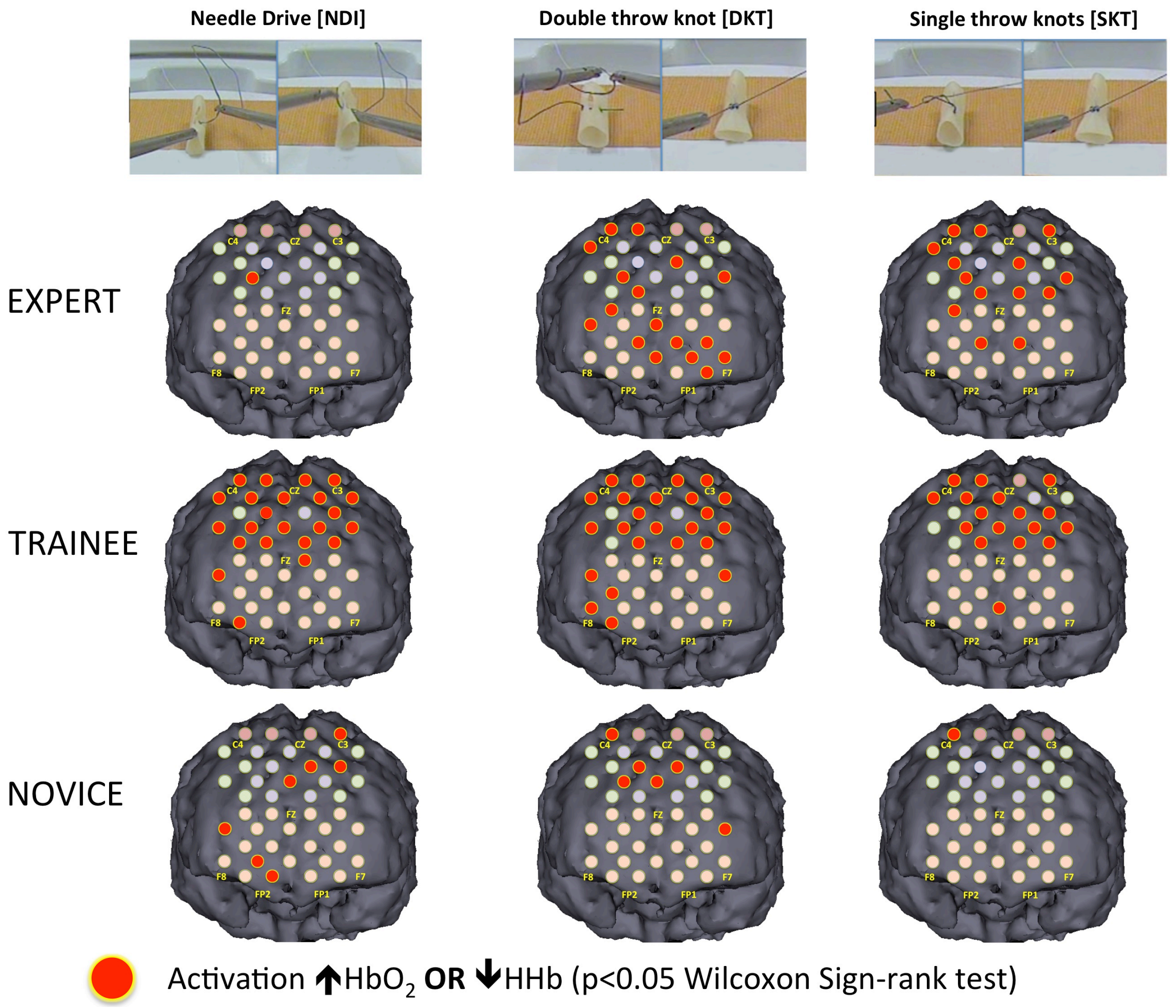

Figure 5.10. Schematic illustration depicting the magnitude of group averaged cortical activation for each LICS sub-task. Channels reaching statistical threshold for activation are denoted in red. Channels failing to reach statistical threshold for activation retain their colour representation within each region of interest (PFC = orange, PMC = green, SMA = violet and M1 = pink).

Experts: Engagement of the PFC was found to be greater during double-throw knots versus single-throw knots and needle drive insertion (%PFC: DKT=45.4, SKT=13.6 and NDI=0). Engagement of the PMC, SMA and M1 was observed to be greater during single throw knots versus double throw knots and needle drive insertion (%PMC = SKT: 50.0, DKT: 25.0 and NDI: 0; %SMA: SKT=40.0, DKT=30.0 and NDI= 10.0; %M1 = SKT=75.0, DKT=50.0 and ND=0).

Trainees: In comparison to experts and novices, trainees displayed cortical responses of greatest intensity in the PMC, SMA and M1 regions. PMC engagement was found to be greater during needle drive insertion versus double-throw and single throw knots (% PMC = NDI: 87.5, DKT: 75.0, SKT: 50.0). SMA engagement was observed in almost all channels across all LICS sub-tasks (%SMA = 90.0). All M1 channels were activated during needle drive insertion and double-throw knots (%M1 = 100).

Novices: Limited recruitment of the PFC was observed during needle drive insertion and double-throw knot tying (%PFC = NDI: 13.6, DKT: 4.5, SKT: 0). Additionally, compared to trainees and experts, recruitment of the PMC, SMA and M1 was less apparent amongst novices. For example, significant activation could not be detected in the PMC during knot-tying manoeuvres and in the SMA during single-throw knots.

Sub-Task	Group	PFC (n=22)	PMC (n=8)	SMA (10)	M1 (n=4)
Needle Drive Insertion [NDI]	Experts	0 (0%)	0 (0%)	1 (10%)	0 (0%)
	Trainees	3 (13.6%)	7 (87.5%)	9 (90%)	4 (100%)
	Novices	3 (13.6%)	1 (12.5%)	2 (20%)	1 (25%)
Double Throw Knot [DTK]	Expert	10 (45.4%)	2 (25%)	3 (30%)	2 (50%)
	Trainee	5 (22.7%)	6 (75%)	9 (90%)	4 (100%)
	Novice	1 (4.5%)	0 (0%)	4 (40%)	1 (25%)
Single Throw Knot [STK]	Expert	3 (13.6%)	4 (50%)	4 (40%)	3 (75%)
	Trainee	1 (4.5%)	4 (50%)	9 (90%)	3 (75%)
	Novice	0 (0%)	0 (0%)	0 (0%)	1 (25%)

Table 5.2**.** Regional number and percentage of activated channels for each group and sub-task are tabulated.

5.3.6 Inter-Group Comparison of Cortical Activation

For brevity, I focus on between-group comparisons in ROI△HbO_2 data (μM•cm), and △HHb results with respect to supporting or refuting the specific hypotheses of interest. Cortical haemodynamic change in all ROI was found to significantly depend on surgical expertise (KW: $p<0.05$), with the exception of M1 during DTK (see tables 5.3 and table 5.4).

Prefrontal Cortex: As depicted (see Figure 5.10), during DTK, greater PFC cortical haemodynamic change was observed amongst novices and trainees when compared to expert surgeons [ΔHbO_2 (μM•cm) median ± IQR: novices = 10.19 ± 21.75, trainees = 7.46 ± 12.57, experts = 1.56 ± 11.37; KW $p<0.05$, novices > trainees > experts MWU: $p<0.001$]. Novices displayed greater PFC excitation than experts for NDI [ΔHbO_2 (μM•cm) median ±

IQR: novices = 6.10 ± 21.68, trainees = 5.24 ± 11.49, experts = 1.49 ± 12.55; KW: $p<0.05$, novices > experts MWU: $p<0.001$].

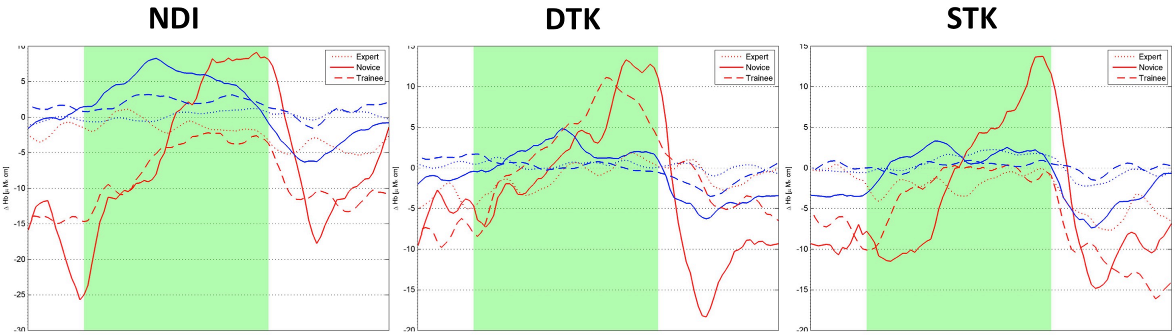

Figure 5.11. Group averaged changes in HbO_2 (red) and HHb (blue) concentration, within a channel centred on the right prefrontal cortex during sub-tasks needle drive insertion (NDI), double throw knots (DTK) and single throw knots (STK) are illustrated. a) Median change in Hb species (µM•cm) is depicted by expertise for the cross-sectional experiment. Dotted, dashed and bold lines represent Haemodynamic signals of experts, trainees and novices respectively.

Premotor Cortex: Regardless of sub-task, the greatest PMC excitation was observed amongst trainees. For NDI, greater cortical haemodynamic change (ΔHbO_2 increase / ΔHHb decrease) was found in trainees compared to novices [NDI: ΔHbO_2 (µM•cm) median ± IQR: novices = 4.03 ± 15.97, trainees = 7.91 ± 14.46, experts = 5.44 ± 14.03; KW: $p<0.05$, trainees > novices MWU: $p<0.001$]. For knot-tying tasks, greater PMC activation was observed amongst trainees compared to experts and novices [DTK: ΔHbO_2 (µM•cm) median ± IQR: novices = 7.71 ± 19.94, trainees = 8.78 ± 16.68, experts = 4.71 ± 14.18; KW: $p<0.05$, trainees > experts MWU: $p<0.001$; STK: ΔHbO_2 (µM•cm) median ± IQR: novices = 1.83 ± 14.82, trainees = 9.54 ± 14.19, experts = 6.61 ± 13.97; KW: $p<0.05$, trainees > experts > novices MWU: $p<0.001$].

Supplementary Motor Area: The magnitude of task evoked SMA cortical haemodynamic change was found to be significantly greater amongst trainees and experts than novices across all LS sub-tasks (see table 5.4 and figures 5.12 and 5.13) [NDI: ΔHbO_2 (µM•cm) median ± IQR: novices = 2.50 ± 14.62, trainees = 6.70 ± 13.02, experts = 6.37 ± 12.17; KW: $p<0.05$, trainees and experts > novices MWU: $p<0.001$; STK: ΔHbO_2 (µM•cm) median ± IQR:

novices = 2.89 ± 10.51, trainees = 6.57 ± 12.65, experts = 7.15 ± 13.30; KW: $p<0.05$, experts and trainees > novices MWU: $p<0.001$].

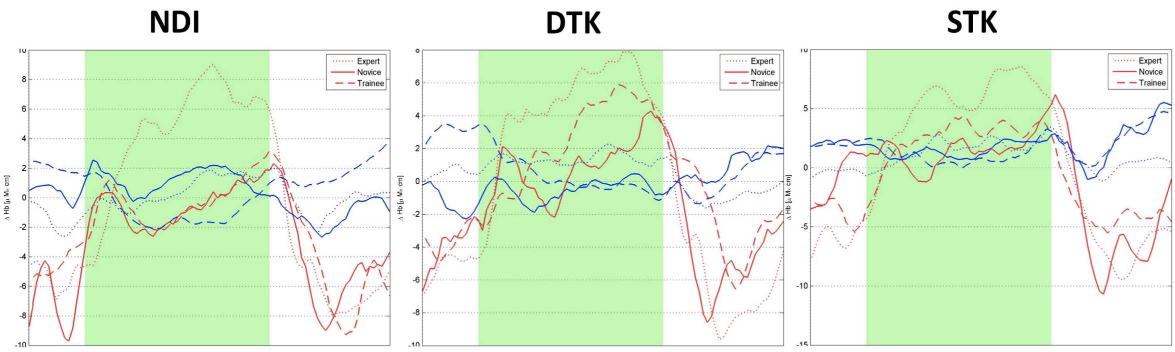

Figure 5.12. Group averaged change in HbO_2 (red) and HHb (blue) concentration within a channel of the supplementary motor area during sub-tasks needle drive insertion (NDI), double throw knot (DTK) and single throw knots (STK) are illustrated. Median changes in µM•cm for experts, trainees and novices are represented in dotted, dashed and bold lines.

Primary Motor Cortex: No expertise related differences in M1 cortical haemodynamic change was observed across knot tying sub-tasks [i.e. DTK and STK]. However for NDI, greater M1 cortical oxygenation was observed in line with expertise [ΔHbO_2 (µM•cm) median ± IQR: novices = 3.35 ± 12.94, trainees = 7.80 ± 12.91, experts = 8.16 ± 12.20; KW: $p<0.05$, trainees and experts > novices MWU: $p<0.001$].

Region	Sub-task	Novices vs. Trainees		Novices vs. Experts		Trainees vs. Experts	
		ΔHbO_2	ΔHHb	ΔHbO_2	ΔHHb	ΔHbO_2	ΔHHb
PFC	NDI	↓	↓*	↓*	↓*	↓*	↓*
	DTK	↓*	↓*	↓*	↓*	↓*	↓*
	STK	↑*	↓*	↔	↓*	↓*	↓*
PMC	NDI	↑*	↓	↔	↓	↔	↔
	DTK	↔	↓	↔	↔	↓*	↑*
	STK	↑*	↔	↑*	↔	↔	↓*
SMA	NDI	↑*	↔	↑*	↔	↔	↔
	DTK	↔	↔	↔	↔	↔	↑*
	STK	↑*	↔	↑*	↔	↔	↔
M1	NDI	↑*	↔	↑*	↔	↔	↔
	DTK	↔	↔	↔	↔	↔	↔
	STK	↔	↔	↔	↔	↔	↔

Table 5.3. Depicts between-group differences in cortical haemodynamic change between paired comparisons of novices, trainees and experts. Regions examined are the prefrontal cortex (PFC), dorsal premotor cortex (PMC), supplementary motor area (SMA) and primary motor cortex (M1). Comparative increase and decrease in oxygenated haemoglobin change ($+\Delta HbO_2$) and deoxygenated haemoglobin change ($-\Delta HHb$) are correspondingly represented in symbols as ↑ and ↓. Significant trends (threshold $p<0.0007$) are additionally denoted with *. ↔ Represents non-significant differences between-groups in cortical haemodynamic change within a given brain region of interest and for a given LICS subtask.

Hb Species, Sub-task	ROI	Novices µmol/cm Median ± IQR	Trainees µmol/cm Median ± IQR	Experts µmol/cm Median ± IQR	*p* values Kruskall - Wallis	*p* values Mann Whitney U		
						Novice vs Trainee	Novice vs Expert	Trainee vs Expert
ΔHbO_2 NDI	PFC	6.10 ± 21.68	5.24 ± 11.49	1.49 ± 12.55	**<0.001**	0.312	**<0.001**	0.001
	PMC	4.03 ± 15.97	7.91 ± 14.46	5.44 ± 14.03	**<0.001**	**<0.001**	0.033	0.019
	SMA	2.50 ± 14.62	6.70 ± 13.02	6.37 ± 12.17	**<0.001**	**<0.001**	**<0.001**	0.609
	M1	3.35 ± 12.94	7.80 ± 12.91	8.16 ± 12.20	**<0.001**	**<0.001**	**<0.001**	0.859
ΔHHb NDI	PFC	2.52 ± 7.66	0.70 ± 5.87	-0.06 ± 3.65	**<0.001**	**<0.001**	**<0.001**	**<0.001**
	PMC	0.76 ± 5.86	-1.95 ± 7.84	-1.64 ± 5.31	**<0.001**	**<0.001**	**<0.001**	0.285
	SMA	-0.28 ± 7.05	-1.32 ± 6.88	-0.80 ± 4.96	0.002	0.007	0.592	0.001
	M1	-0.95 ± 7.65	-0.56 ± 5.71	-0.26 ± 4.53	0.838	0.574	0.660	0.982
ΔHbO_2 DTK	PFC	10.19 ± 21.75	7.46 ± 12.57	1.56 ± 11.37	0.001	**<0.001**	**<0.001**	**<0.001**
	PMC	7.71 ± 19.94	8.78 ± 16.68	4.71 ± 14.18	0.001	0.050	0.009	0.001
	SMA	5.52 ± 19.10	6.59 ± 12.44	5.53 ± 12.51	0.067	0.037	0.473	0.076
	M1	7.57 ± 20.85	6.85 ± 12.11	4.10 ± 13.62	0.212	0.864	0.249	0.057
ΔHHb DTK	PFC	3.01 ± 9.13	1.08 ± 5.67	0.09 ± 4.16	0.001	**<0.001**	**<0.001**	**<0.001**
	PMC	0.96 ± 7.41	-1.60 ± 6.65	0.46 ± 5.12	0.001	**<0.001**	0.302	**<0.001**
	SMA	-0.16 ± 7.70	-1.21 ± 7.22	0.16 ± 5.43	0.001	0.009	0.138	**<0.001**
	M1	0.81 ± 8.16	-0.92 ± 6.42	-0.05 ± 4.14	0.250	0.122	0.445	0.301
ΔHbO_2 STK	PFC	3.97 ± 19.09	8.50 ± 15.95	4.20 ± 11.45	0.001	**<0.001**	0.585	**<0.001**
	PMC	1.83 ± 14.82	9.54 ± 14.19	6.61 ± 13.97	0.001	**<0.001**	**<0.001**	**<0.001**
	SMA	2.89 ± 10.51	6.57 ± 12.65	7.15 ± 13.30	0.001	**<0.001**	**<0.001**	0.870
	M1	5.20 ± 13.60	7.50 ± 15.69	6.07 ± 11.03	0.042	0.025	0.039	0.920
ΔHHb STK	PFC	2.39 ± 7.16	1.14 ± 6.54	0.03 ± 3.82	0.001	**<0.001**	**<0.001**	**<0.001**
	PMC	-0.23 ± 6.21	-1.12 ± 8.94	-1.03 ± 5.56	0.332	0.214	0.830	0.199
	SMA	-0.53 ± 5.88	-1.74 ± 7.38	-0.42 ± 4.26	0.001	0.003	0.061	0.001
	M1	-0.59 ± 8.03	-0.49 ± 6.59	0.007 ± 3.90	0.649	0.883	0.377	0.451

Table 5.4. Comparison of the derived regional ΔROI Hb for HbO_2 and HHb species for each sub-task, namely: needle drive insertion (NDI), double throw knots (DTK) and single throw knots (STK). Regions interrogated were the prefrontal cortex (PFC), dorsal premotor cortex (PMC), supplementary motor area (SMA) and primary motor cortex (M1). Comparisons reaching statistical significance are represented in bold.

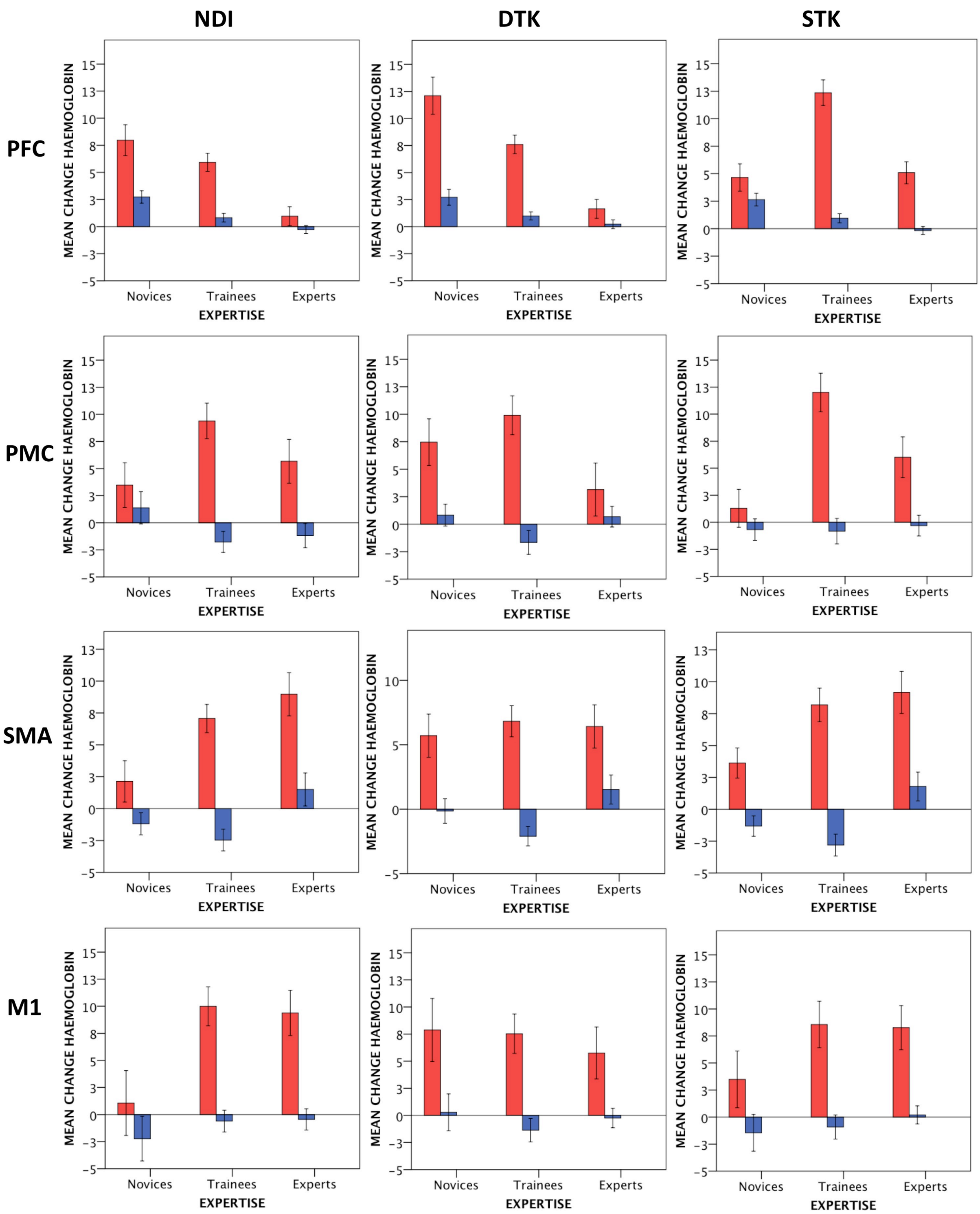

Figure 5.13. Bar charts illustrate comparative regional changes in HbO_2 (red bars) and HHb (blue bars) amongst novices, trainees and experts during LS subtasks: needle drive insertion (NDI), double throw knot-tying (DTK) and single throw knots (STK). Data represent mean and standard error bars. Regions interrogated are the prefrontal cortex (PFC), dorsal premotor cortex (PMC), supplementary motor area (SMA) and primary motor cortex (M1) respectively. Change haemoglobin (µm•cm) is represented on the y axis.

5.3.7. Directional Interference and Sub-Task Haemodynamic Response

As summarised in Table 5.5, significantly greater PFC cortical oxygenation change was observed amongst experts during *'anti-directional'* knot-tying tasks (KT) compared to *'iso-directional'* NDI tasks [ΔHbO_2 (µM•cm) median ± IQR: KT = 2.67 ± 11.30, NDI = 1.49 ± 12.55 MWU: $p<0.05$]. Conversely, needle drive evoked greater haemodynamic changes in PMC than knot-tying tasks [ΔHHb (µM•cm) median ± IQR: KT = -0.30 ± 5.36, NDI = -1.64 ± 5.37; MWU: $p<0.05$].

Haemoglobin species	Region of Interest	Needle Drive Insertion	Knot tying tasks (DTK & STK)	***p* values *Mann Whitney U***
ΔHbO_2	PFC	1.49 ± 12.55	2.67 ± 11.30	**<0.001**
	PMC	5.44 ± 14.03	5.30 ± 14.32	0.471
	SMA	6.37 ± 12.17	6.16 ± 13.08	0.394
	M1	8.16 ± 12.20	5.73 ± 12.11	0.037
ΔHHb	PFC	-0.06 ± 3.65	0.06 ± 4.04	0.329
	PMC	-1.64 ± 5.37	-0.30 ± 5.36	**<0.001**
	SMA	-0.80 ± 4.96	0.18 ± 5.03	0.077
	M1	-0.26 ± 4.53	-0.01 ± 3.98	0.419

Table 5.5. Comparison of derived regional ΔROI Hb for HbO_2 and HHb species for sub-task based on directional interference. *'In-phase'* needle drive insertion (NDI) was compared to *'anti-phase'* knot-tying sub-tasks namely double throw knots (DTK) and single throw knots (STK) in experts within the cross-sectional study. Regions interrogated were the prefrontal cortex (PFC), dorsal premotor cortex (PMC), supplementary motor area (SMA) and primary motor cortex (M1). Comparisons reaching statistical significance are represented in bold.

5.3.8. Networks and connectivity analysis

Expertise related variation in graph theory econometrics derived from nine derived frontal brain networks showed variations. Mean characteristic path length was observed to decrease with increasing expertise (mean ± S.D. = novices 0.23 ± 0.02, trainees 0.22 ± 0.04 and experts: 0.18 ± 0.02)**.** The resulting trainee and expert group networks consisted of nodes with fewer functional connections (i.e. edges) than that of novices (mean network edges ± S.D: novices = 356 ± 98.76, trainees = 256 ± 54.72, experts = 264 ± 39.95). On the other hand, modularity of the network was observed to increase with expertise (mean networks modularity ± S.D. = novices 0.26 ± 0.06, trainees 0.33 ± 0.09, experts: 0.36 ± 0.03).

As illustrated in Figure 5.14 and Table 5.6, the SWI varied according to surgical expertise. Expert brain networks were found to be smaller world than those of trainees who in turn generated networks with a greater SWI than novices ("*Combined LICS*" task SWI median ± IQR: novices = 1.46 ± 0.41, trainees = 1.68 ± 0.97, experts =1.87± 0.26, experts > trainees and novices MWU p<0.001, trainees > novices MWU p<0.001). Sub-task analysis confirmed a trend towards greater small worldness with increasing expertise, particularly for knot-tying episodes (SWI median ± IQR: DKT novices = 1.21 ± 0.01, trainees = 1.54 ± 0.07, experts = 1.88 ± 0.07, experts > trainees and novices MWU p<0.001, trainees > novices MWU p<0.001; SKT novices = 1.67 ± 0.07, trainees = 1.55 ± 0.06, experts = 2.10 ± 0.17, experts > trainees and novices MWU p<0.001, novices > trainees MWU p<0.001). However, in the needle-drive phase, interestingly trainees were observed to generate network topologies that were smaller world than those of both experts and novices (SWI median ± IQR: novices = 1.46 ± 0.03, trainees = 2.75 ± 0.37, experts = 1.71 ± 0.08, trainees > experts and novices MWU p<0.001, experts > novices MWU p<0.001).

Task	**Expert [E]** Median ± IQR	**Trainees [T]** Median ± IQR	**Novices [N]** Median ± IQR	***p* values**
Overall	1.87 ± 0.26	1.68 ± 0.97	1.46 ± 0.41	E > T > N **p<0.001**
Needle Drive Insertion [NDI]	1.71 ± 0.08	2.75 ± 0.37	1.46 ± 0.03	T > E > N **p<0.001**
Double Throw Knot [DTK]	1.88 ± 0.07	1.54 ± 0.07	1.21 ± 0.01	E > T > N **p<0.001**
Single Throw Knots [STK]	2.10 ± 0.17	1.55 ± 0.06	1.67 ± 0.07	E > N > T **p<0.001**

Table 5.6. Comparisons of group networks by Small World Index according to expertise, LICS task and sub-tasks.

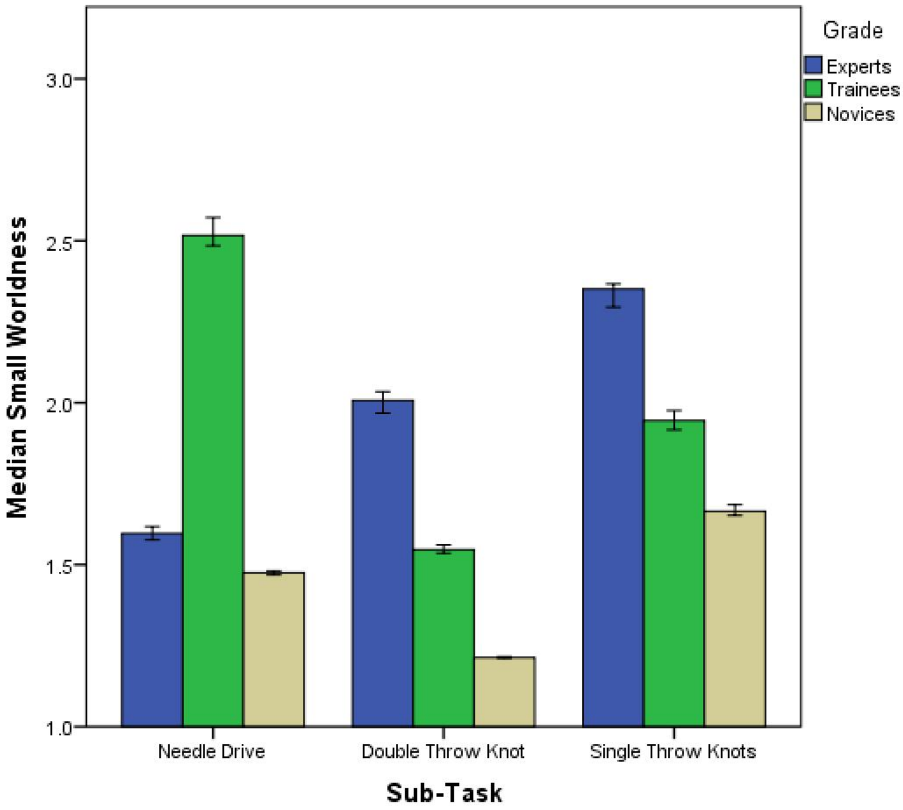

Figure 5.14. Graphical representation of median network's Small World Index (SWI) according to expertise for each sub-task.

5.4. Discussion

In this study functional neuroimaging was employed to investigate expertise-related disparities in recruitment of cortical substrates and changes in cortical networks within the entire frontal lobe that are likely to be vital for complex laparoscopic skills acquisition. Mental workload reduced whilst technical performance improved in line with operator experience and brain function varied according to surgical expertise. The primary observation is that cortical network architectures became progressively more small-world in line with increasing surgical technical skill. Conceivably, with practice, surgeons learn to retain only vital long-range cortical connections whilst simultaneously tuning local connectedness; wiring specifically designed to maximise neural efficiency[338]. Contrary to the hypothesis and the findings of a similar study evaluating surgeons brain function during LICS[29], novices do appear to activate the PFC to a greater extent than trainees and experts, a finding commensurate with motor learning theory[42,279] and studies of open surgical skills acquisition[259,260]. Despite this, experts may still recruit centres of executive control during episodes of high mental workload; indeed the magnitude of PFC recruitment amongst expert surgeons appears to depend on the degree of task complexity. One advantage of entire frontal lobe neuro-monitoring is the capacity for assessment of brain excitation beyond isolated evaluation of the PFC[29,259-261,326]. Indeed, trainees and experts appear to recruit frontal motor regions (PMC, SMA & M1) to a greater extent than novice laparoscopists, compatible with the notion that they have progressed beyond the "cognitive" phase of motor learning[18,25,42,246,277,325].

It is perhaps unsurprising that experts outperform trainees and novices when objectively assessed on LICS. However, the current study provides further data to reinforce the construct validity of the FLS assessments and adds to a wider body of literature regarding objective assessment of technical laparoscopic skills per se[320,324]. However, as previously discussed, scoring systems such as FLS provide a one-dimensional assessment of technical skill and are not designed to evaluate attention and concentration or the amount of neuronal resources invested in performance which arguably better reflect the degree of task internalisation that defines the extent of motor learning. Therefore, surgeons with similar technical performance scores may differ significantly with respect to task automation and hence capacity for multitasking and attention sharing. For example, in a study by Stefanidis and colleagues[80] whilst trainees and experts were indistinguishable on the basis of LICS technical skills, significant differences in their ability to perform a secondary attentional demanding task was exposed[80]. Whilst the introduction of secondary tasks may provide clues to the extent of task internalisation and phase of motor learning, the cognitive burden of the primary technical task can be provided by either indirect (e.g. NASA TLX) or direct assessments of mental workload (e.g. brain function).

Experts develop automaticity with deliberate practice presumably liberating mental resources to deal with new or unfamiliar tasks, a surrogate of which is amelioration in mental workload[83]. Commensurate with this theory, novices found performing LICS subtasks significantly more burdensome than expert surgeons. Interestingly, despite significant differences in FLS performance, novice and trainees were indistinguishable on the basis of differences in workload scales. Therefore, whilst the NASA TLX is a useful instrument for workload evaluation, between-group differences in workload may not be sufficiently construct valid for assessment on LICS, and it may be that performance gains made early in practice are independent of changes in task-evoked cognitive burden. Interestingly, other MIS studies have failed to demonstrate significant differences in mental workload scales between surgeons with significantly different technical performance[339]. Other limitations of the NASA TLX include context effects (ratings influenced by environmental effects prior to rating), range or anchor effects (rator tailors of use of scales to experience in experimental environment) as well as being labour intensive to administer[332]. Despite these limitations there is evidence to suggest that increased task familiarity as a result of practice is associated with both an alleviation of the cognitive burden as indexed from workload scales and attenuation in PFC excitation[340].

The PFC located in the anterior frontal lobe is an area of executive control that plays a vital role during the early 'cognitive' phase of learning when novice performance is unskilled, inconsistent, highly dependent on sensory feedback and effortful[279]. During early stages of skills acquisition it serves as a scaffolding framework to help cope with novel task demands[279]. Prefrontal activity is observed to attenuate with practice often associated with performance improvements[259,277,312,325]. Indeed, unlike novice surgeons who activate the PFC during open surgical knot-tying[326] a relative prefrontal redundancy was identified in experts. In this context, there is a misalignment between the parsimonious observed PFC activation and the anticipated prefrontal response amongst novice laparoscopists. One theory accounting for this low-level of PFC activity is the complexity of LICS is such that a limited 2-hour training session is simply insufficient for subjects to enter the "cognitive phase" of learning. Izzetoglu et al[340] observed that PFC oxygenation initially rose linearly with increasing task difficulty until a maxima where overwhelmed participants no longer engaged, reflected in a drop in PFC oxygenation. Conceivably, novices overwhelmed by the difficulty of LICS struggled to adequately engage the PFC. Perhaps, as has been observed by Iacoboni et al[341] further time in practice is necessary to facilitate development of a cognitive strategy leading to detectable PFC excitation.

Although novices displayed seemingly parsimonious channel based PFC activation, the magnitude of PFC oxygenation change was significantly greater than that of experts. This is contradictory to a similar OT study[29], which failed to significantly differentiate between PFC oxygenation of novices from experts. Only when novices were provided more practice did they display significantly higher PFC oxygenation change than experts. Unlike Ohuchida et al[29] the findings of the current study were congruent with motor learning literature[25,245,246,277,312]. These differences are possibly attributed to a difference in OT equipment, the study's experimental design where participants performed LICS as per the FLS curriculum and were asked to focus on the quality of the task instead of being provided a fixed time period where performance was assessed by the number of knots performed, which may lead to increased temporal demand.

Unlike open hand knot-tying[259] demonstrable differences in technical skill of LICS was observed between trainees and experts, which was reflected in significantly greater PFC oxygenation amongst trainees. Motor learning behaviour and associated motor performance is task specific[25]. Open hand knot tying is comparatively easier, well rehearsed and overlearned compared to LICS, which permits trainees to rapidly progress to autonomous stages whereby like experts they no longer need to engage the PFC. Interestingly, the presence of PFC activation during performance of the most burdensome phase the *double throw knot* may signify that certain aspects of this manoeuvre require far longer periods of practice, similar to complex motor sequences performed by musicians[25]. Alternatively, PFC guided selective focussed attention during this critical phase may be vital for successful execution as it necessitates very precise bimanual movements.

Unsurprisingly, trainees and experts accordant to their stage of skills acquisition engaged secondary motor regions (PMC, SMA and M1) greater than novices based on number of activated channels and magnitude in activation on both haemoglobin species (HbO_2 and HHb), which lends support to the secondary hypothesis of this study. Experts who assumedly with extended practice had begun to exhibit a shift in engagement from anterior regions associated with early learning to posterior and subcortical regions appropriately exhibited less activation than trainees in the PMC. The PMC is recruited for retrieval of learned motor sequences at early stages[25,246,312,325]. SMA activation is observed at late stages of motor learning[25,245,246,277,312,325] where it too is involved in generation of learned sequences but unlike PMC, it is implicitly driven. Lefebvre et al[342] highlighted the importance of engagement of the SMA, which was observed amongst individuals who were better performers and had made early performance gains[342]. Similarly, as observed by Karni et al[317] trained participants (experts and trainees) displayed greater and a wider map of activation in the M1 than novices. Finally, Puttemans et al[312] through 10,500 cycles of practice uniquely tracked learning-related neuroplasticity extensively beyond motor performance reaching asymptotic levels. In this longitudinal study of bimanual motor learning, at different stages down-regulation of all cortical motor learning centres were observed whereas attenuation of SMA and M1 was observed at late stages when participants continued to practice beyond asymptotic levels. Shifts in activation foci and up-regulation of activation were only observed in sub-cortical structures namely the putamen and the anterior cerebellum at the final autonomous stage of learning[312]. In this regard, attenuated SMA and M1 activation observed in experts suggests a shift from cortical to subcortical activity related to over practice.

5.4.1. Sub-task Differences in Cortical Activation

Extensive research on bimanual movements has revealed that humans show a basic tendency to symmetrical movements (*in-phase*) as opposed to directionally different asymmetric movements (*anti-phase*) [343,344]. Allocentric constraints in human motor behaviour result in *'anti-phase'* movements being less stable, more difficult to perform, requiring greater attention and practice [330]. Recent evidence suggests a distributed brain network of fronto-parietal regions involved in bimanual coordination [330,345-348]. The added value of LS sub-tasks selected here was to expose potential differences in the neural correlates underlying differences in bimanual co-ordination according to demands. Expert surgeons provide a rare opportunity to explore patterns of bimanual coordination that are both exceptional and relatively free from learning confounds. Despite displaying significantly lower PFC excitation across all tasks, upon comparison of sub-task response *'anti-phase'* knot-tying manoeuvres led to significantly greater PFC oxygenation than the *'in-phase'* NDI task. These results suggest the need for greater demands on attention and action selection during knot-tying episodes relative to needle insertion during suturing. However, unlike other studies implicating the pre-SMA and SMA in *'anti-phase'* or incongruent bimanual co-ordination [291,349-351] greater SMA activation during knot-tying relative to needle drive was not observed in the current study. Finally, comparatively lower M1 excitation during knot-tying is commensurate with other fNIRS studies suggesting that complex uni-manual tasks burden M1 to a greater extent than bimanual tasks [352]. Here, needle insertion is best regarded as a set of complex uni-manual manoeuvres with the dominant hand adopting a manipulative and the non-dominant a stabilizing role.

5.4.2. Expertise related differences in Small World Index

Connectivity measures (inter-regional activity) have been found to be more sensitive than regional cortical activation when evaluating neuroplasticity associated with long-term training[220]. The final hypothesis underpinning this thesis chapter was that with increasing expertise the cortical network of the motor learning regions would evolve into an architecture characterised by a shorter path length (indicative of greater global efficiency of network), higher clustering (robustness to random error) and more modular structure (presence of sub-systems or neighbourhoods)[328], all of which ultimately increase the network SWI. All groups exhibited networks with SWI of greater than 1, suggesting LICS subtasks result in brain networks with small world topology regardless of expertise. However, significantly greater small world architectures were observed in experts compared to trainees who in turn exhibited greater small worldness than novices.

Many real world complex networks possess small-world topology which is characterised by dense local clustering between neighbouring nodes to form neighbourhoods, yet possess a shorter path length between any pair of nodes due to the presence of few long range connections (edges)[338]. In a seminal paper, Watts et al[327] found that most real world complex networks either biological (Neural network of the worm *C.elegans)*, technological (power grid of the United States) and social network (collaboration graph of film actors) were shown to be small world-networks. In order to maximize information transfer (efficiency) at a low wiring cost due to the finite space in which the human brain is contained, it is found to possess small world architecture at both a macro and micro scale[338]. Langer et al[329] not only observed similarly a greater SWI in higher performing individuals on a working memory task but small worldness also improved with training.

Cortical network analysis has recently been applied in longitudinal motor learning studies to evaluate training regimens[353], type of feedback[354] and assistive technology[253]. Taubert et al[355] demonstrated a tight correlation between functional connectivity and actual structural changes in the grey matter of SMA and PFC in a longitudinal motor learning task. Variations in the technical prowess of expert surgeons performing advanced MIS have recently been exposed which has a direct impact on their operative outcomes[4]. The magnitude of activation in motor centres of the brain and the strength of connectivity between these regions such as SWI could potentially bolster surrogate objective markers of expertise in credentialing of expert surgeons based on efficient neural behaviour. Future work to strengthen our observations would focus on longitudinal studies of procedural motor learning of LICS over multiple sessions. The type of feedback (positive, negative and neutral) also has differential effects on procedural learning in terms of performance gains and retention[356]. It would be desirable to investigate the underlying neuroplasticity associated with each form of feedback, which could guide structuring of current training methods for complex laparoscopic skills acquisition either in the controlled laboratory or live surgical environment.

5.4.3. Limitations of the study

The total sample size (n=32) and the number of experts (n=9), represents a large OT dataset investigating surgical skill. Limited availability of experts who perform enough procedures involving LICS prevented recruitment of more experts. Indeed, only bariatric or upper gastrointestinal surgeons who most often perform this manoeuvre were eligible. To facilitate comparison between this group and novices after limited training, the LICS task was assessed in sub-phases since a discrepancy in the time taken for completion between novices and experts was predicted. In the process of segmenting the task and analysing both haemoglobin species a plethora of results (see Table 5.4) was obtained which required summary. A variable "*combined LICS*" was created, which represents the average haemodynamic change for all sub-tasks within a group to help summarize our sub-task results. We recognize that if LICS was performed without segmentation the observed brain behaviour may not approximate the results of "*combined LICS*". Network econometrics such as characteristic path length, number of edges and modularity could not be compared for statistical significance as were limited by the generation of nine group-averaged brain networks representative of each sub-task. Finally, OT cannot quantify deeper brain regions that are associated with motor learning such as the basal ganglia and cerebellum. On the other hand DOT is advantageous as it neither restricts surgeons from performing technical manoeuvres in unrealistic positions[265,266] i.e. lying supine nor leads to exposure to ionising radiation[266].

Chapter 6

Effects of Practice on Frontal Lobe Neuroplasticity in MIS: A Longitudinal Study

Contents from this chapter have been published in:

Shetty K, Leff DR, Orihuela-Espina F, Yang GZ, Darzi A. Persistent Attentional Demands Despite Laparoscopic Skills Acquisition. *JAMA Surgery* 2016 Jul 1; 151(7): 682-684.

Crewther B, Shetty K, Jarchi D, Selvadurai S, Cook CJ, Leff DR, Darzi A, Yang GZ. Skill acquisition and stress adaptations following laparoscopic surgery training and detraining in novice surgeons. *Surgical Endoscopy* 2016 Jul; 30(7): 2961-2968.

6.1. Introduction

Practice dependent neuroplasticity describes changes in brain activation or network connectivity that occur with improved speed and task accuracy as a result of motor skills training, and which can be mapped to traditional models of learning [25]. Procedural training and expertise development is associated with neuroplasticity in primary and secondary motor regions in the brain [25,245,246,277,306]. Regarding frontal lobe plasticity, empirical data suggests an *"associative/premotor"* network including the dorsolateral prefrontal cortex (DLFPFC) and dorsal premotor (PMC) which supports novel demands early in practice [311,312], and a *"sensorimotor"* network that involves the supplementary motor area (SMA) and primary motor cortex (M1), recruited in advanced stages of training [299,307,325,357].

Surgeons represent a unique population in whom to study such learning-related neuroplasticity due to the intensity of training, and the precision and complexity of the bimanual motor skills they need to assimilate. These skills comprise bimanual sequence learning under error-based paradigms requiring adaptation to the external environment and inter-limb co-ordination [245,281,358-360]. Critically, functional neuroimaging provides objective data to inform the stage of learning for trainee assessment; technical aptitude for trainee selection [259,313] and enables evaluation of strategies geared to enhance learning or re-learning [255], especially important in high-risk industries such as healthcare.

Studying ambulant subjects performing complex tasks with ferromagnetic equipment imposes experimental challenges for which traditional scanning environments such as functional magnetic resonance imaging (fMRI) are ill suited. Investigators have therefore exploited recent technical developments in electroencephalography (EEG)[267,270,272] and multi-channel functional near infrared spectroscopy (fNIRS)[253,254,258,259] to study motor skills acquisition, implicit and explicit skills learning schedules [272], and neural feedback to improve motor training in surgeons[255]. For certain motor skills such as bimanual knot-tying, these studies demonstrate greater prefrontal responses amongst novice versus experts [259,361] and prefrontal attenuation following a week of practice [260]. However, the time course of shifts in activation foci that support learning on more complex visual spatial tasks such as minimially invasive surgery (laparoscopic) remain to be fully investigated.

Amongst operative manoeuvres, laparoscopic suturing (LICS) is putatively the most challenging motor skill to acquire [114,319,362,363]. LICS represents a complex bimanual task requiring the operator to overcome paradoxical movements due to the fulcrum effect under restricted kinematics (degrees of freedom), execute a precise sequence movements using long rigid instruments via miniature incisions that amplify hand tremors, besides imposing sensorimotor disruptions such as the distortion of tactile feedback, movement adaptation according to the visual scalar difference and the loss of depth perception [319].

Contrary to published literature implying an expertise-associated gradient in prefrontal responses [259,260,312,325,361], experts and novices could not be differentiated based on prefrontal responses during LICS, and further training led to incremental PFC engagement and not attenuation [29]. Importantly, extended training was not provided, practice associated prefrontal attenuation was not observed, and secondary motor areas such as the SMA which may better characterise differences in cognitive processing between expert and novices were not studied [29,364].

In the current study, a longitudinal experiment was designed to interrogate frontal lobe plasticity associated with LICS acquisition using fNIRS. Based on established motor learning literature and the results of the cross-sectional study (see Sec 5.3 of Chapter 5) our hypotheses were as follows: (i) attenuation in PFC activation, (ii) an initial increase followed by a decrease in PMC and SMA activation, (iii) an initial reduction followed by a late increase in M1 activation and (iv) sustained increases in PMC and SMA activation upon retention testing due to *off-line* learning gain. A fortnight of deliberate practice is expected to lead to performance improvement and redistribution of cortical activation in line with motor learning hypotheses (see figure 6.1).

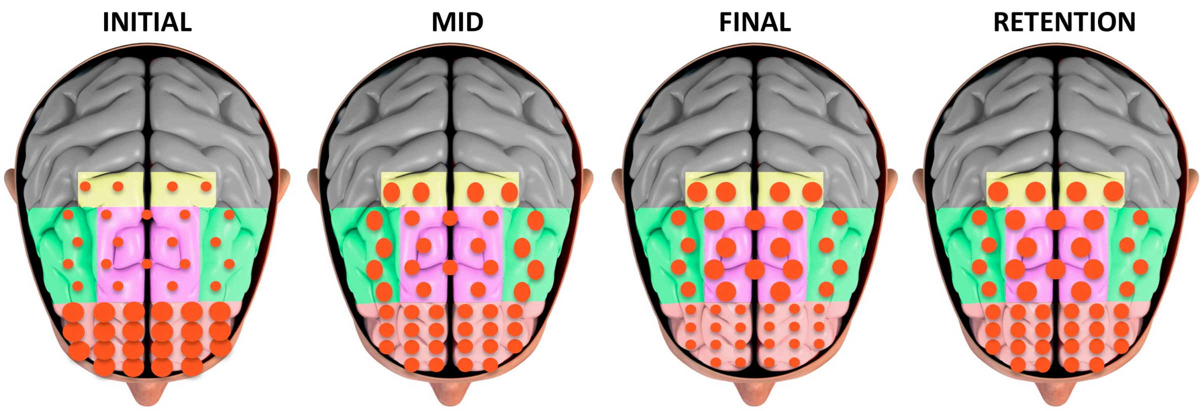

Figure. 6.1. Research hypotheses depict functional redistribution of cortical activation within the frontal lobe in relation to training on laparoscopic intracorporeal suturing (LICS). The functional regions interrogated within the frontal lobe are the prefrontal cortex (pink), premotor cortex (green), supplementary motor area (purple) and primary motor cortex (yellow). The size of the red circles in each of the aforementioned regions represents the intensity of the anticipated cortical activation commensurate with the length of training received.

6.2. Materials and Methods

6.2.1. Experimental and Task Paradigm

LICS is a serial motor skill requiring complex movement sequences to achieve stitching and knot-tying within a confined compartment. The operator uses long rigid instruments under multi-sensory adaptation to perform a motor task in a confined (see figure 6.2). Here, LICS was deconstructed into three sub-tasks (see figure 5.5 in Chapter 5) namely needle drive insertion (NDI), double throw knot (DTK) and a pair of single throw knots (STK) [319,331]. Task deconstruction is necessary for a complex serial skill acquisition since it facilitates part-practice [52], and allows identification of rate limiting sub-tasks for LICS acquisition.

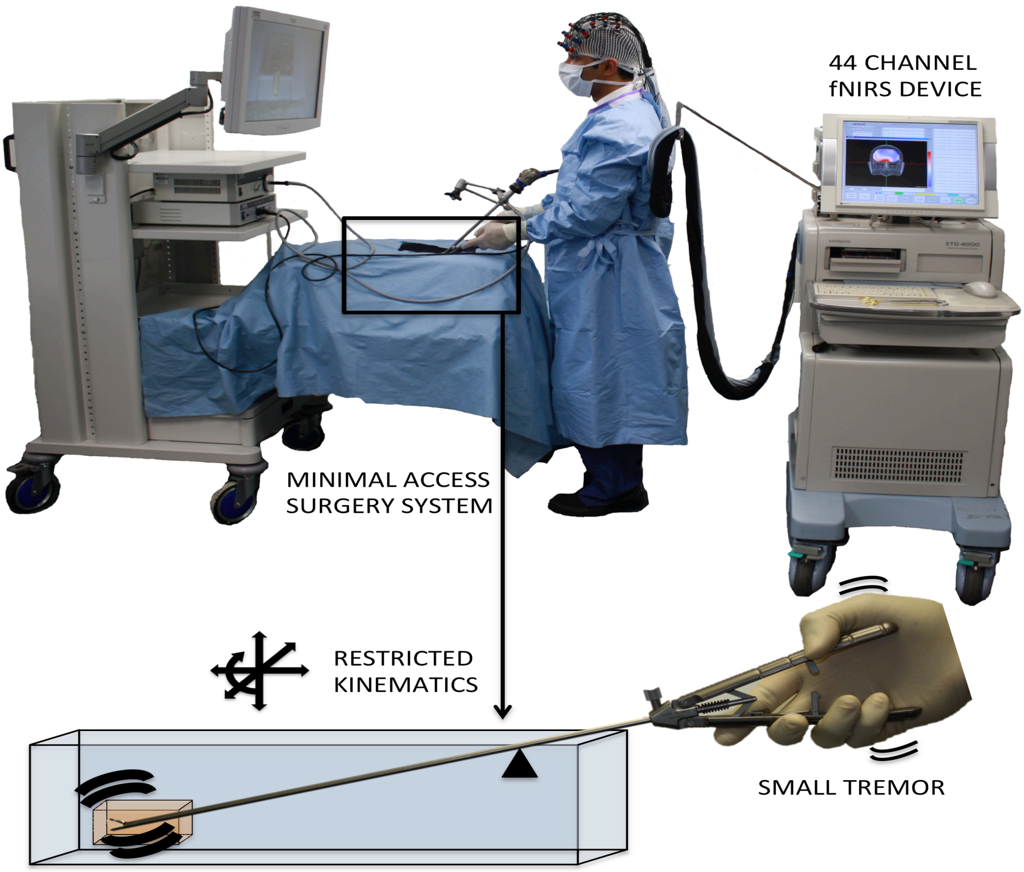

Figure 6.2. Experimental set-up illustrating a surgeon performing laparoscopic intracorporeal suturing (LICS) using hand held laparoscopic needle holders within a box trainer under two-dimensional visual feedback. Cortical haemodynamic data is captured using a 44-channel fNIRS device (ETG-4000, Hitachi Medical Corp., Japan). The magnified workspace depicts the "fulcrum effect" at the entry point into the body (blue box), which results in paradoxical movements of the operator's hand and instrument tip. Fine hand gestures and tremors are amplified at the effector's tip, within the confined workspace (orange box). Instrument movement is restricted to four degrees of freedom.

Methodologically, task deconstruction also aimed to reduce the variance in task durations (i.e. between novice and experienced operators) and facilitate assessment of the evoked brain response by each sub-task. In the first sub-task NDI (see Figure 5.5 in Chapter 5) the needle is passed through pre-marked entry and exit points of a slit fashioned on a penrose drain. DTK involves at first synchronised precise bimanual rotatory to enable twice looping the long end of suture around the needle-holder. The short end of suture is then grasped and the knot tied through opposing forces on the suture, which apposes the wound edges. The final sub-task STK requires formation of a single loop, tying a knot and repeating this sequence by inverting the hand gestures. DTK is noticeably more exigent because the property of the suture makes it prone to uncoiling, meaning that looping the suture twice mandates a greater degree of stable bimanual movement control.

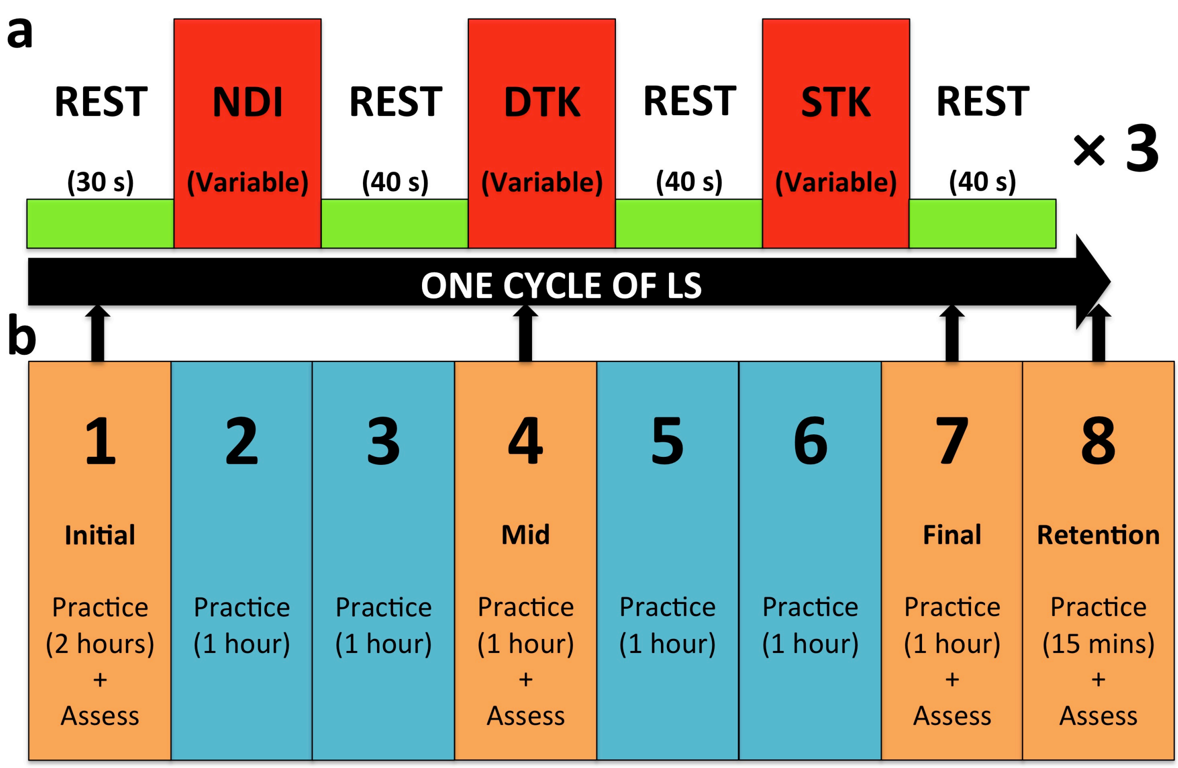

Figure 6.3. Schematic illustration of (a) block design of the experiment composed of alternating fixed episodes of rest periods (30 seconds) and sub-task periods (variable duration). LICS was performed thrice along this paradigm. (b) Illustrates the experimental design of the study that involved multiple practice sessions (blue), and combined practice with neuroimaging assessment sessions (orange).

A block design experimental paradigm was conducted, consisting of baseline rest (30s), alternating blocks sub-task execution (variable duration) and post task motor rest (40s). During rest periods, participants were instructed to view the centre of the screen. Experiments were conducted in a controlled laboratory bereft of natural light and distracting auditory stimuli (see Figure 6.2). Adopting LICS task description from the Fundamentals of Laparoscopic Surgery curricula[26], all participants were required to perform three cycles of LICS in a box trainer [iSurgicals, UK]. The participants performed LICS using a pair of laparoscopic needle holders (Karl Storz GmBH & Co, Tuttlingen, Germany) and a fixed length of 15 centimetres of 2/0 Polysorb suture material (Covidien Ltd, Ireland).

6.2.2. Participants and Practice Schedule

Local Research Ethics Committee (project number: 05/Q0403/142) approval was obtained. 13 right-handed, task naive, males medical students were recruited from Imperial College London (mean age ± S.D. = 21.6 ± 0.8 years). All participants were screened for neuropsychiatric illnesses and handedness [365]. All participants withheld from caffeine or alcohol consumption for 24 hours prior to neuroimaging assessments given the influence on regional cerebral haemodynamics [199]. Participants were consented for a total of eight sessions including LICS practice and neuroimaging sessions as outlined (see Figure 6.3). During the first session referred to as *"initial session",* novices received two hours of task training from a single trained, certified proctor (author KS). Training comprised a ten-minute instructional video, followed by one to one instruction for two hours. At the end of training, a minimal level of competency i.e. execution of one cycle of LICS within 600 seconds was ascertained. This threshold was both logistical and methodological. A pilot study determined that 2 hours was the duration required to train 80% (12/15 participants) to reach basic proficiency [228] and it was impractical to study subjects who took too long to execute one cycle of LICS. Similarly, if during the experiment a participant required more than 300 seconds to perform one LICS sub-task, data collection was terminated. Competent novice participants progressed to the block designed experiment (see Figure 6.3) as described in Section 6.2.1. Novices then received six further one-hour sessions of deliberate practice. Neuroimaging assessments were conducted after one hour of practice of LICS, and following the fourth ("mid") and seventh sessions ("final") respectively. The interval between each of the first seven sessions was two days. Skills retention (R) was assessed four weeks after the final session. Prior to retention testing, only 15 minutes of task familiarisation was permitted, akin to that allowed for trainees and experts in the cross-sectional study. At each assessment session (initial, mid, final and retention) participants were assessed in an identical manner to the cross-sectional experiment in which they were neuro-monitored during three cycles of LICS.

6.2.3. Behavioural Performance

A single assessor 'blinded' to the identity of participants retrospectively assessed performance. Task completion times (seconds), end-product analysis of the sutured drain and knot quality were incorporated into a single performance score. The validated formula adopted accounted for time and accuracy [FLS Score = 600 - (time in seconds) - (penalties × 10)] [27]. Penalties were levied on inaccuracies namely distance (mm) from entry and exit points marked on the Penrose drain, distance (mm) between the cut edges of the drain and knot quality. At every assessment point, median FLS scores in arbitrary units (A.U.) were calculated across three LICS procedures for each subject.

Subjective mental effort was measured at the end of each assessment using the NASA Task Load Index (NASA-TLX) questionnaire [81]. This scale has been used to evaluate the demands on operators' attention during surgery [83] or in other high-risk fields [314]. Here the instrument was used to identify amelioration of subjective mental effort with training or expertise, and stage-specific sub-task difficulty. Between-group FLS and NASA-TLX scores were compared using non-parametric statistical tests of significance (IBM SPSS v22; SPSS Inc., Chicago, IL, USA). Inter session data was compared using the Friedman test (statistical threshold $p<0.05$) for multiple comparisons and the Wilcoxon Sign Rank test for paired comparisons (Bonferroni adjustment = $p<0.008$).

6.2.4. Functional Near Infra-Red Spectroscopy

A 44-channel fNIRS device (ETG-4000; Hitachi Medical Corp, Tokyo, Japan) was used to acquire functional neuroimaging data. Optodes comprising emitters and detectors were placed three centimetres apart and held in two thermoplastic 3 × 5 array holders. The International 10-10 system was employed in tandem to fixed reference points on the subjects scalp (i.e. nasion, inion, vertex and bilateral external auditory meati) to locate ten inferred scalp markers [200]. The fixed reference points and all optode positions were registered using a 3D digitizer (Isotrak, Polhemus Inc, Vermont, US) to provide locations of each intended channel position [333]. 22 channels were situated over the PFC, 8 channels over the PMC, 10 channels over the SMA and 4 channels over M1. Employing 3D Composite Display Unit (Hitachi Medical Corp, Tokyo, Japan), positional data of a representative case was projected onto an individual MRI to visualize and ratify frontal lobe regions underlying each channel (see Figure 6.4).

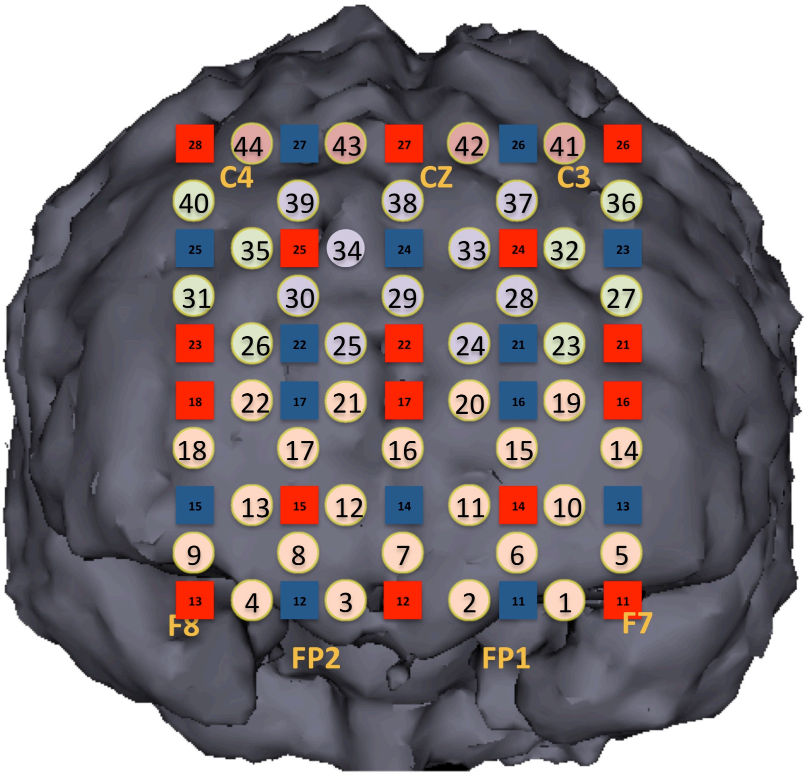

Figure 6.4. Channel locations (1-44) were obtained by transferring topographical data to a 3D cortical surface of a T1-weighted MRI image. Relative International 10-10 cranial markers are illustrated (orange). 22 channels (channel number 1-22) were positioned over the Prefrontal Cortex (pink) and 22 channels were positioned over the dorsal Premotor Cortex (green), Supplementary Motor Area (lilac) and Primary Motor Cortex (red) (channel number 23-44). The approximate position of each emitter (red squares) and detector (blue squares) are shown in relation to 10-10 surface markers.

6.2.5. Cortical Haemodynamic Data Processing

Bespoke MATLAB based software named Imperial College Neuroimage Analysis (ICNA) [334] was employed to process fNIRS data. Relative changes in light data were transformed to changes in HbO_2 and HHb utilising the modified Beer-Lambert law. Derived haemoglobin data was subsequently down sampled from 10Hz to 1Hz and linearly detrended to reduced physiological noise and systemic drift respectively. Finally, integrity checks were conducted to eliminate unreliable data due to saturation artefact or abrupt optode displacement [199]. To overcome temporal variance in data sampled from sub-task periods, data from each OT channel was resampled. Resampled rest data comprised 10 samples from each rest period (baseline and recovery) captured 15 seconds before sub-task onset to allow for the recovery of cerebral haemodynamics. Resampled task data comprised 55 samples from each sub-task period starting six seconds after task onset to allow inertia in task-evoked changes in cerebral haemodynamics. For each LICS sub-task, average Hb data was generated by block averaging resampled task data.

A new variable ΔHb was created for each given haemoglobin species (HbO_2 and HHb), to estimate the magnitude of task-evoked cortical activation. For each given channel and haemoglobin species, ΔHb was calculated by subtracting rest haemoglobin data from task haemoglobin data (i.e. ΔHb = ΔHbTask – ΔHbRest). Regional changes in cerebral blood flow (ROIΔHb) for each region of interest (ROI) namely PFC, PMC, SMA and M1 was calculated by averaging ΔHb across all channels located in that ROI for a given sub-task. ROIΔHb was compared between training sessions using Friedman (Fr) tests. If statistically significant ($p<0.05$), a post-hoc paired comparison was performed using Wilcoxon Sign rank (WSR) tests respectively. To account for the total number of potential comparisons (x 92) experiments, Bonferroni correction was applied, adjusting the threshold for significance to $p<0.0005$ (0.05/92).

6.2.6. Brain-Behavioural Performance Correlation Analysis

The relationship between longitudinal changes in PFC cortical haemodynamics and behavioural improvements in technical performance was investigated using correlation analysis. A crude fixed-effect model (Hb ~ 1 + Session + FLS score + Session:FLS score) was created to investigate the possibility of a linear correlation between PFC ΔHbO_2 and FLS performance.

6.2.7. Random Effects Model for Cortical Data

Clustering of ΔHb data was anticipated due to inter-subject variability associated with longitudinal studies [366]. Therefore, for each Hb species each (ΔHbO_2 and ΔHHb) ΔHb data was incorporated into a random effects model (Intercooled Stata. v13, Stata Corporation, Texas, USA) to interrogate the effect of practice session (i.e. initial, mid, final and retention) on task-induced change regional in cortical haemodynamics across frontal ROIs.

6.2.8. Heart Rate Variability and Stress

The effect of stress on the autonomic nervous system can induce changes in cerebral blood flow [232]. Task induced stress may influence the autonomic system measured as a change in heart rate (HR) and/or derived heart rate variability (HRV). To quantify stress-induced changes in systemic haemodynamics, a portable electrocardiogram (Bioharness v2.3.0.5; Zephyr Technology Limited, Annapolis, USA) was fastened over each participant's chest to capture HR and derive HRV data. R waves were detected using bespoke MATLAB software and then visually inspected and verified by two investigators. Capturing the R-R intervals and R-R standard deviation (SD_{RR}), enabled mean HR and HRV to be derived. An increase in HR and decrease in HRV (SD_{RR}) represent an increase in stress [367]. Between-group comparisons in HR and HRV data was performed using the one-way analysis of variance (ANOVA) test ($p<0.05$).

6.3. Results

6.3.1. Participants, Exclusions and Data Integrity

12 of 13 participants completed all eight scheduled sessions. One participant withdrew due to an inability to adhere to sequential training sessions. Following data integrity checks [199] a number of channels were excluded due to artefacts or noise [457 out of a total of 2068 (22.09%)].

	Initial (I)	Mid (M)	Final (F)	Retention (R)	*p* values Friedman	*p* values Wilcoxon Sign Rank
FLS (a.u.)	302± 128	421 ± 86	471 ± 40	468 ± 56	**<0.001**	F & R > M > I **(0.001)** F > R (0.808)
NASA-TLX (a.u) Overall	59.5 ± 19	57 ± 29	43 ± 30	42 ± 44	**<0.001**	I > F **(0.001)** I > R **(0.002)** M > F **(0.001)** M > R **(0.002)**
NASA-TLX (a.u) NDI	53 ± 20	58 ± 35	51 ± 33	39 ± 34	0.407	Nil significant
NASA-TLX (a.u) DTK	66 ± 23	50 ± 25	37 ± 27	43.5 ± 41	**0.005**	I > F **(0.005)**
NASA-TLX (a.u) STK	62 ± 13	61 ± 31	30 ± 31	44 ± 47	**0.010**	I > F (0.010) I > R (0.037) M > F (0.013)

Table 6.1. Technical performance (FLS) and subjective cognitive load (NASA-TLX) data for each practice session. Group comparisons were analysed using Friedman and Wilcoxon Sign Rank test. Significant p values are highlighted in bold.

6.3.2. Behavioural data (FLS score)

FLS performance data are summarised (see Table 6.1), and illustrated (see Figure 6.5). Technical performance improved with practice (median FLS scores ± IQR: initial session = 302 ± 128, mid session = 421 ± 86, final session = 471 ± 40, retention session = 468 ± 56, Fr: $p<0.001$). Significant improvements were observed from initial to mid (WSR: $p<0.001$) and mid to final sessions (WSR: $p<0.001$). On retention testing no significant deterioration in performance was observed (final session = 471 ± 40, retention session = 468 ± 56; WSR: $p=0.808$).

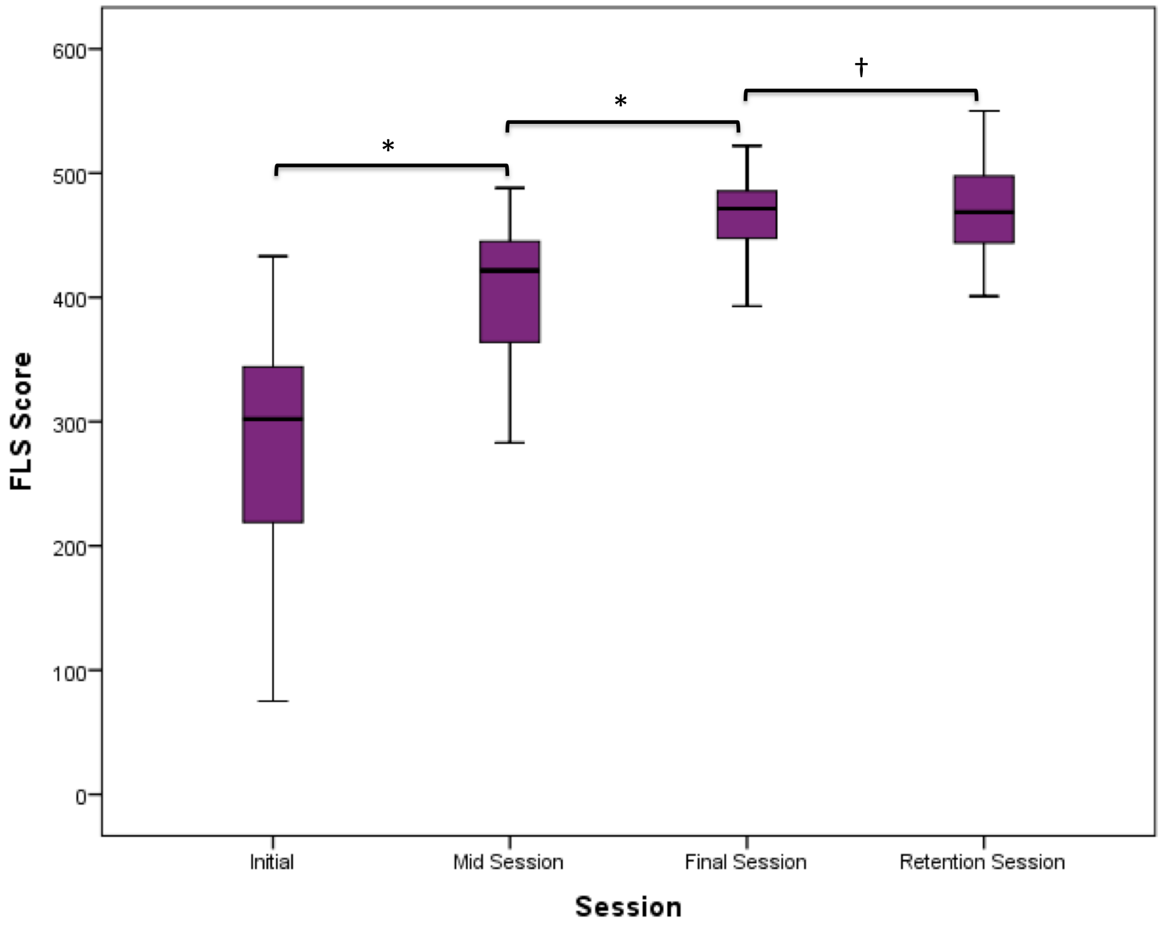

Figure 6.5. FLS scores ranging from 0 to 600 (A.U.). Box plots represent technical performance based on technical progress through points of practice. * denotes a statistically significant paired comparison ($p \leq 0.001$ on WSR test) and † denotes a non-statistically comparison ($p<0.05$).

6.3.3. Behavioural data, Subjective Assessment of Cognitive Workload (NASA-TLX)

Practice-related attenuation in subjective workload was observed for sub-tasks DTK (NASA TLX scores median ± IQR: initial session = 66 ± 23, mid session = 50 ± 25, final session = 37 ± 27, retention session = 43.5 ± 41, Fr: $p<0.05$) and STK (NASA TLX scores median ± IQR: initial session = 62 ± 13, mid session = 61.5 ± 31, final session = 30 ± 31, retention session = 44 ± 47, Fr: $p<0.05$). However, post-hoc tests revealed significant workload alleviation for DTK only [initial vs final: $p=0.005$].

6.3.4. Heart Rate Variability and Stress

No between-group difference was found in mean HR across practice sessions [mean HR (beats / min) ± S.D: initial session = 91.08 ± 20.86, mid session = 85.02 ± 17.01, final session = 85.00 ± 19.94, retention session 93.05 ± 18.95, ANOVA: p =0.645).

6.3.5. Practice Related Differences in Frontal Lobe Cortical Activation

Longitudinal changes and trends in ROI ΔHb data by subtask are summarised (see Table 6.2). Comprehensive analysis with data for changes in ROI ΔHb of each Hb species are presented in Table 6.3.

ROI	Sub-task	Initial – Mid		Initial – Final		Initial – Retention		Mid – Final		Mid-Retention		Final – Retention	
		HbO_2	HHb	HbO_2	HHb	HbO_2	HHb	HbO_2	HHb	HbO_2	HHb	HbO_2	HHb
PFC	NDI	↔	↔	↓*	↔	↔	↔	↔	↔	↔	↔	↑*	↔
	DTK	↔	↓*	↔	↔	↔	↔	↔	↔	↔	↓*	↔	↔
	STK	↓*	↔	↔	↔	↔	↔	↑*	↔	↑*	↔	↔	↔
PMC	NDI	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↑*	↔
	DTK	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔
	STK	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔
SMA	NDI	↔	↔	↔	↔	↔	↓*	↔	↔	↔	↔	↔	↔
	DTK	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔
	STK	↔	↔	↔	↔	↔	↔	↔	↓*	↔	↔	↔	↔
M1	NDI	↔	↔	↔	↓*	↔	↔	↔	↔	↔	↔	↔	↔
	DTK	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔	↔
	STK	↔	↔	↔	↔	↔	↔	↔	↓*	↔	↔	↔	↔

Table 6.2. Longitudinal variations in cortical haemodynamic change associated with learning LICS, from initial practice (2 hours) to mid practice (5 hours), through to final practice (8 hours), and finally from practice termination to retention testing four weeks later. Brain regions examined are the prefrontal cortex (PFC), dorsal premotor cortex (PMC), supplementary motor area (SMA) and primary motor cortex (M1). Increases and decreases in the magnitude of oxygenated haemoglobin (+ΔHbO_2) and deoxygenated haemoglobin (-ΔHHb) are correspondingly represented as ↑ and ↓. Significant trends (p<0.0005) are additionally denoted with *. ↔ Represents non-significant between-session differences in cortical haemodynamic change.

Hb Species, Sub-task	ROI	Initial µmol/cm Median ± IQR	Mid µmol/cm Median ± IQR	Final µmol/cm Median ± IQR	Retention µmol/cm Median ± IQR	*p* values Friedman	*p* values Wilcoxon Sign Rank					
							Initial vs Mid	Initial vs Final	Initial vs Retention	Mid vs Final	Mid vs Retention	Final vs Retention
HbO_2 NDI	PFC	12.55 ± 24.13	10.43 ± 19.88	6.92 ± 14.79	10.07 ± 18.97	**<0.001**	0.125	**<0.001**	0.936	0.112	0.006	**<0.001**
	PMC	8.34 ± 17.20	7.47 ± 18.92	3.78 ± 14.89	8.37 ± 14.22	**0.006**	0.692	0.002	0.848	0.007	0.733	**<0.001**
	SMA	8.23 ± 17.35	7.31 ± 13.89	4.20 ± 12.17	5.71 ± 13.28	**0.003**	0.759	0.001	0.089	0.001	0.015	0.353
	M1	13.39 ± 17.42	9.14 ± 16.84	4.43 ± 14.93	9.03 ± 17.37	**0.002**	0.575	**<0.001**	0.278	0.006	0.639	0.001
HHb NDI	PFC	0.54 ± 8.15	0.42 ± 7.42	0.59 ± 5.70	0.70 ± 7.88	0.276	0.048	0.929	0.271	0.002	0.144	0.284
	PMC	-0.76 ± 8.49	-2.39 ± 8.87	-1.49 ± 7.50	-1.91 ± 7.08	0.319	0.013	0.205	0.206	0.710	0.206	0.263
	SMA	-1.22 ± 7.54	-2.38 ± 6.51	-1.91 ± 6.40	-2.79 ± 7.04	**0.024**	0.049	0.015	0.008	0.724	0.258	0.031
	M1	-0.79 ± 8.14	-3.07 ± 8.54	-2.97 ± 8.43	-2.87 ± 8.03	**0.046**	0.627	0.399	0.018	0.777	0.440	0.580
HbO_2 DTK	PFC	8.08 ± 19.95	7.67 ± 21.12	7.77 ± 17.91	8.88 ± 19.92	**0.030**	0.098	0.727	0.787	0.007	0.004	0.988
	PMC	6.26 ± 19.33	5.73 ± 17.22	4.20 ± 15.54	5.24 ± 14.72	0.740	0.647	0.626	0.608	0.203	0.533	0.184
	SMA	4.84 ± 16.22	4.80 ± 14.71	3.77 ± 13.16	3.42 ± 11.34	0.558	0.085	0.547	0.467	0.673	0.150	0.416
	M1	6.30 ± 17.81	7.78 ± 15.69	4.27 ± 18.48	6.01 ± 16.17	**0.005**	0.001	0.702	0.978	0.298	0.030	0.719
HHb DTK	PFC	1.17 ± 7.17	0.49 ± 7.37	0.65 ± 5.22	1.41 ± 6.10	**<0.001**	**<0.001**	0.037	0.375	0.051	**<0.001**	0.001
	PMC	-0.51 ± 9.09	-1.30 ± 8.43	-0.68 ± 8.84	-0.62 ± 7.08	0.670	0.062	0.086	0.294	0.563	0.116	0.128
	SMA	-0.91 ± 6.91	-1.49 ± 6.62	-1.19 ± 6.56	-0.82 ± 6.05	0.749	0.510	0.150	0.439	0.347	0.820	0.854
	M1	-1.74 ± 4.14	-1.74 ± 7.22	-1.02 ± 7.98	-1.84 ± 5.74	0.406	0.612	0.149	0.507	0.010	0.136	0.060
HbO_2 STK	PFC	8.59 ± 19.21	3.86 ± 14.24	7.31 ± 12.09	8.77 ± 17.47	**<0.001**	**<0.001**	0.315	0.252	**<0.001**	**<0.001**	0.011
	PMC	6.49 ± 15.59	5.69 ± 14.22	5.27 ± 15.23	7.63 ± 14.63	0.235	0.181	0. 274	0.406	0.356	0.092	0.002
	SMA	7.15 ± 16.91	6.40 ± 16.99	5.14 ± 12.74	6.48 ± 11.99	0.767	0.083	0.278	0.834	0.294	0.905	0.608
	M1	10.26 ± 18.24	8.96 ± 15.22	6.70 ± 13.69	9.75 ± 13.64	0.937	0.877	0.511	0.372	0.451	0.266	0.075
HHb STK	PFC	0.83 ± 6.79	0.50 ± 7.13	0.93 ± 6.07	1.03 ± 5.70	0.188	0.019	0.086	0.984	0.116	0.014	0.433
	PMC	-0.26 ± 7.3	-1.27 ± 8.22	-0.76 ± 7.1	-0.67 ± 5.35	0.338	0.260	0.691	0.145	0.849	0.581	0.479
	SMA	-1.57 ± 6.3	-0.69 ± 6.61	-1.60 ± 6.60	-1.65 ± 6.15	0.055	0.181	0.865	0.719	0.008	0.032	0.479
	M1	-0.17 ± 7.40	0.02 ± 8.95	-3.42 ± 8.04	-0.83 ± 6.99	**0.009**	0.321	0.009	0.286	**<0.001**	0.591	0.014

Table 6.3. Illustrates the regional derived ΔROI Hb for HbO_2 and HHb species for each sub-task, namely: needle drive insertion (NDI), double throw knots (DTK) and single throw knots (STK). Regions interrogated were the prefrontal cortex (PFC), dorsal premotor cortex (PMC), supplementary motor area (SMA) and primary motor cortex (M1). Comparisons, which reached statistical significance, are represented in bold.

Prefrontal Cortex: PFC cortical haemodynamic change decreased with practice for NDI only [ΔHbO_2 (μM•cm) median ± IQR: initial = 12.55 ± 24.13, mid = 10.43 ± 19.88, final = 6.92 ± 14.79, retention = 10.07 ± 18.97; Fr: $p<0.05$, initial > final WSR: $p<0.001$]. However, PFC attenuation was not sustained and a greater response was observed during retention (retention > final WSR: $p<0.001$). Similarly, contrary to the attenuation hypothesis, paired between-session comparisons suggested that PFC excitation during STK was greater during latter ('final' and 'retention') versus intermediate ('mid') practice runs [ΔHbO_2 (μM•cm) median ± IQR: initial = 8.59 ± 19.21, mid = 3.86 ± 14.24, final = 7.31 ± 12.09; retention 8.77 ± 17.47, Fr: $p<0.05$; WSR: final and retention > mid $p<0.001$]. Indeed, once clustering was considered in a random effect model (see Table 6.4), practice session was not found to be a predictor of PFC cortical haemodynamic change ($p=0.065$). Sequential haemodynamic time courses (see figure 6.6) further reflect sustained PFC activation (see figure 6.7).

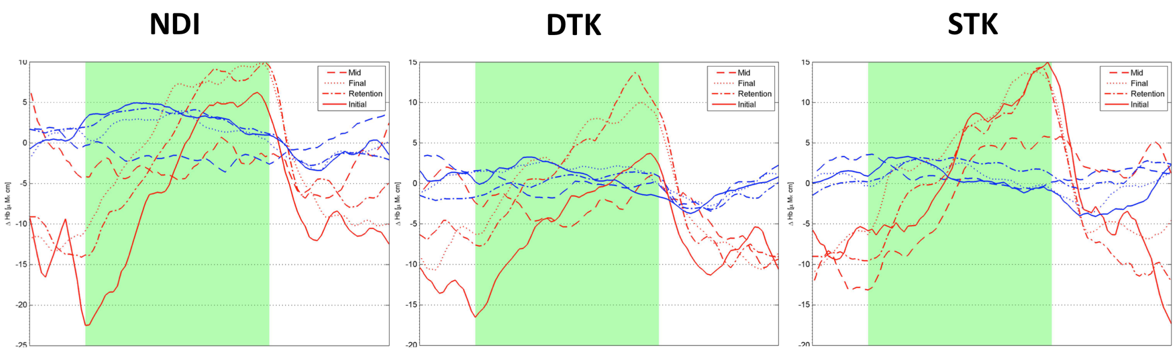

Figure 6.6. Illustrates group averaged changes in HbO_2 (red) and HHb (blue) concentration, within a channel centred on the right prefrontal cortex during sub-tasks needle drive insertion (NDI), double throw knots (DTK) and single throw knots (STK). Median change in Hb species (μM•cm) according to stage of practice. Initial, mid, final and retention session signals are correspondingly represented as bold, dashed, dotted and dash-dot lines.

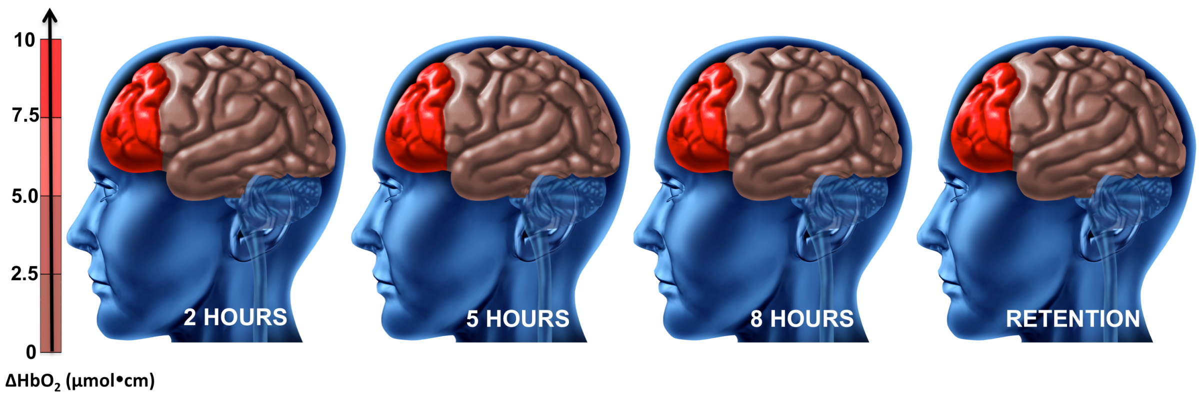

Figure 6.7. Graphic illustration of persistently raised PFC activity despite a week of practice. Group median rise in cortical oxygenation is depicted on a colorimetric scale and reflects the intensity of activation.

Premotor Cortex*:* For NDI, practice-related attenuation was identified in PMC haemodynamic responses from initial to final practice, followed by a subsequent significant increase at retention [ΔHbO_2 (µM•cm) median ± IQR: initial = 8.34 ± 17.20, mid = 7.47 ± 18.92, final = 3.78 ± 14.89, retention = 8.37 ± 14.22; Fr: $p<0.05$, retention > final WSR: $p<0.001$]. No significant differences in PMC haemodynamic change were observed between practice-sessions during knot-tying sub-tasks (DTK and STK).

Supplementary Motor Area: During NDI, SMA cortical haemodynamic change decreased with practice (see figure 6.8) from initial to final practice, and increased between final practice and retention [ΔHbO_2 (µM•cm) median ± IQR: initial = 8.23 ± 17.35, mid = 7.31 ± 13.89, final = 4.20 ± 12.17, retention = 5.71 ± 13.28; Fr $p<0.05$]. However, after Bonferroni adjustment, post-hoc paired comparison failed to reach significance. Practice-related changes in SMA activation during knot-tying subtasks failed to reach statistical threshold. However, session was observed to be a significant predictor for ΔHHb in SMA channels upon REM analysis (see Table 6.4).

Motor Cortex: Only during NDI was M1 activation found to decrease until final practice, followed by an increase on retention [ΔHbO_2 (µM•cm) median ± IQR: initial = 13.39 ± 17.42, mid = 9.41 ± 16.84, final = 4.43 ± 14.93, retention = 9.03 ± 17.37; Fr: $p<0.05$; initial > final WSR: $P<0.001$]. A similar trend was observed for STK, although changes failed to reach statistical threshold [ΔHbO_2 (µM•cm) median ± IQR: initial = 10.26 ± 18.24, mid = 8.96 ± 15.22, final = 6.70 ± 13.69, retention = 9.75 ± 13.64; Fr: $p=0.937$]. However, HHb analysis implied a strengthening of M1 activation as training progressed (see Table 6.2).

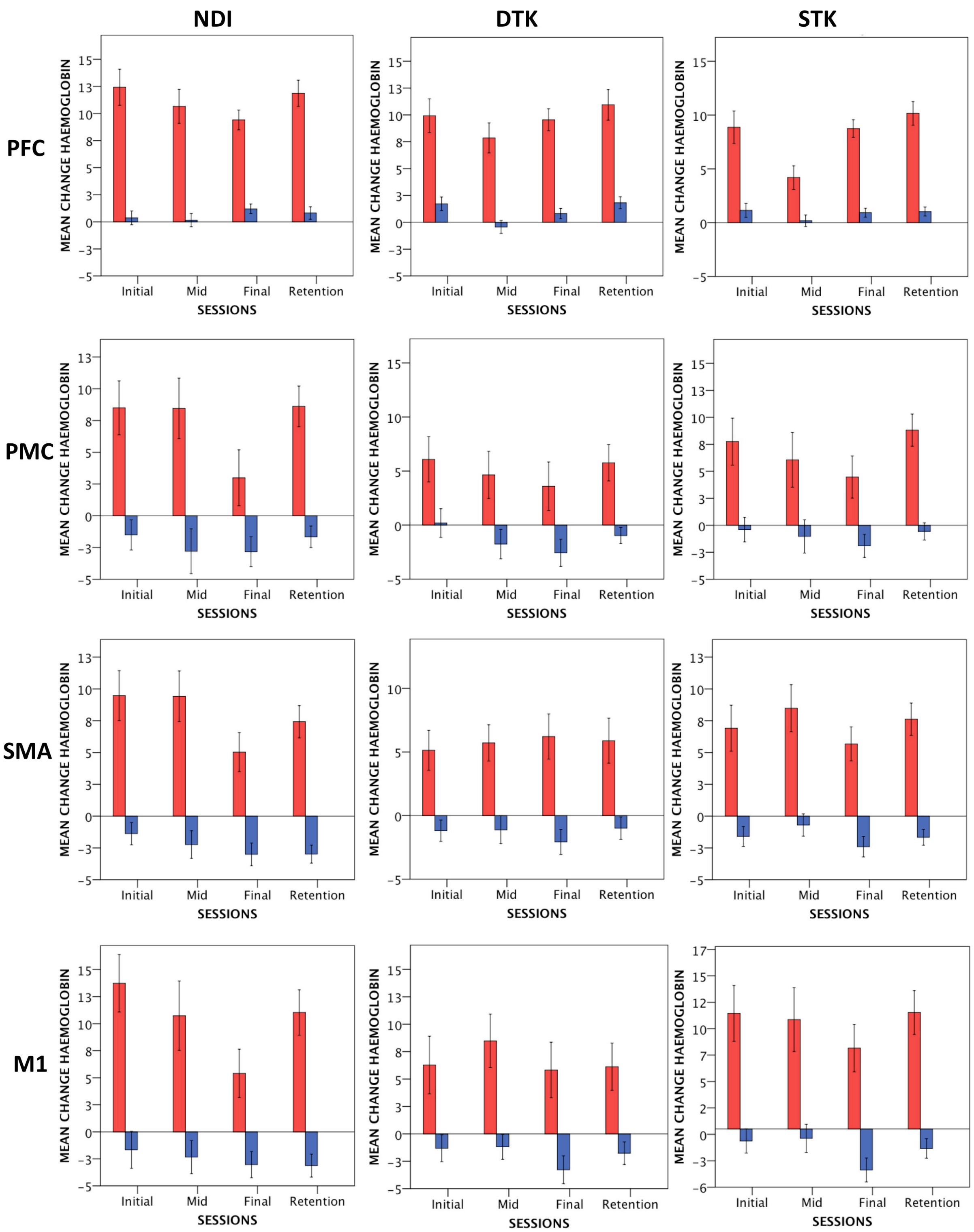

Figure 6.8. Bar charts illustrating regional changes in HbO_2 (red bars) and HHb (blue bars) at the initial, mid, and final practice sessions, and at retention whilst performing needle drive insertion (NDI), double throw knot (DTK) and single throw knots (STK). Data represent mean and standard error bars. Regions interrogated are the prefrontal cortex (PFC), dorsal premotor cortex (PMC), supplementary motor area (SMA) and primary motor cortex (M1). Change haemoglobin is represent µm•cm on the y axis.

ROI ΔHb	Variable	β	s.e.	z	p>z	95% C.I.	σ_u	σ_e	ρ fraction of variance due to u_i
PFC	Session	0.30	0.16	1.84	0.065	-0.02 to 0.62	6.93	16.5	0.15
ΔHbO_2	Constant	8.86	2.05	4.32	0.000	4.84 to 12.89			
PFC	Session	0.09	0.07	1.24	0.215	-0.05 to 0.22	2.66	7.17	0.12
ΔHHb	Constant	0.64	0.79	0.81	0.417	-0.91 to 2.20			
PMC	Session	-0.11	0.27	-0.39	0.696	-0.64 to 0.43	4.10	16.52	0.06
ΔHbO_2	Constant	6.39	1.41	4.52	0.000	3.62 to 9.17			
PMC	Session	-0.25	0.16	-1.53	0.127	-0.57 to 0.07	1.98	9.80	0.04
ΔHHb	Constant	0.64	0.79	0.81	0.417	-0.91 to 2.20			
SMA	Session	-0.25	0.22	-1.12	0.262	-0.67 to 0.183	4.88	13.93	0.11
ΔHbO_2	Constant	7.94	1.54	5.16	0.000	4.92 to 10.95			
SMA	Session	-0.34	0.12	-2.81	**0.005**	-0.58 to -0.10	1.44	7.69	0.03
ΔHHb	Constant	-0.84	0.54	-1.57	0.116	-1.90 to 0.21			
M1	Session	-0.77	0.33	-2.35	**0.019**	-1.42 to -0.13	5.39	13.57	0. 14
ΔHbO_2	Constant	11.53	1.83	6.30	0.000	7.94 to 15.12			
M1	Session	-0.48	0.16	-2.95	**0.003**	-0.80 to -0.16	2.15	6.67	0.09
ΔHHb	Constant	-1.03	0.78	-1.32	0.186	-2.55 to 0.50			

Table 6.4 Random effect analysis of the effect of practice session on task induced change in regional (ROI) cortical haemodynamic change (ΔHb). Regions interrogated were the prefrontal cortex (PFC), premotor cortex (PMC), supplementary motor area (SMA) and primary motor cortex (M1). Change in cortical haemodynamics of each haemoglobin species namely oxygenated haemoglobin (ΔHbO_2) and deoxygenated haemoglobin (ΔHHb) were examined.

6.3.6. Prefrontal Haemodynamics and Performance Improvement

The fitted surface of the linear fixed effect model correlates PFC oxygenation change with FLS across training (see Figure 6.9). The adjusted R^2=0.00445 (ANOVA F=0.789, p-value = 0.502) suggests the relationship departs from linearity. Therefore, this implies that changes in PFC cortical haemodynamic and technical performance were not significantly correlated across practice.

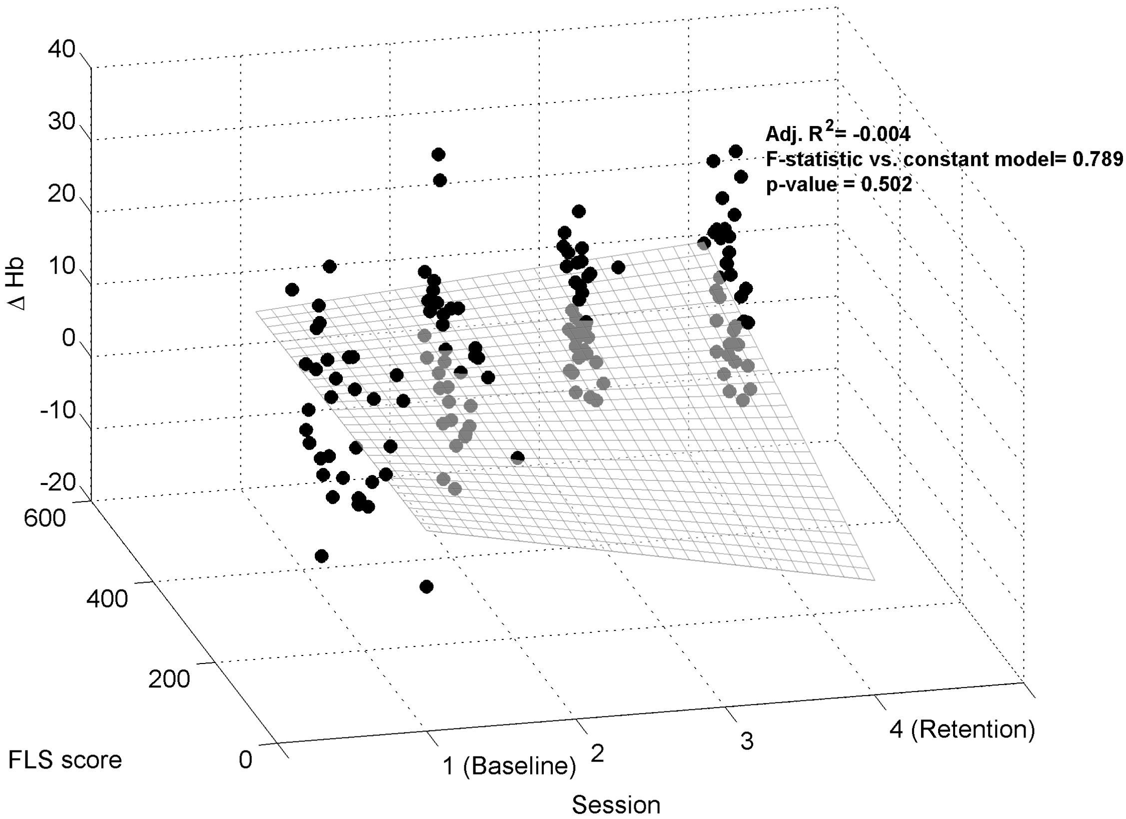

Figure 6.9. The fitted surface of the linear fixed effect model relates cortical haemodynamic change (ΔHbO_2) with training progress (session) and exhibited technical performance (FLS score). The relationship is clearly non-linear. Observations are depicted as dots (grey represent those lying behind the surface).

6.4 Discussion

Frontal lobe neuroplasticity associated with bimanual complex skills acquisition was interrogated in a longitudinal study design. The motor skill studied is both occupationally relevant for a high-risk field and requires a precise sequence of movements under multi-sensorimotor adaptations. Direct one to one practice led to significant improvements in motor skills and behavioural performance stabilization. Motor performance adaptation was associated with predicted longitudinal changes (increases) in haemodynamic changes in secondary motor areas such as the SMA, but not across areas of executive control. Indeed, in spite of behavioural gains, and at odds with motor learning literature [25,246] predicted attenuation in prefrontal responses was not identified, and no significant correlation between longitudinal changes in performance and brain behaviour was observed. Rather, persistent PFC recruitment was evident even after eight hours of deliberate practice.

6.4.1. Technical Performance

As expected, trained novices displayed improvements from the beginning until the end of practice. Indeed, performance gains were relatively greater at the early "fast" learning phase (initial to mid) and were followed by a late slow learning phase (mid to final). A retention test four weeks after completion of practice confirmed performance stabilisation.

6.4.2. Subjective Workload

NASA-TLX has been used as a training assessment tool to gauge operator comfort for a given task whereby a reduction in subjective workload would indicate preparedness to progression to the next level of training [80]. Moreover, subjective assessments of mental workload such as the NASA-TLX correlate with PFC haemodynamic change during visuospatial tasks [314]. In the current study, as predicted, an inverse relationship between burden and duration of practice was demonstrated. Greatest impact in reducing burden was associated with double knot-tying, since mental workload indices on this subtask were significantly lower upon practice termination. This notwithstanding, significant stepwise gains in technical performance between each training session were not necessarily accompanied by significant sequential reductions in subjective burden. Indeed, early performance gains (initial to mid) were made without significant alleviation in burden. This suggests that performance improvements early in training are possible even if demands are high.

6.4.3. Learning Related Changes in Cortical Activation

Following 8 hours of training over the course of two weeks, the cohort of trained novices made significant gains towards expert levels of technical performance. As postulated, longitudinal increases in PMC, SMA and M1 activation was suggested from progressively larger changes in HHb. Employing fMRI, Karni et al[317] detected an inter-practice session increase in M1 activity, and suggested that this reflected an expanded representation of the trained sequence. In the current study, random effect analysis indicated that practice was a significant predictor of HHb change across SMA channels reflecting embedding of the trained motor sequence. Increased activation in M1 and SMA amongst learners implies increase in processing time or encoding of additional neuronal units [277,368].

The anticipated pattern of redistribution expected with practice is an increase in activation in primary and secondary motor regions [277,300]. However, no such sustained, significant or sequential decrease in PFC activation was identified with short-term training. Results of the random effect analysis and longitudinal correlation analysis imply that PFC haemodynamic change was independent of practice-related changes in performance. Considering the complexity of this task, the duration of practice may have been too limited and more training or even overtraining may be required before demands on executive control are offset towards an expert pattern of brain behaviour. Redistribution of cortical activity might only occur after further refinement, protracted training and years of daily practice on LICS.

The five-stage Dreyfus model [369] of skill acquisition may better explain the lack of prefrontal attenuation than the model proposed by Fitts and Posner [42]. During the 'novice', 'advanced beginner' and 'competent' stages attentional processing is necessarily high to support motor performance. With practice, novices progressively shift their focus of attention onto different nuances of the task in order to refine performance. At the fourth stage, denoted as 'proficiency', the performer becomes emotionally involved in a task in order to strengthen successful from unsuccessful perspectives [51,369]. Prefrontal activation is also associated with reward and motivation, whereby with improvements in performance as a result of practice, participants gain positive feedback [370]. Therefore, hypothetically, attention-related PFC attenuation might have been camouflaged by an increase in reward associated PFC activation.

In the context of technical performance improvement and alleviation of subjective mental workload burden as a result of training, it is fascinating to observe persistent PFC excitation. Several studies suggest that escalating cognitive workload demands manifest as greater PFC haemodynamic change [314,371,372] but this is not universal [373]. Conversely, it follows that alleviation of burden as a result of training may prompt prefrontal attenuation, as observed in flight simulation training [314]. Therefore, the disconnect between improving workload scores on the one hand and sustained PFC activation merits further discussion. Firstly, these results imply that at a given level of performance, measurements of neuronal activation may be especially useful in differentiating expertise. Specifically, whilst differences in behavioural performance between trained novices and experts were not readily apparent, functional neuroimaging data suggests that trained novices were investing on average four-five times more attention and concentration than experts (see Table 5.4 in Section 5.3.6 of Chapter 5). Learners may misperceive the level of attention, concentration and on-going load, falsely perceive a reduction in their subjective workload or that the NASA-TLX instrument suffers from recall bias. Alternatively, given the complexity of LICS, persistent prefrontal recruitment may reflect the need for on going executive functions (e.g. planning, set shifting, motor attention) or for other 'non executive' PFC functions such as rewards and motivation [374,375].

Finally, in accordance with the hypothesis, upon retention testing following several weeks without practice, progressive prefrontal excitation was observed in all sub-tasks, and was significant for needle insertion. It is anticipated that the lack of deliberate practice may have led to a rebound increase in recruitment of executive control regions to support performance. Interestingly, progressive PMC activation between final practice and retention was observed. Improvements that occur without physical practice are referred to as "*off-line learning"* and dependent on either sleep or rest [376], and this suggests that off-line learning may have continued despite practice cessation.

6.4.4. Spatial and Longitudinal Disparity between Haemoglobin Species

The results of this study revealed asynchronous temporal cortical haemodynamic changes within the PMC, SMA and M1. For example, hypothesised longitudinal decreases in ΔHHb across the SMA signifying increased activation were not accompanied by progressive increases in ΔHbO_2. Similarly, across PFC channels increases in both ΔHbO_2 and ΔHHb were observed concomitantly. Clearly, the relationship between changes in ΔHbO_2 and ΔHHb in an activated brain are complex [377-379] and comprise the cumulative effects of a mismatch between an increased cerebral blood flow - CBF (increase in HbO_2 and decrease in HHb), increased cerebral blood volume (increase in HbO_2 and HHb) and cerebral metabolic rate of oxygen – $CMRO_2$ (decrease in HbO_2 and increase in HHb)[186]. Therefore, the observed increase in ΔHHb in relation to the increased HbO_2 may suggest decreased oxygen consumption in relation to an increase in HbO_2 or arteriolar dilatation with a regional cerebral blood flow increase. HHb increases are associated with cerebral venous dilatation leading to cerebral congestion as has been observed in a rat model [181] and attributed to an immature cerebral haemodynamic system in neonates [380]. Similarly, the post-stimulus undershoot in the BOLD response that arises because cerebral blood volume is slower to recover than CBF and $CMRO_2$ [377] may be measured as HHb increases and yet still reflect genuine activation responses. Finally, inadvertent placement over superficial scalp veins that dilate as a result of stimulation or stress may also manifest as elevations in HHb.

Importantly, the disparity in longitudinal trends between Hb species exposed here (i.e. change in one chromophore as predicted whilst the other remains stable or progresses in a juxtaposed direction) are challenging to contextualise since numerous fNIRS studies base their interpretations on cortical oxygenation changes alone [29,314,381,382] and there is a paucity of repeated measures fNIRS studies. One approach to reconcile the disconnect between observed patterns of asynchronous cerebral haemodynamic change and the underlying physiological processes that may be responsible is to employ a mathematical model of brain circulation and energy metabolism to aid with data interpretation [383]. Models such as these have already been utilised to better understand the relationship between demands on the brain and changes in cerebral haemodynamics and importantly to elucidate the underlying physiological mechanisms from which the measured fNIRS signals arise [384].

6.5. Conclusion

The pattern of practice related changes in frontal lobe activation on an LICS task is commensurate with established motor learning models that describe changes in secondary motor areas. Specifically, eight hours of distributed intensive practice led to improvements in skilled motor performance and strengthening of activation in secondary motor regions. However, concurrent prefrontal attenuation was not observed suggesting that either the challenges in executing complex bimanual tasks such as LICS demand sustained attention and executive control despite performance gains, or else that other cognitive factors such as rewards may lead to persistently high prefrontal responses. For highly complex bimanual skills, only through long-term training, repeated daily exposure and accuracy refinement that occur through years of practice may demands on executive control centres be released.

Study Limitations:

In comparison to fMRI, the established gold standard for studying human brain function, fNIRS has several limitations including: lower signal-to noise ratio; poorer spatial resolution; and the inability to capture activity of deep cortical and subcortical such as the cerebellum and basal ganglia that are associated with motor learning. However, for the purpose of studying LICS, fNIRS has several advantages, such as: resistance to motion artefacts; ability to maintain normal posture and enables the use of ferro-magnetic tools. Precise placement of fNIRS optodes over the frontal lobe enables interrogation of structures linked to various stages of motor learning. However, limitations in the coverage of optodes meant that only four channels were positioned over the primary motor cortices, limiting inferences from data obtained in this region.

Scalp flow is known to influence cortical haemodynamic data [199,232,385,386] and is modulated by stress. Specifically, experiments aiming to de-couple cortical and skin flow responses using near (5mm) and far (30mm) inter-optode distances [385,386] identified that a significant proportion of task-evoked changes assumed to be from cortical origin (far channels) were explained by changes in scalp haemodynamics (near channels) [386]. This notwithstanding, the rigid inter-optode distance of the current head-gear and the fixed emitter-detector multiplexing configuration of our system (ETG-4000) and the sparse numbers of channels across the PFC made it technically impractical for us to adopt both near and far optode arrangements. Instead, to control for likely disparity in stress responses between groups, HR and HR variability were monitored. Since no significant difference was observed between practice groups in HR it is unlikely that between-group differences in cortical haemodynamic change were mitigated by stress induced changes in systemic effects.

Behavioural confounds such as economy of kinematics associated with practice should also be considered. Whilst one study observed no effect of expertise on motion characteristics such as movement speed of either hand for any LICS sub-task [331], more recent data suggests that the forces exerted by novices during laparoscopic surgery supersede those of the more experienced [121], and that force perception is improved with experience [387] who presumably acquire knowledge regarding permissible forces through experience. I acknowledge that differences in movement intensity [273,388], velocity [291], amplitude [389], and frequency [390-392], which were not measured here may have influenced my results. However, given that increasing motor force or intensity is associated with increased M1 activation [273,388], and that novices are anticipated to use greater force, then the prediction would be greater M1 activation at the early phase versus late phase of learning, which is the inverse of the observed trend.

The experiment would have been strengthened with the use of a control group who did not receive practice but whom were repeatedly assessed with fNIRS. A control group would have enabled me to better decouple short-term learning related changes in cortical responses from habituation effects. However, if the findings were dominated by habituation then one might expect a decrease in the PFC response or signal as a result of boredom on the task, when in fact the converse was observed. Additionally, there are ethical considerations when seemingly only providing training to one group of medical students and not to another.

Subjects are also known to display idiosyncratic patterns of activation during motor execution, which may be masked by group analyses due to clustering effects. Variability between subjects in fNIRS responses has been previously reported [393]. To address clustering caused by inter-subject and within subject variability, the study was analysed using a random effect model. Finally, it should be acknowledged that the current experiment was deliberately controlled (familiarisation, minimising distractions, etc), to monitor task-specific cortical haemodynamics, and as such responses may not accurately reflect those derived during real surgery.

Chapter 7

Frontal Lobe Connectivity Aids Expertise Classification

Contents from this chapter have been published in:

Andreu-Perez J, Leff DR, Shetty K, Darzi A, Yang GZ. Disparity in Frontal Lobe Connectivity on a Complex Bimanual Motor Task Aids Classification of Operator Expertise. *Brain Connectivity*. 2016 Jun; 6(5): 375-88.

7.1. Introduction

Neuroergonomics captures the study of brain behaviour at work with the goal of improving performance, safety and efficiency[243,394]. This is particularly relevant to fields that warrant operator vigilance, technical skill levels and decision-making such as command control in aviation[340] and surgery[395]. Of these, surgery is unique in that the operator's technical and cognitive skill levels have a direct impact on patient safety. In theory, enhanced neuronal efficiency in surgeons acquired through practice and training or through improved ergonomics may manifest as improved patient safety. Deepening our understanding of operator brain behaviour may have value when applied to improve performance, particularly in high-risk procedures. Advances in functional neuroimaging technology have made it possible to monitor operators in more realistic settings and track evolution in brain behaviour that accompanies motor skill levels learning. To this end, there has been an increasing number of research studies focused on studying evoked cortical response to complex motor behaviour in the context of open and minimally invasive surgery (MIS)[29,259,260,265,266,272,361]. In this study we used functional near-infrared spectroscopy (fNIRS), which has raised increasing interest in recent years for performing less constricted, hence more naturalistic neuroscience experiments[396].

Previous studies have exposed skill level-related differences in brain behaviour and longitudinal changes in cortical excitation in line with technical skill levels acquisition[194,361]. For example, data suggest that cortical responses may be skill level-dependent[259]. Greater activation within executive control centres such as the prefrontal cortex (PFC) has been observed in novices and PFC excitation appears to attenuate following practice[260]. In a study of 18 subjects, decreasing ratios of oxygenated haemoglobin were observed in supplementary motor area (SMA) and preSMA when learning a motor skill[227]. Similarly, a decrease in cortical activation of the sensorimotor cortex was reported during the learning of a multi-joint discrete motor task[397]. This is commensurate with evidence that re-organisation of brain function (i.e. neuroplasticity) accompanies motor skill levels learning, such that operator skill level may be best reflected in the magnitude of regional brain excitation or shifts in activation foci[246,277,398]. This notwithstanding, investigations of highly complex motor skill levels such as MIS, have failed to demonstrate differences between naïve subjects and expert operators[29]. One theory is that for highly complex motor tasks such as MIS that require 2D to 3D perceptual transformation and precise inter-manual co-ordination, skill *level*-related disparity may manifest as differences in frontal lobe connectivity rather than changes in activation per se. This can be deduced from the cortical network differences reported in Sun et al[221] across early and late motor learning states, as well as the variations in cortical connectivity presented in James et al[253] during skill acquisition for a surgical task.

Whilst longitudinal changes in network topologies have previously been studied in surgical learners, cortical interactions have only been tracked over days[253] and not months or years. In this study from our group[253], changes in the network cost across days of the practice were found amongst subjects of similar skill level. One advantage of evaluations that incorporate master operators is that they facilitate interrogation of motor plasticity as a result of repeated practice of a technical skill level over many years. Indeed, few studies have investigated learning-related changes in connectivity across such timescales, and current reports are restricted to minutes or weeks[25,221,299,354,399]. Nevertheless, existing literature seems to suggest greater connectivity between frontal and cortical motor regions in 'early' versus 'late' learning[221] and longitudinal attenuation in functional integration in motor-related networks following extended practice[299]. Therefore, conceivably functional connectivity between associative and premotor regions may be expected to decrease in line with continued practice and increasing operator skill level.

Differences in functional connectivity may help discriminate operators based on their skill level. Classification of operator proficiency based on brain behaviour may prove invaluable for objective assessment of technical skill levels, evaluation of trainee progress, and if "interfaced" to the operator or team, may improve performance or aid patient safety through cognitive biofeedback[400-402]. Whilst amplitude of the evoked response has been used for classification of operator states previously[403-405], there have been no such reports in surgeons. Indeed, whilst differences in amplitude signal change in executive control and motor cortical regions related to surgical skill level have been previously observed[29,259], it has not been possible to discriminate operators' performance based solely on these signal characteristics. Classification of operator proficiency based on functional connectivity data represents a non-trivial high dimensionality problem for which conventional statistics are ill-posed to solve. Machine learning (ML) techniques such as least-squares support vector machine (LS-SVM) generate predictive models that can discriminate data based on the complexity of multivariate relationships such as those arising from functional connectivity.

In this study, we aim to investigate and discriminate operator skill level during laparoscopic (key-hole) surgical manoeuvres using brain connectivity derived from evoked optical imaging responses. Specifically, we employ Optical Topography (OT) to monitor operator brain function during a highly complex surgically relevant motor task (i.e. key-hole surgical suturing). We hypothesize that frontal lobe connectivity in associative (prefrontal) and premotor seed regions will decrease in line with increasing operator experience. Functional connectivity computed from one-second epochs of filtered oxygenated and deoxygenated haemoglobin signals (HbO_2 and HHb, respectively) was used to classify operator skill level. Additionally, an analysis of the discriminatory performance of local (within-region) and inter-regional connectivity was performed. To facilitate this comparison, a single metric of performance named Multiclass Mathew's Correlation Coefficient (MMCC)[406] was employed.

7.2. Materials and methods

7.2.1. Experimental set-up and neuroimaging data acquisition

Local Research Ethical Council approval was obtained (05/Q0403/142). Participants were screened for handedness[365], gender and neuropsychiatric illness. Participants abstained from consumption of alcohol and caffeine for 24 hours before the study date[199]. Thirty-two right-handed male surgeons were recruited from the National Health Service (NHS) and Imperial College London to perform a complex visual-spatial task, namely simulated laparoscopic suturing (LS) (i.e. key-hole surgical stitching), in a box trainer (i-SIM, iSurgicals, UK). The cohort included 12 novices (with no prior experience of MIS, mean age 22.4 ± 1.6 years), 11 trainees (limited MIS experience of less than 50 cases involving LS, mean age 33.8 ± 2.9 years) and 9 expert consultants (minimum of 50 independent cases requiring LS, mean age 42.7 ± 3.6 years). Participants were required to perform the task three times as per the fundamentals of laparoscopic surgery curriculum (FLS) (FLS, 2015). Statistical analysis on the FLS was performed using Kruskal Wallis (KW), followed by Dunn's Test (DT) for multi-comparisons. During each of the three sessions, the surgical procedure was segregated sequentially into three subtasks (see Figure 7.1), namely 'needle insertion', 'double-throw knot tying' and 'single-throw knot tying'.

Prior to recording, the participants were allowed a brief familiarisation session (15 minutes). The experimental design consisted of a sequence of continuous episodes of baseline motor rest (30 seconds), subtask phases during which given LS manoeuvres were executed (variable time i.e. self-paced) and inter-trial recovery periods (40 seconds). During the rest episodes the participants were asked to remain still and regard the centre of the monitor. Observable technical performance in LS was assessed by adoption of the FLS score [Score = 600 – (time in seconds) – (penalties $\times$10)]. Penalties were measured as per needle entry and exit distance points from pre-marked points (mms) as well as gap between the edges of wound (mms), i.e. an objective quantitative assessment. The derived score based on time and accuracy of performance was compared across experience groups using statistical tests of significance.

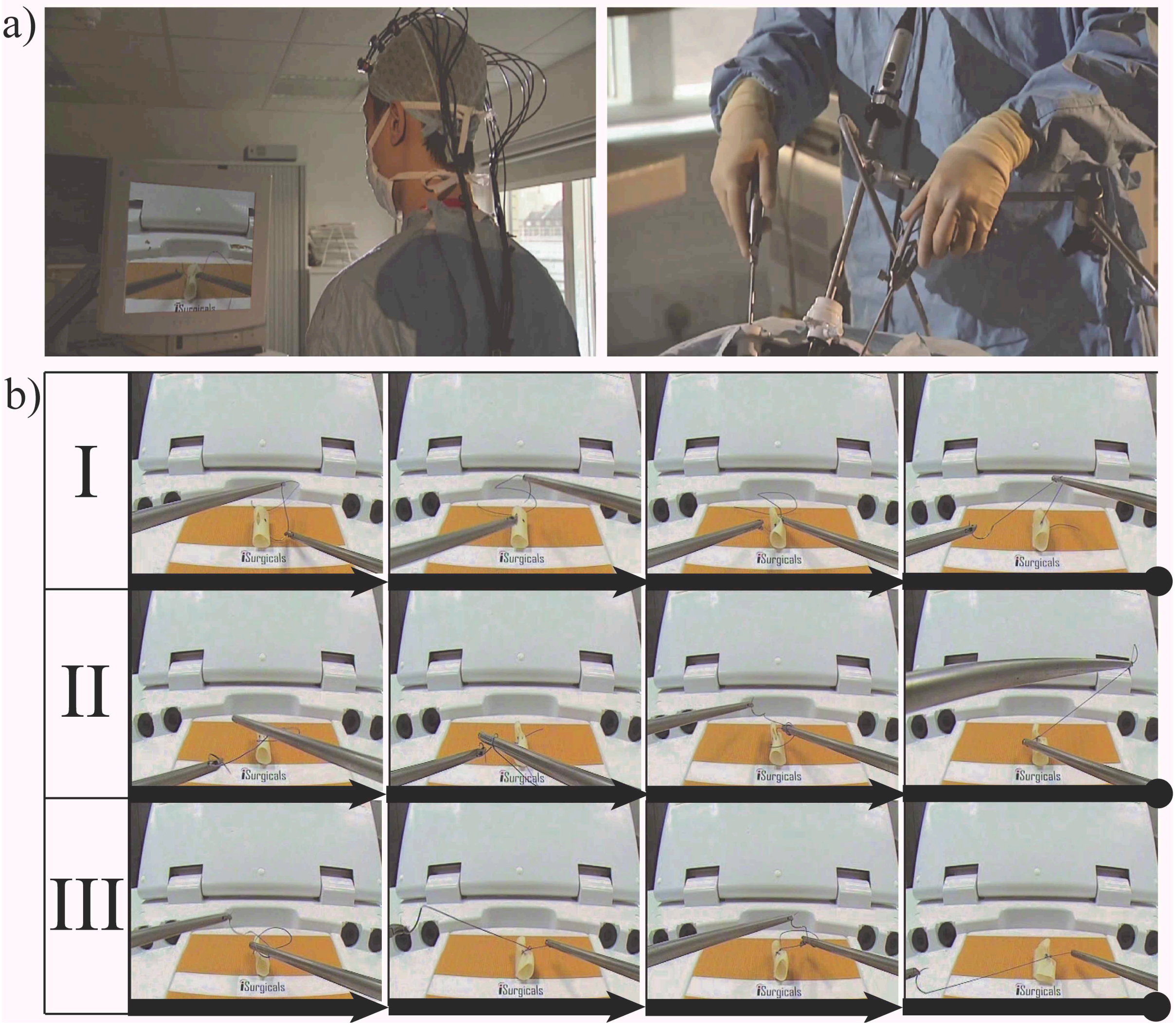

Figure 7.1. A participant (a) performing LS in a box trainer with simultaneous recordings of cortical brain activity using OT; (b) The three subtasks of the surgical procedure, namely 'needle insertion' (I), 'double throw knot' (II) and 'single throw knot' (III) are illustrated.

Cortical hemodynamic data were recorded at 10 Hz using a 44-channel OT system (ETG-4000, Hitachi Medical Corp, Japan). To capture functional behaviour relating to motor skill levels acquisition, channels were positioned over the PFC, premotor cortex (PMC), SMA and motor cortex (M1) regions of the brain according to landmarks of the International 10-10 system[200] (see figure 7.2). Optode positions were measured using a 3D digitizer. A stand-alone registration method was used to project NIRS probe positions into a MNI (Montreal Neurological Institute) coordinate space[407]. As OT data may be influenced by changes in the systemic circulation[408,409], a portable electrocardiogram sensor was attached to the chest of each subject to continuously monitor the heart rate (HR) during the course of the experiment. Mean heart rate (MHR) and heart rate variability (MHRV) was extracted from the electrocardiogram sensor readings using R-R wave intervals. Statistical significance was computed using analysis of variance ANOVA over these measurements.

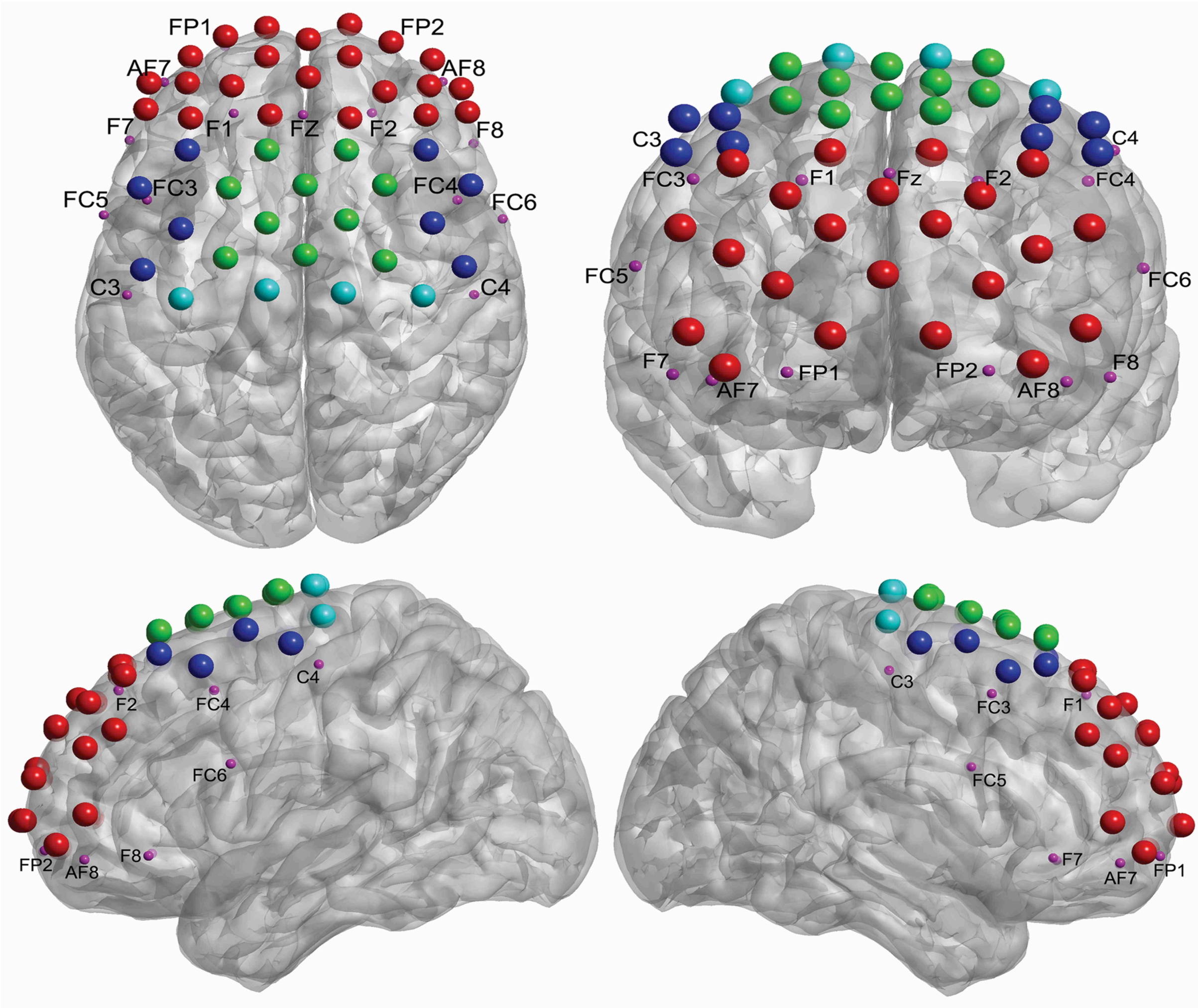

Figure 7.2. Registration of channel positions in MNI space illustrate the approximate locations over the prefrontal cortex (red), supplementary motor area (green), premotor cortex (dark blue) and primary motor cortex (soft blue) relative to International 10-10 markers (magenta).

7.2.2. De-noising and filtering

For each channel of data, an average baseline was computed considering 10s of data prior to each task onset to allow for haemodynamic normalisation after task offset. During this period, subjects were asked to refrain from any motion prior to task onset and not to engage in deliberate cognitive work. For each subtask, the first 6s of data were discarded, thereby allowing for the temporal delay between task onset and cortical haemodynamic change. At each channel, changes in the concentration of HbO_2 and HHb were reconstructed from variations in light attenuation using the modified Beer-Lambert law[188]. To reduce systemic interference, haemodynamic data were low-pass filtered, linearly detrended to eliminate system drift and integrity checked to remove noisy channels using Imperial College Neuroimage Analysis (ICNA) software[199].

7.2.3. Between-group differences in functional connectivity

The methodology for computing the between-group differences consists of the following sequential steps:

1) For every sample of haemoglobin data, the difference between the current value and the average baseline is computed. The resultant metrics are denoted by ΔHbO_2 or ΔHHb.
2) ΔHbO_2 and ΔHHb from channels belonging to the same region of interest (ROI) are grouped and averaged, resulting in $\mu_{\mathrm{HbO_2}}$ and μ_{HHb} for each ROI.
3) To obtain a single metric of inter-connectivity, the Rv coefficients[410,411] are calculated between the pairs of $\mu_{HbO_2}$and μ_{HHb} from two ROIs. Rv coefficients form a metric of multivariate statistical relationship to measure the dissimilarity between two sets of data. It is formulated as follows:

$$r = \frac{trace(\alpha\, \alpha^T\, \beta\, \beta^T)}{\sqrt{trace((\alpha\, \alpha^T)^2) \times trace((\beta\, \beta^T)^2)}}; \qquad \begin{matrix} \alpha = \left[\mu_{HbO_2}^{\mathrm{ROI}^1}, \mu_{HHb}^{\mathrm{ROI}^1}\right] \\ \beta = \left[\mu_{HbO_2}^{\mathrm{ROI}^2}, \mu_{HHb}^{\mathrm{ROI}^2}\right] \end{matrix} \qquad (1)$$

where we denote Rv coefficients with the function r taking as inputs the mean-centered α and β matrices; and the indexes ROI^1 and ROI^2 represent two distinct regions of interest. Once computed, Rv coefficients return values between 0 and 1 that can be interpreted as an approximation of the squared Pearson correlation coefficient[410,411].
4) Fisher transformation is applied to the Rv coefficients yielding the Gaussian distributed z scores. In order to obtain a single informational value, the z scores are averaged and the inverse Fisher transformation is applied, transforming the average back to its original Rv coefficient. The resultant average value is then considered as

the inter-regional connectivity between two brain areas of a subject performing a given LS trial (session).

Statistical analysis of normalized z-scores for inter-regional connectivity was performed using the ANOVA test, followed by post-hoc analysis using Tukey's honest significance difference (HSD).

7.2.4. Machine learning from functional connectivity for operator skill level discrimination

7.2.4.1. Functional connectivity datasets for automated discrimination

Two sub-datasets with different granularity are studied to classify operator skill level from fNIRS data: the *session based networks* and *time-course based networks*. The former aims to classify operators based on data from an entire session while the later from every one-second time epoch within a session. Both are derived from the ΔHbO_2 and ΔHHb readings of the original dataset captured at 10 Hz.

a) *Session based networks:* Rv coefficient is a scalar that determines the relationship of the joined signals, ΔHbO_2 and ΔHHb, between two channels during a session (i.e. the duration of one trial). Each correlation matrix represents a session. Each element within the matrix corresponds to a scalar that represents the 2D relationship between two channels when jointly considering both signals of interest, namely ΔHbO_2 and ΔHHb. The resulting dimensionality of the correlation matrix is 44×44, i.e. one element for each channel, out of which 22 are on the PFC, 10 are on the SMA, 8 are on the PMC, and 4 on the M1.

b) *Time-course based networks*: Due to the averaging used to construct an exemplar network across a whole session, some latent discriminatory information might be omitted. A short period connectivity network can be constructed using Spearman's correlations between channels, this time considering ΔHbO_2 and ΔHHb separately. To increase the granularity, correlations are computed within one-second epochs across the whole session. Hence, each correlation matrix corresponds to a single epoch and each element within the matrix is a single correlation of either ΔHbO_2 or ΔHHb between two channels. As a result, two 44x44 correlation matrices are generated, one for each haemoglobin species.

7.2.4.2. Classifier and parameter settings

For the analysis of the functional connectivity datasets, a least-squares support vector machine (LS-SVM) with a non-linear radial basis function kernel is used. In LS-SVMs the parameters of the separating hyper-plane are formulated as a closed-form linear system of equations[412,413]. The entire data are first divided into 5 cross-validation subsets and one is held out for testing. Before each test, optimization of the L2 internal regularization parameter is performed over a 5-fold cross-validation process using only the training set. This prevents over-fitting over the presence of irrelevant features. The optimization algorithm used is simulated annealing[414].

7.2.4.3. Performance measures for classification

To evaluate the performance of the classifier for each of the two sub-datasets defined in 7.2.4.1, standard metrics are computed from the true/false positive/negative (i.e. TP, TN, FP, FN) frequencies. These metrics of performance are:

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN};\ \ precision = \frac{TP}{TP + FP}$$
$$sensitivity = \frac{TP}{TP + FN};\ \ specificity = \frac{TN}{TN + FP} \tag{2}$$

Considering the metrics presented in (2), every class would lead to an independent metric of performance, which is not desirable for assessing the overall performance. Instead, a practical metric of classification performance is the Mathews' correlation coefficient (MCC), which considers all TP, TN, FP and FN simultaneously. For binary class problems the MCC is formulated as:

$$MCC_{binary} = \frac{(TP * TN - FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}; \quad (3)$$

MCC can be re-formulated to be multiclass and define a global performance metric representative of all classes[406,415]. In order to make it multiclass, the whole multiclass confusion matrix can be used as the reference to compute the MCC statistics. The multiclass MCC (MMCC) is obtained from marginalizing the different dimensions of the confusion matrix C, which elements include all true/false positive/negative frequencies for each class. Each row of C represents a class prediction while each column accounts for the true class. Hence, formula (3) becomes:

$$MMCC_{multiclass} = \frac{\sum_{q,e,p} C_{q,q} * C_{e,p} - C_{q,e} * C_{p,q}}{\sqrt{\sum_{q} \left(\sum_{e} C_{e,q}\right)\left(\sum_{e',q' \neq q} C_{e',q'}\right)} \sqrt{\sum_{q} \left(\sum_{e} C_{q,e}\right)\left(\sum_{e',q' \neq q} C_{q',e'}\right)}} \quad (4)$$

Where $C_{(.)(.)}$ is a single element of the confusion matrix and all matrix indexes (q, e, p) start with 1 and finish with K, unless when specified. Formula (4) generalises MCC and can be used either for multiclass or binary problems. The value obtained with MMCC is always between -1 and 1. If the coefficient is -1, the MMCC portrays a total disagreement between predictions and the ground truth, whereas 0 indicates a random prediction and 1 means a ground truth match.

7.2.4.4. Predicting operator skill level

After training the classifier using *a priori* data, operator skill level can then be predicted for any new, unseen (i.e. unlabelled) set of fNIRS data. Figure 7.3 summarises the processing steps necessary to obtain these predictions:

1) A signal processing step is composed of the following tasks:
    a. Signal extraction: 1Hz of haemoglobin data captured by each OT channel is stored in an independent buffer.
    b. Signal filtering (section 7.2.2): Each signal in a buffer is denoised, smoothed and baseline-subtracted.

2) Functional connectivity: coefficients are computed as a proxy for functional connectivity for all between-channel signal pairs, thus generating a functional connectivity matrix. The elements of the resulting matrix are translated into a sample vector, in which each vector dimension corresponds to an element of the matrix. In the case where two functional connectivity matrices are generated (one for each haemoglobin species), the elements of both matrices are linked together within a same sample vector.

3) Skill level estimation: In advance of testing, a LS-SVM algorithm is trained and parameter optimization is performed as described in 7.2.4.2. The sample vector obtained in step 3 is tested against a support vector machine model by using one of the multiclass settings described in 7.2.4.3. New data samples are analysed and a prediction is made regarding the likely skill level.

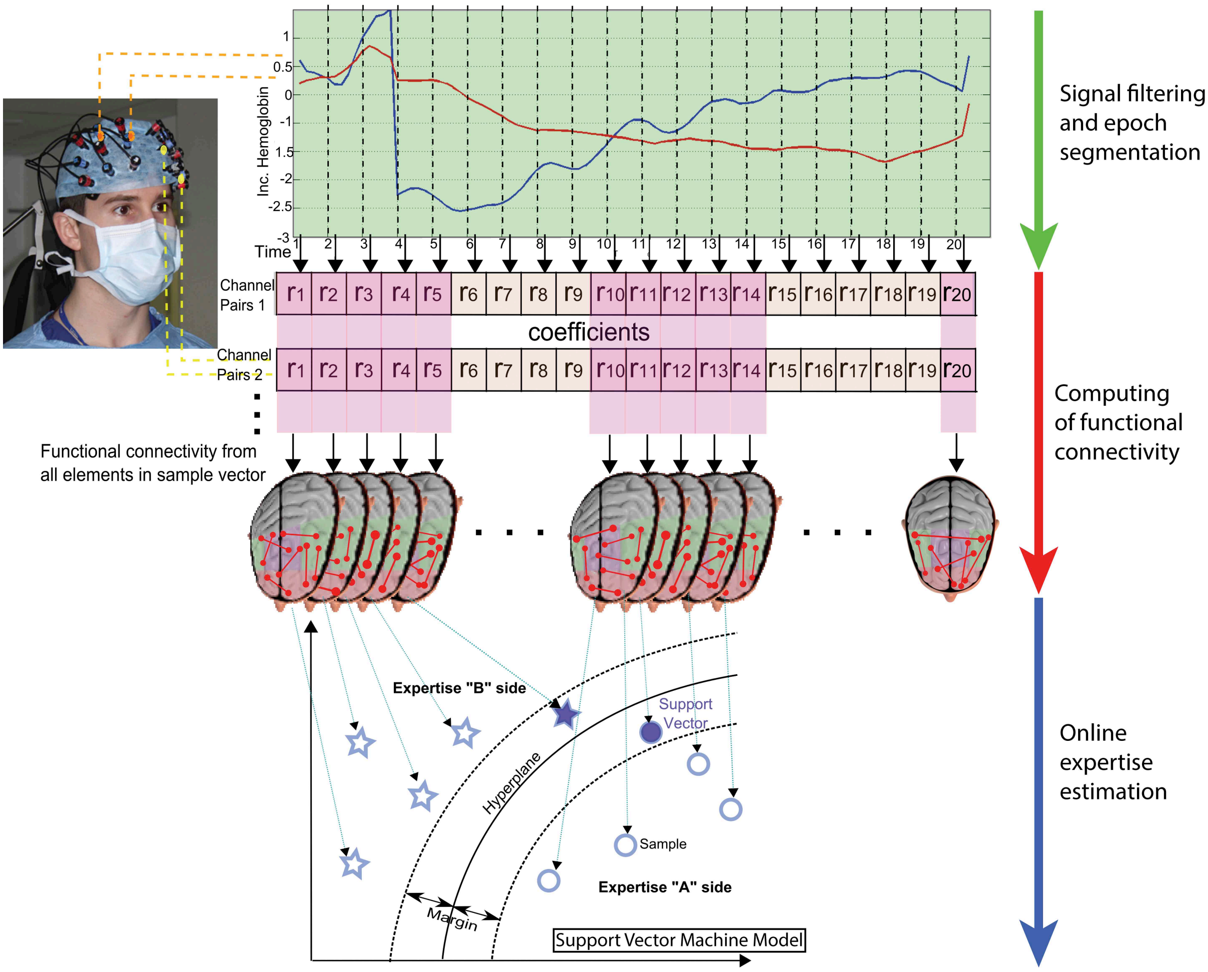

Figure 7.3. Schematic illustration of the processing steps and classification of the extended dataset using a large-margin classifier, i.e. LS-SVM.

7.2.5. Evaluating discrimination capabilities by groups of connectivity

It is valuable to explore the relative importance of each brain area, by considering the connectivity within and between brain areas in the frontal lobe, as opposed to solely considering classifier performance based on the fully connected adjacency matrix. The extended dataset is used in this analysis. In this regard, the classifier is adapted to consider only the within-region (local) or between-region (inter-regional) connectivity for a given brain region of interest (ROI). As illustrated in Figure 7.4, for a given ROI the local, inter-regional or combined functional connectivity data may be considered and defined as:

a) *Local connectivity (within-region)*: Considering observed connections within a given brain region. In other words, it refers to the internal functional connectivity of a particular brain region without considering long-range connections to other brain regions.

b) *Inter-regional connectivity:* Involves observed connections between channels of different brain areas, capturing distant functional relationships between areas of the brain, pruning local connections within the region.

c) *Combined (within-region and inter-regional)*: Includes all possible functional connections, both within a given ROI and between that ROI and other brain regions, thereby emphasizing all possible functional relations exhibited by a specific area.

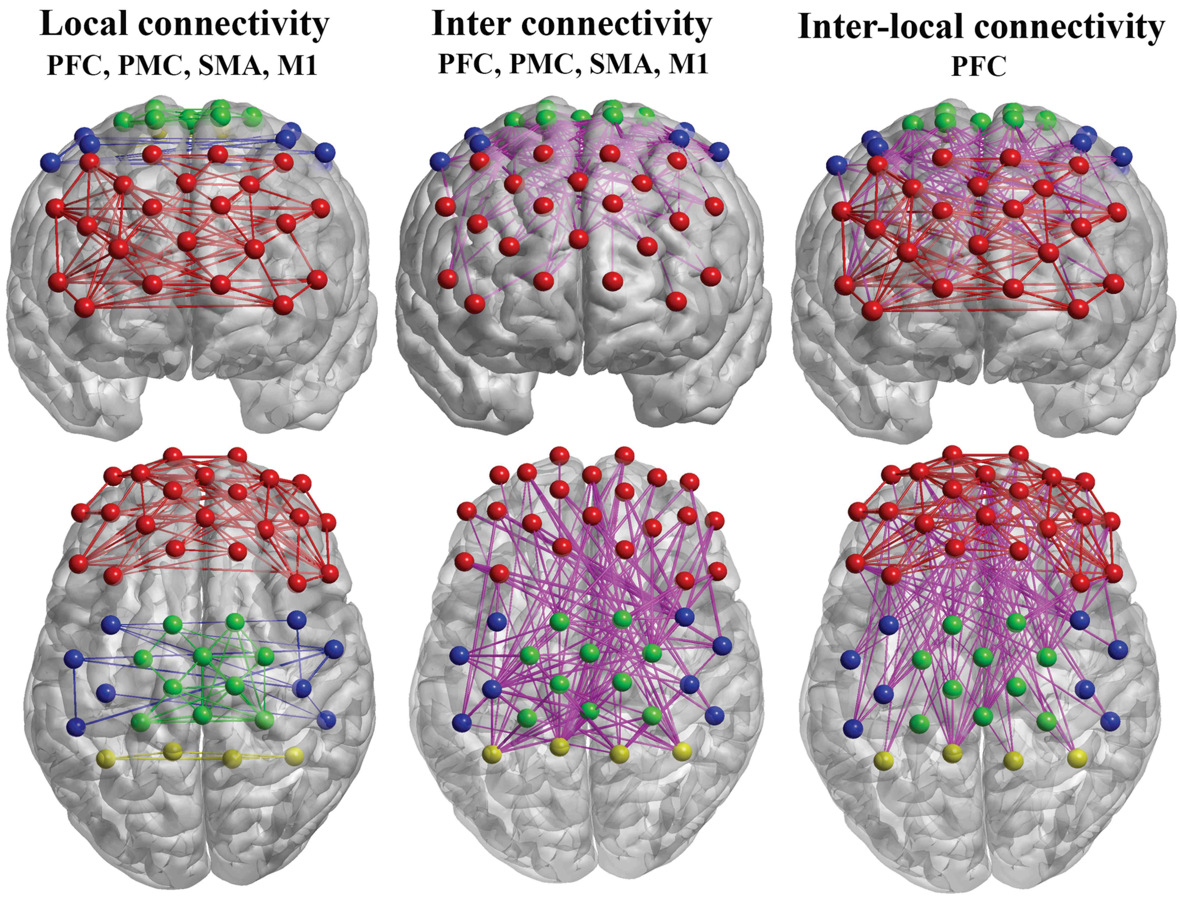

Figure 7.4. Graphical representation of local, inter-regional and combined (inter-local regional) connectivity showing channels over the PFC (red), PMC (blue), SMA (green) and M1 (yellow). Each edge represents a functional connection between channels. Within-region connections (colour coded by sub-region) and inter-regional connections (magenta) are highlighted.

7.3. Results

7.3.1. Behavioural performance and cognitive load

Table 7.1 summaries the results of task-related change in MHR, MHRV and assessment of motor laparoscopic skill levels (FLS). As anticipated, technical performance was significantly different between groups (p<0.001, Chi-square=58, df=2, KW). Experts displayed superior performance (lower FLS scores) compared to trainees (p<0.001, z=-32.2, DT) who in turn outperformed novices (p=0.002, z=-22.8, DT). No statistically significant differences in task-related MHR (p=0.87, F-value=0.145, df=2; ANOVA) and MHRV (p=0.83, F-value=0.182, df=2; ANOVA) were found between groups. For all groups, the null hypothesis is retained under Shapiro-Wilk test of normality for HRV (Novices: p=0.42, Trainees: p=0.42, Experts: p=0.50) and MHRV (Novices: p=0.61, Trainees: p=0.21, Experts: p=0.33).

Values		**Experts [E]**	**Trainees [T]**	**Novices [N]**	**Significance test**
Heart Rate	Mean beats per min.	80.71 ± 12.03 (Mean ± S.D.)	84.31 ± 11.2 (Mean ± S.D.)	82.98 ± 15.66 (Mean ± S.D.)	*p*=0.87 *F-value*=0.145 *df*=2
Heart Rate Variability	Mean R to R wave interval	0.76 ± 0.12 (Mean ± S.D.)	0.72 ± 0.1 (Mean ± S.D.)	0.75 ± 0.13 (Mean ± S.D.)	*p*=0.83 *F-value*=0.182 *df*=2
Motor Performance &Technical Skill (FLS score)		487 ± 53 (Median ± IQR)	400 ± 90 (Median ± IQR)	334 ± 76 (Median ± IQR)	*p*<0.001 *Chi-square*=58 *df*=2

Table 7.1. Between-group differences in task-related heart rate and technical skill (FLS).

7.3.2. Learning-related changes in local and inter-regional connectivity

For all groups, the null hypothesis is retained under Shapiro-Wilk's (SW) test of normality (Novices: p=0.31, Trainees: p=0.11, Experts: 0.09).

7.3.2.1. Needle insertion

Frontal lobe connectivity during needle insertion varied according to surgical skill level (Figure 7.5). Correlations between PFC and PMC seed regions and other frontal motor cortical regions were observed to be lower in trainees and experts than novices. Overall, a significant main effect of skill level was observed for the interactions PFC-SMA (p<0.01, F-value=20.06, df=2; ANOVA), PFC-PMC (p<0.01, F-value=16.05, df=2; ANOVA), SMA-PMC (p<0.01, F-value=36.67, df=2; ANOVA). Upon post-hoc analysis no significant differences between trainees and experts were observed (PFC-SMA: p=0.56, PFC-PMC: p=0.97, SMA-PMC: p=0.09; Tukey's HSD). However, comparisons between novices and trainees (PFC-SMA: p<0.01, PFC-PMC: p<0.01, SMA-PMC: p<0.01; Tukey's HSD), and novices and experts were statistically significant (PFC-SMA: p<0.01, PFC-PMC: p<0.01; SMA-PMC: p<0.01; Tukey's HSD). No main effect of skill level was observed for M1-PFC or M1-SMA seed interactions. However, a main effect of skill level was observed for M1-PMC interactions (p=0.01; ANOVA). Post-hoc analysis revealed no significant difference between trainees and experts while, differences between novices and trainees (p=0.01; Tukey's HSD) and novices and experts (p=0.03; Tukey's HSD) reached statistical threshold.

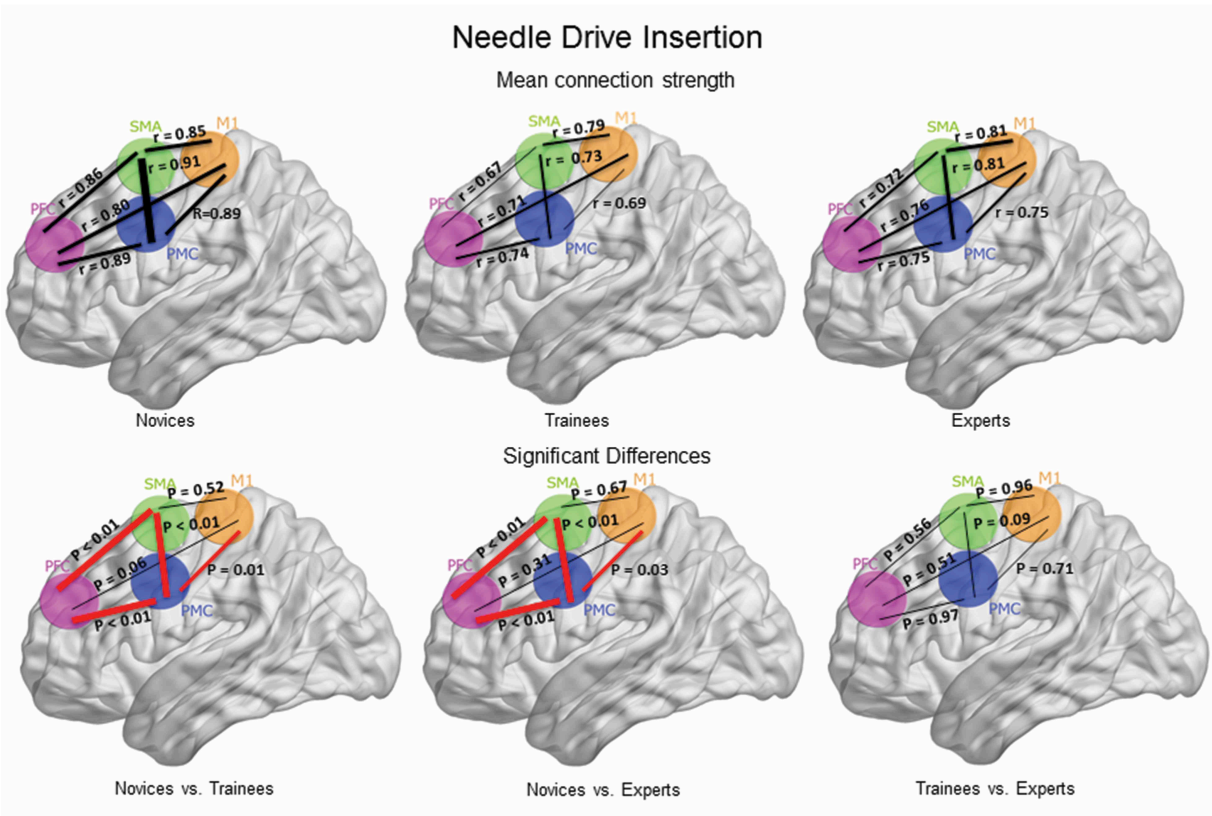

Figure 7.5. Sagittal brain views of learning-related changes in connectivity for the needle insertion sub-task. The three sagittal brains at the top display the mean Rv connectivity strength, while the following three at the bottom show the significance strength of the statistical test. Areas are depicted as PFC (magenta), SMA (green), M1 (orange) and PMC (blue).

7.3.2.2. Double-throw knot tying

The strength of PFC-PMC interactions was not observed to depend on skill level (Figure 7.6.) for double-throw knot tying. No significant differences were found for PFC seed interactions. However, interactions between the PMC and other frontal lobe seed regions varied with skill level. For example, a main effect of skill level was observed for interactions between the SMA and PMC (p<0.01, F-value=8.574, df=2; ANOVA). The interaction strength of this connection was significantly lower in experts than novices (SMA-PMC: p<0.01; Tukey's HSD). Of all frontal lobe interactions, differences between trainees and experts were only observed in SMA-PMC, the interaction strength being significantly lower in experts (SMA-PMC: p=0.04; Tukey's HSD).

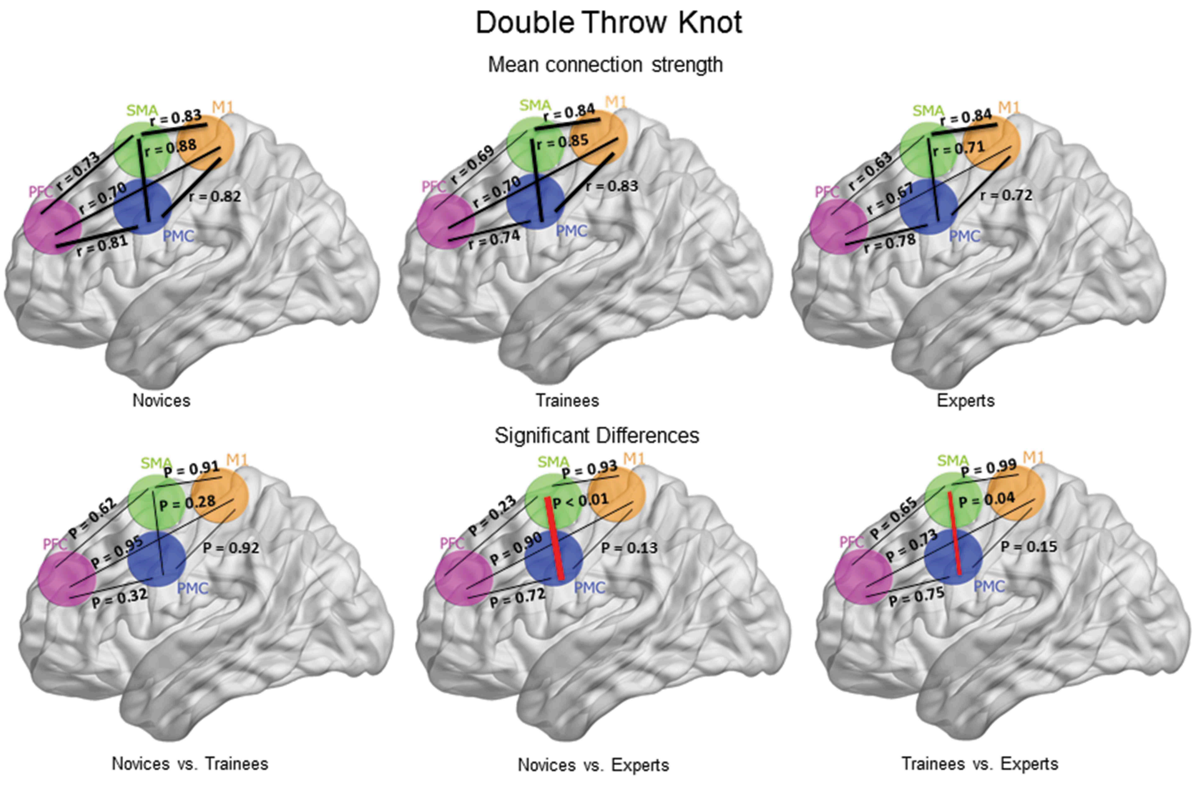

Figure 7.6. Sagittal brain views of learning-related changes in connectivity for the needle double-throw knot subtask. The three sagittal brains at the top display the mean Rv connectivity strength, while the following three at the bottom show the significance strength of the statistical test. Areas are depicted as PFC (magenta), SMA (green), M1 (orange) and PMC (blue).

7.3.2.3. Single-throw knot tying

A significant main effect of skill level was observed for the interactions PFC-PMC (p<0.01, F-value= 11, df= 2; ANOVA), PFC-SMA (p=0.03, F-value= 4.14, df= 2; ANOVA) and SMA-M1 (p=0.01, F-value=6.49, df=2; ANOVA). For PFC seed interactions, the strength of association was observed to be significantly lower in experts than novices (PFC-SMA: p=0.03; PFC-PMC, p<0.01; Tukey's HSD) and also lower for trainees between PFC-PMC (p=0.03, Tukey's HSD) (Figure 7.7.). Interestingly, differences between experts and trainees were only evident in SMA-M1 interactions, the strength of association being significantly greater in experts (SMA-M1: p=0.01; Tukey's HSD).

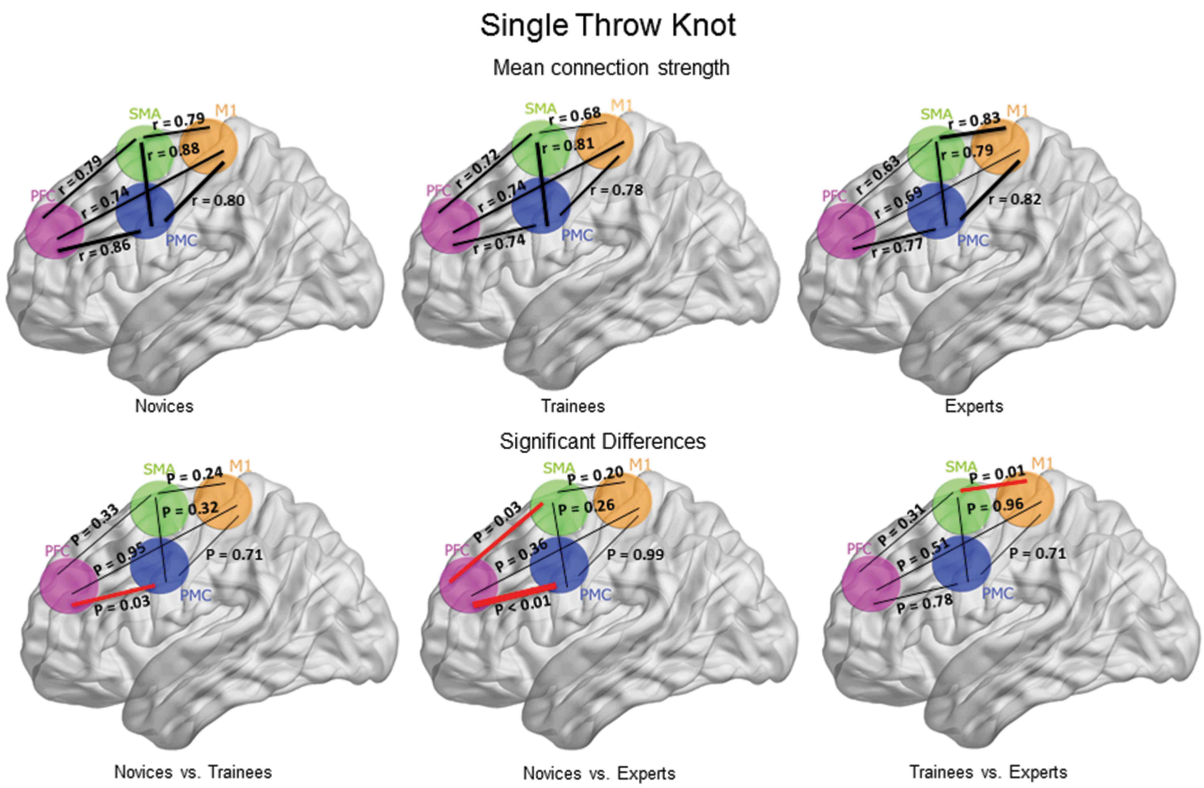

Figure 7.7. Sagittal brain views of learning-related changes in connectivity for the single-throw know subtask. The three sagittal brains at the top display the mean Rv connectivity strength, while the following three at the bottom show the significance strength of the statistical test. Areas are depicted as PFC (magenta), SMA (green), M1 (orange) and PMC (blue).

7.3.3. Classification of operator skill level

Tables 7.2 and 7.3 show the results of the classifier using the "session based networks'' (see Table 7.2) and the 'time-course based networks' (see Table 7.3).

a) Needle insertion	**Precision**	**Specificity**	**Sensitivity**
Novices	0.66	0.83	0.66
Trainees	0.65	0.76	0.76
Experts	0.77	0.92	0.60
Model accuracy: 0.68	F-measure: 0.68	MMCC: 0.52	
a) Double-throw knot	**Precision**	**Specificity**	**Sensitivity**
Novices	0.71	0.84	0.74
Trainees	0.65	0.77	0.70
Experts	0.68	0.89	0.59
Model accuracy: 0.68	F-measure: 0.68	MMCC: 0.50	
a) Single-throw knot	**Precision**	**Specificity**	**Sensitivity**
Novices	0.66	0.85	0.61
Trainees	0.59	0.70	0.73
Experts	0.57	0.85	0.45
Model accuracy: 0.61	F-measure: 0.61	MMCC: 0.41	

Table 7.2. Classification results for the session based networks.

a) Needle insertion	**Precision**	**Specificity**	**Sensitivity**
Novices	0.83	0.80	0.92
Trainees	0.85	0.94	0.78
Experts	0.8	0.8	0.67
Model accuracy: 0.83	F-measure: 0.81	MMCC: 0.71	
a) Double-throw knot	**Precision**	**Specificity**	**Sensitivity**
Novices	0.83	0.90	0.85
Trainees	0.83	0.86	0.86
Experts	0.76	0.95	0.67
Model accuracy: 0.82	F-measure: 0.81	MMCC: 0.70	
a) Single-throw knot	**Precision**	**Specificity**	**Sensitivity**
Novices	0.82	0.82	0.91
Trainees	0.85	0.92	0.81
Experts	0.8	0.97	0.63
Model accuracy: 0.82	F-measure: 0.81	MMCC: 0.70	

Table 7.3**.** Classification results for the time-course based networks.

Classification using *session-based networks* yields a greater number of true positives than false positives in all subtasks, with low MMCC scores (<0.5) for certain trials. In contrast, the *time-course based networks* consistently achieves high MMCC scores (≥ 0.70) across subtasks. For this latter case, the classifier appears to perform equally well in each LS subtask. Overall, novice identification was highly precise (≥ 0.8 for all three subtasks) and sensitive (0.92 for needle insertion, 0.85 for double-throw knot and 0.91 for single-throw knot) whilst trainee and expert identification was precise and specific (e.g. precision for trainees ≥ 0.83 and specificity ≥ 0.86 across all subtasks). The discrimination of experts also exhibits high specificity (≥ 0.95) and precision (≥ 0.76), but with lower sensitivity versus the other operator groups (0.63-0.67). These results suggest that the variability in cortical haemodynamic responses between groups based on experience is potentially greater than the variability in responses across subtasks within a given experience group.

7.4. Discussion and Conclusion

Notwithstanding established models of motor skill levels learning[307] and evidence supporting learning-related changes in brain function[221,299,354], disparate technical performance does not necessarily translate into differences in functional activations on highly complex bimanual co-ordination tasks[29]. Our hypothesis is that differences in technical skill level on tasks of such complexity may be better reflected in differences in functional connectivity within associative and/or sensorimotor brain networks[299,307,354,416,417]. Here, skill level-related differences in frontal lobe connectivity were revealed, summarised as a reduction in connectivity strength between cortical regions involved in the associative network. This suggests that frontal lobe cortico-cortical connectivity on a bimanual co-ordination task varies according to operator skill level and technical skill level, thereby confirming the dynamic nature of coupling in cognitive-motor circuitry. In particular, interaction strength between prefrontal and premotor seeds and other motor-related cortical regions were found to attenuate with skill level. In other words, novices appear to depend on the interactions between associative and motor cortical networks more than experts. Of these frontal lobe interactions, PFC-SMA, PFC-PMC and SMA-PMC connections were consistently stronger in novice operators whereas SMA-M1 interactions remained stable or even increased in expert compared with trainee operators. These data align with models of motor learning implying a high level of integration which decreases with practice in the associative/premotor network and between this and the sensorimotor network[307]. Moreover, the results are consistent with empirical data that suggest large-scale functional re-organisation accompanies motor skill levels learning[221,299,338,354,417].

Among studies investigating spatial and temporal changes in recruitment of brain regions involved in motor skill levels learning[277,307], few have interrogated functional connectivity[221,299,338,354,417,418]. Whilst investigations have explored changes in modularity and allegiance of network nodes[338], variation in functional network econometrics[354] and hierarchical integration within associative and sensorimotor networks[299], only one study specifically interrogated changes in cortico-cortical connectivity during explicit bimanual motor skill level learning[221]. Commensurate with our findings, this study by Sun and colleagues[221] demonstrated enhanced cortico-cortical network connectivity during early phases of explicit bimanual motor sequence learning. Interestingly, connectivity between higher cognitive centres (PFC) and the motor network (PMC) increased only during early learning and subsequently decreased during late within-session learning when subjects improved their performance[221]. Fast motor learning is characterised by increased functional connectivity between dorsolateral PFC and PMC, possibly related to the heightened attentional demands required at this stage of skill levels acquisition[25,307].

Consistent with our findings, studies of slow motor learning suggest longitudinal changes in functional connectivity in premotor associative networks over extended practice schedules[299]. Slow learning (weeks) is accompanied by decreased integration, a metric reflecting functional interactions amongst several brain regions, in a premotor-associative-striatum cerebellar network[299]. The current analysis both supports and extends these findings, suggesting that attenuation in associative cortico-cortical network connectivity is durable across many years of practice, i.e. from novice to trainee. More importantly, it appears that progression from intermediate (trainee) to advanced (expert) phases of skill level execution are potentially independent of further changes in connectivity strength within this associative cortico-cortical network. Contrary to variation observed in the strength of connectivity between prefrontal and premotor regions associated with skill level, SMA-M1 connectivity strength was more stable and skill level-related differences were less apparent. Despite models of learning suggesting that integration in the sensorimotor network is anticipated to increase[307], stability in the strength of SMA-M1 integration across motor skill levels learning has been observed previously, further supporting the results of the current analysis[299]. It is plausible that gradual on-going refinements in complex motor skill levels do not require sustained or progressive increase in SMA-M1 connectivity even if the representation of the trained task continues to gradually expand in these brain regions in association with slow skill levels learning[318].

Moreover, correlation data obtained at one-second intervals of the time course were subsequently used as an input to a LS-SVM algorithm towards automated discrimination of operator skill level. Overall, the approach was found to precisely classify skill level, although it was significantly influenced by the subset of connectivity data under consideration. Specifically, local (within-region) connectivity significantly improved classifier performance. Passive Brain Computer Interfaces (pBCI) offers the potential to feedback implicitly derived information regarding user states. In this regard, through immediate automated categorisation of short epochs of brain inputs, optical neuroimaging data if appropriately decoded may be interfaced with a learner toward improvements in motor skill levels training[255]. Previous research on fNIRS-BCI has focused on modelling patterns of signal amplitude[403-405], failing to incorporate functional connectivity data. Here, discriminatory patterns of operator skill level based on functional connectivity have been exposed through the multi-class machine learning approach, and were not revealed with a conventional analytical framework[29,259,260,265]. The algorithm for classification of operator skill level states has proven to be accurate, sensitive and specific for detection of operator skill level during LS, regardless of the subtask. This suggests that the algorithm is capable of capturing experience-related differences in frontal lobe connectivity equally well in each task phase. Whilst each LS subtask has a unique and specific goal, the rudimentary bimanual co-ordination required to achieve these goals may be similar and hence generic learning-related differences in frontal lobe connectivity may exist, thereby aiding discrimination. Moreover, the results highlight the importance of within-region correlations in discrimination of operator skill level. Finally, classification results using M1 connectivity appeared inferior to the results obtained from other frontal regions possibly owing to a relative paucity of channel information and hence a smaller number of nodes. More importantly, the fact that classifier results were not consistently superior for any specific ROI suggests that the entire frontal lobe is equally informative.

When results of the classifier are provided online, these may be used as a pBCI to feedback data regarding skill level to the operator in the hope of improving the cognitive abilities required to execute the task. Automated classification of skill level based on cortical connectivity rather than technical performance or abstracted end-points may serve to improve skill levels learning by interfacing decoded brain data to the learner, trainer or team. Incorporation of instrument motion tracking in conjunction with brain function may further improve workflow segmentation and proficiency detection. Future work will focus on measuring the impact of neural feedback training to determine if learning and performance can be improved. In this way, online neural feedback regarding operator skill level may minimise the dependency on human assessors allowing learners to rapidly access objective assessments, thereby facilitating self-directed training. In addition, we acknowledge potential confounds due to individual characteristics that were not controlled for, such as the potential influence of intellectual ability (IQ), which could be of interest to consider in future studies. Likewise, incorporating electrophysiological signals that overcome some of the limitations of optical imaging technology such as inferior temporal resolution may improve performance and be more practical for pBCI in surgery. Finally, future research could address the limitations associated with a cross-sectional approach by implementing longitudinal studies to extend on the present findings.

In summary, the current study highlights differences in frontal brain connectivity underpinning operator skill level on a complex bimanual co-ordination task. Operators well versed in bimanual co-ordination appear to depend less on associative and premotor connectivity than novices. Moreover, the disparity in learning-related changes in connectivity facilitates an algorithmic solution toward online classification of operator skill level which is both highly accurate and precise. Automated classification of operator states may facilitate interfaces designed to enhance learning, performance and patient safety.

Chapter 8

Conclusion and Future Perspectives

8.1. Achievements of this thesis

Technical skills in surgery have become more complex since the advent of MIS as mastery requires longer learning curves and under adaptation to multiple sensorimotor distortions. In addition, reforms to surgical careers, cumulative pressures from increased public scrutiny[15] and medical litigation[14] have necessitated multiple stakeholders to implement a framework to facilitate MIS skills and assessment[16]. Current methods of objective assessment of surgical skills primarily utilise speed (time) and accuracy, which have been useful in separating novices from experts but may fail to adequately differentiate experts from proficient trainees[80]. Understanding motor learning theory and the changes in cortical brain activation that accompany learning may provide valuable insights into tracking MIS skills acquisition, and differentiating experts from proficient performers. Motor learning is an internal process that occurs within the central nervous system. However, validated assessments of MIS skills acquisition are reliant on surrogate markers that measure technical performance. However, they do not utilise objective physiological markers of cognition, mental workload, nor measure the intensity of attention, concentration and effort.

The motivation of this thesis was to utilise functional neuroimaging to identify objective neurophysiological markers that accompany motor learning of complex bimanual MIS skills. For this purpose fNIRS and laparoscopic suturing were chosen as the investigative modality and motor task respectively. Selection of fNIRS was based on its features of non-invasiveness, portability, resistance to motion artefacts, and its adaptability to the experimental environment. Despite its limited spatial resolution it allows adequate interrogation of multiple cortical sites linked to various stages of motor learning. The MIS task was chosen from the Fundamentals of Laparoscopic Surgery curriculum[26] on the basis of its task complexity and difficulty to master, which enabled better between group classification in technical skills (especially between specialty registrars in training and experts) than has been demonstrated for open surgical techniques[259,260].

There is now a decade worth of literature describing interrogations of surgeon's brain behaviour using a variety of functional neuroimaging modalities. Chapter 4 provides a systematic review on the valuable role of brain behaviour analysis in differentiating between surgical experts and non-experts, evaluating the impact of task complexity[264], surgical technology[269,270], fatigue[261] and determining visuo-spatial ability[256]. In addition neuroimaging is a useful adjunct in tracking technical surgical skills acquisition[257,260] and allows comparison of different training regimes[253-255,272]. However, prior to this thesis, data regarding functional neuroimaging results acquired during laparoscopic surgery from other groups did not parallel motor learning theory regarding the role of the prefrontal lobe in novel task performance. Indeed, the magnitude of PFC activation during initial execution of complex MIS tasks[29] amongst novices did not resemble that of simpler open surgical tasks[259] or follow established motor learning principles[25]. However several methodological flaws were identified in the design of the study[29]. The complex task was not deconstructed, spatial information on the frontal lobe was not provided and temporal demand during the task was induced. To overcome this gap in understanding, subsequent chapters investigated frontal lobe neuroplasticity, for a complex bimanual laparoscopic task using a cross-sectional and longitudinal design. The task was deconstructed based on interference in bimanual co-ordination, temporal demand was not induced and most importantly multiple motor associated regions in the frontal lobe were interrogated for patterns of functional segregation and integration.

Although fNIRS has been previously used to measure brain function during surgical skills acquisition[253,254,258-263,361,395,419], to our knowledge this is the first evaluation of complex MIS skills acquisition by interrogation of multiple cortical regions situated in the frontal lobe associated with motor skill learning. In Chapter 5, a cross-sectional design was used to explore disparity in regional cortical activation and small-world index a measure of brain connectivity. In relation to expertise, technical performance improved and subjective mental workload decreased respectively. The task recruited the greatest PFC response among novices and the least among experts, commensurate with motor learning literature[25,245,277,279], whereby the PFC is thought to function as an area for attention serving to support novel task demands[300].

Trainees and experts displayed greater engagement of PMC and SMA regions that are responsible for storage of practiced sequential motor skills, which is also commensurate with motor learning of either sequence or under sensory adaptation[25,246,312,420-422]. Although experts displayed comparatively low level PFC activation, sub-tasks of laparoscopic suturing were compared on the basis of directional interference. Commensurate with motor learning literature, anti-phase bimanual sub-tasks (laparoscopic knot –tying) evoked a greater PFC response than the in-phase sub-task (laparoscopic needle manipulation).

In Chapter 6, a longitudinal design was employed to display practice induced neuroplasticity that progressed towards the pattern of brain behaviour exhibited by expert surgeons. Distributed training over a fortnight provided to laparoscopic novices led to significant improvements and stabilisation of technical performance up to expert levels. Commensurate with motor learning literature PMC, SMA and M1 regions showed progressive increase in cortical activation. However, improved technical performance was not accompanied with corresponding decrease in prefrontal activation. This interesting finding suggests that although expert benchmarks were achieved for technical performance based on time and accuracy, unlike experts who have attained automaticity a high degree of attention was required to support performance. These finding demonstrate the value of functional neuroimaging as an adjunct to assess expertise in motor skills and the effects of overtraining on PFC activity warrant further scrutiny.

In Chapter 7, functional connectivity was used to explore variations in frontal lobe activity on the basis of expertise. Commensurate with concurrent motor learning literature[221], associative cortico-cortical motor network located within the frontal lobe attenuated with increase in skill level suggesting that novices were more reliant on interactions between associative and motor cortical networks. Correlation data inputed into a least-squares support vector machine algorithm provided precise automated classification of expertise-level. To our knowledge this is the first study to employ functional neuroimaging to automatically detect user status from cognitive behaviour of surgeons.

8.2. Future Perspectives

Findings of this thesis represent early, albeit important steps towards understanding neuroplasticity that underlies complex MIS skills acquisition. Several questions outlined below remain unanswered, warranting further scrutiny and forms the basis of future research in this field towards technical performance enhancement.

8.2.1. Effect of over training on PFC neuroplasticity

In Chapter 5 by utilising a cross-sectional design, novices exhibited greater recruitment of PFC than experts suggestive of attentional demand during novel task performance. However in Chapter 6, a longitudinal study where despite novices achieving expert levels of proficiency in task performance, persistently high engagement of PFC observed was suggestive of failure to off-load attentional demands. Persistently raised PFC activity may either be attributed to the challenges faced in proficiently executing a complex bimanual task such as laparoscopic suturing or other cognitive factors such as reward related PFC activity. This raises the question at what point during training will PFC activity begin to attenuate for complex bimanual skills? Practice beyond achievement of asymptotic levels of proficient performance is known as overtraining. An extended longitudinal study where participants are serially monitored cognitively after achievement of proficiency in task performance might identify the length of training required before PFC activity begins to significantly attenuate (see Figure 8.1).

Unlike for simple surgical tasks, trainees may spend extended periods of time in each phase of learning a complex bimanual task, thus making the transition process more apparent. Rather than adopting the three-stage Fitts model[42], the five-stage Dreyfus model[2] may better explain the transition from novice to expert. As explained in Chapter 2 despite achieving proficiency, trainees are embodied with the task and still rely on analytic decision-making as opposed to intuitive decision-making ability of experts. The two aforementioned cognitive processes might result in persistent PFC activity observed among trained novices in Chapter 6 as opposed to the lack of PFC activity among experts in Chapter 5, which heralds automaticity.

The relevance of continuous research in this stream is vital to improve current assessment methods of technical competence in surgery, which are arguably one-dimensional and cognitive monitoring can serve as a useful adjunct. Aspects of the task that are particularly demanding can potentially be identified and targeted training can be provided to overcome individual-specific hurdles.

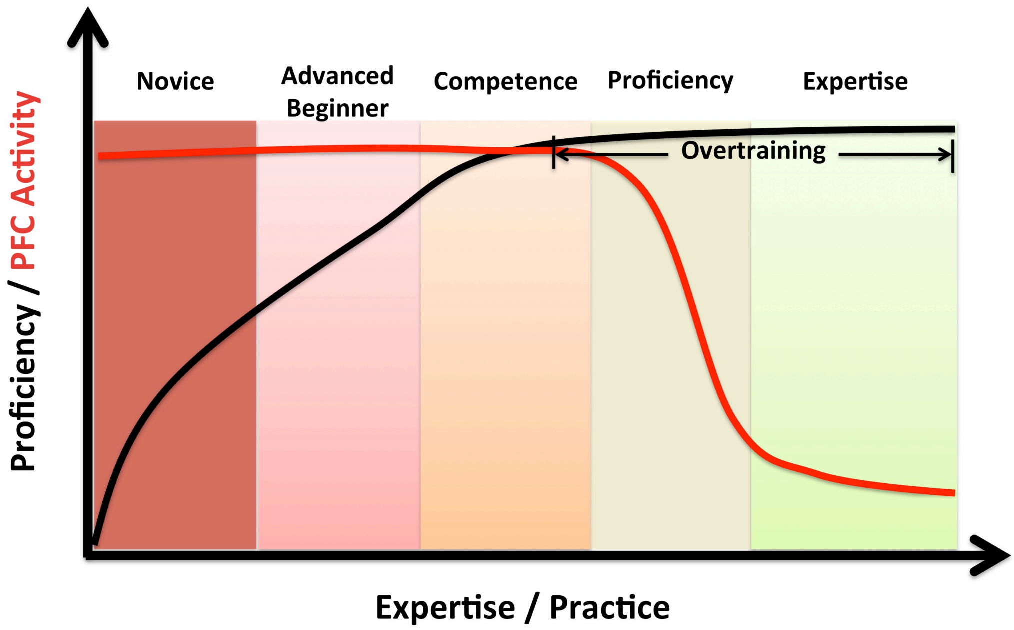

Figure 8.1 Graphical illustration of hypothetical attenuation in PFC activity induced by extensive training. Proficiency (black line) and PFC activity (red line) are plotted in relation to practice or expertise of a complex bimanual MIS task. The five-stage Dreyfus model transitions from novice to expert[2]. PFC activity may only attenuate after a period of overtraining.

8.2.2. Effect of feedback on motor learning

The type of practice and feedback play an important role in improving motor performance, consolidation and retention of practiced motor tasks. Retention of practiced motor tasks is of particular relevance as trainees often spend a period of time practicing certain operations before moving on to another post where they train to perform different types of operation, as a result are at risk of forgetting the practiced task. Surgeons develop technical skills under various feedback conditions, where good performance is rewarded with positive feedback and bad performance results in negative feedback. Although positive, negative and neutral feedback result in improvement in behavioural performance, underlying cognitive processes may vary. Identifying disparity in the neural substrates engaged may provide the key to understanding the preferential benefits of positive feedback on consolidation of practised motor task as opposed to forgetting under negative or neutral feedback[356].

The effects of differential feedback during skills learning in open[255] and laparoscopic surgery[272] have been undertaken, however comparison between effects of positive, negative and neutral feedback remains to be explored. Research in this arena might provide important neuroscientific basis to guide surgical mentors chose the right type of feedback dispensed to trainees, especially in an environment already burdened by the mounting challenges towards improving surgical education.

8.2.3 Automated Classification, Brain Computer Interface

The use of more channels to interrogate a greater number of functional brain regions in this thesis have aided expertise classification based on brain activation and connectivity, yet the analytic process remains complex. Machine learning algorithms enable interpretation of complex high-density neuroimaging data by identifying multivariate relationships between multiple functional brain regions as demonstrated in Chapter 7. In addition, features from neuroimaging data enabled automated classification that has the added advantage of providing instant feedback. A major drawback of surgical assessments like OSATS or GOALS is the unavailability of expert surgeons to assess performance and feedback provided is often delayed.

The results of automated classification generated from neuroimaging data has the potential advantage to be instantly fed back to the trainee providing the stimulus for further training. As a result feedback would be objective and forms an implicit loop in the form of a passive brain computer interface (see Figure 8.2). A number of studies have demonstrated the feasibility of implementing fNIRS data to form a BCI[403,405,423-425], although the majority have been devoted towards rehabilitation processes in individuals with movement disorders. Research towards improving surgeons' dexterity via auto-regulated neural feedback has the potential to be adopted in other domains of movement science that incorporates healthy subjects such as sports science.

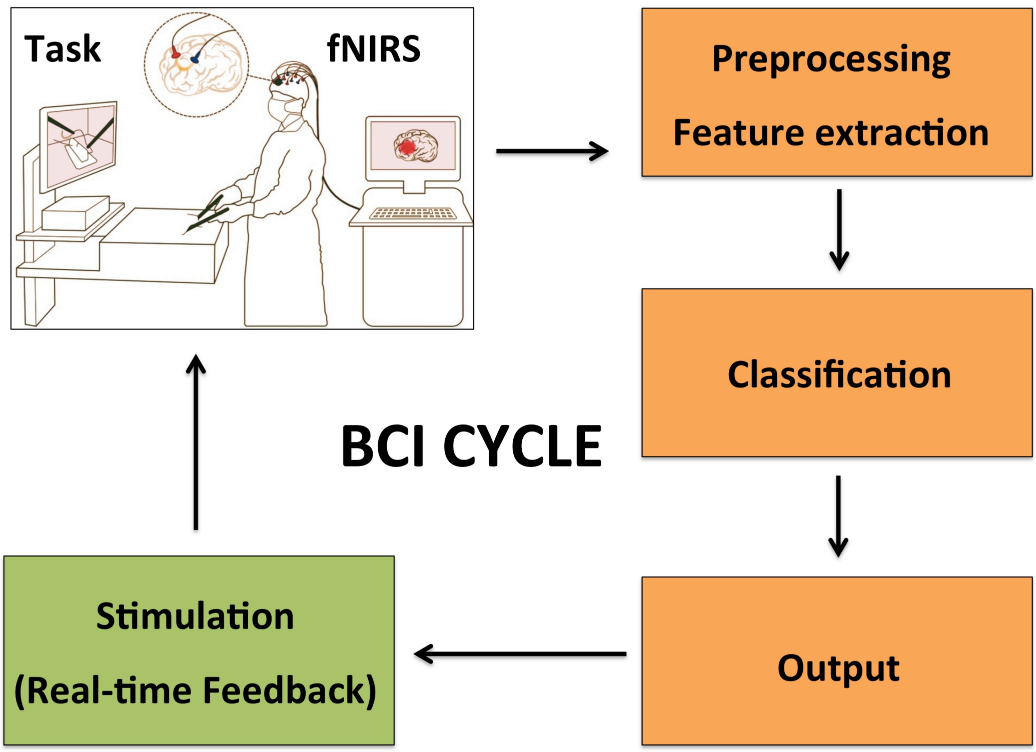

Figure 8.2. Schematic illustration of passive BCI cycle providing implicit feedback to the trainee based on interpretation of task induced neurophysiological response.

8.2.4 Dual imaging – EEG; fNIRS

A major limitation of fNIRS towards brain computer interfacing is its dependence on the haemodynamic response to neural activation, which has a lag of approximately 6 to 8 seconds. Ultimately this results in a delay in feature extraction necessary for classification and production of output in the BCI cycle (see figure 8.2). EEG a direct neuroimaging modality provides excellent temporal resolution but poor localisation of neural responses. However fNIRS and EEG complement each other and overcome their individual limitations towards BCI application in a real-world setting[426].

In Appendix A, an exploratory study demonstrates the ability to detect user's cognitive state from EEG data of surgeon's monitored during performance of surgical tasks. LDA, a machine-learning algorithm enabled detection of whether the user was performing a low-load (open hand knot-tying) or a high-load task (laparoscopic knot-tying). The potential applications of BCI in surgery are vast. Detected brain signals can be used to control instrumentation via an active BCI format, however it would require rigorous testing and ethical approval for understandable patient safety concerns. On the other hand real-time identification of a surgeon's cognitive status might serve as an alternative communication channel to the rest of the surgical team for example in reducing distractions that frequently occur in the OR. Surgeon's whilst performing a cognitively burdensome aspect of an operative task, can potentially be detected in real-time via available wireless neuroimaging sensors and then conveyed seamlessly to the rest of the surgical team (change in tone of background OR lights) thereby promoting a change in the behaviour of the operative team to facilitate a conducive environment.

In conclusion, in this thesis fNIRS has been used to identify neuroplasticity underlying acquisition of complex bimanual motor skills required in MIS. Monitoring multiple motor regions located in the frontal lobe enabled identification of expertise dependent disparity in cortical activation patterns commensurate with motor learning literature. Interaction between brain regions via functional connectivity analysis aided automated expertise classification. Furthermore, a longitudinal study revealed training induced neuroplasticity of secondary motor regions but persistence in PFC engagement warrants further research into examining the effects of overtraining in off-setting attentional demands. Further research is also necessary towards identifying the effects of feedback on motor learning and application of brain-computer interfacing.

Appendix A

Automated Task Load Detection with EEG: Towards Passive Brain-Computer Interfacing

The content from this appendix has been accepted for publication in:

Zander T, Shetty K, Lorenz R, Leff DR, Krol LR, Darzi AW, Gramann K, Yang GZ. Automated Task Load Detection with Electroencephalography: Towards Passive Brain-Computer Interfacing in Robotic Surgery. *Journal of Medical Robotics Research* (Accepted May 2016).

A.1 Introduction

The last thirty years have witnessed a radical transformation in the operative environment with the introduction of Minimally Invasive Surgery (MIS), propelled by patient demand, smaller operative incisions and faster recovery. However, surgeons take significantly longer to reach proficiency in MIS[427] and find it cognitively more burdensome compared to traditional open operative procedures[10]. For example, open hand knot tying (OHKT), a routinely performed surgical manoeuvre, is far less challenging than knot-tying performed in minimally invasive surgery (MISKT), and an inability to perform MISKT efficiently limits the surgeon from performing advanced MIS procedures[27]. This is attributed to poorer instrument ergonomics such as the loss of depth perception, reduced degrees of freedom of movement, amplification of tremors from using long instruments, a lack of tactile feedback, and paradoxical movements as a result of the fulcrum effect[10]. Additionally, the influx of new technology that supports MIS necessitates the operators' vigilance to attend to auditory alarms that alerts the surgeon to a faulty technical device or declining status of the monitored patient. If an alarm suggesting failure of a device is undetected, it may cause potential harm to the patient.

Surgeons are unique amongst doctors because they are often required to make decisions based on the presentation of the real-time problem whilst operating, and therefore should not just possess the technical capability but also the cognitive resources to deal with unanticipated scenarios[21]. Moreover whilst operating, surgeons on average are interrupted 13.5 times[428] and at times these interruptions warrant immediate decisions on management of a critically unwell patient outside the theatre. The impact of such secondary tasks (decision-making, detection of sensory stimuli) has been observed to degrade performance of the primary task (technical performance) albeit to a lesser extent in experts who have presumably achieved automaticity[79,80,241,429]. According to resource theory, humans have a finite pool of resources that can be allocated or shared across tasks[430] with the assumption that the more challenging a task, the greater the resources it consumes. Therefore, a cognitively challenging surgical task can not only impair situational awareness but also degrade optimal performance, which ultimately could jeopardize patient safety. To enhance surgical ergonomics with the view of improving patient safety it is imperative that we be able to characterize these variations in cognitive demand, brought about by different tasks or different contingencies during the same task.

Compared to traditional cognitive state measures such as behavioural correlates[431] or subjective questionnaires such as the NASA Task Load Index (NASA-TLX)[432] the analysis of neural mechanisms[433] has a number of advantages. The measurement can be done continuously and in real time, i.e. in the very moment the relevant events take place, and the measurement does not interfere with the actual task. Also influences of memory, e.g. primacy and recency effects, do not play a role in such measurements. Most importantly however, measures of brain activity may provide more detailed and fine-grained insights into the state of the surgeon than traditional measures. Neural correlates may be used to accurately identify current tasks, but potentially also carry information on e.g. task load, attention, and error detection. Studying brain behaviour in surgeons enables the impact of novel technologies on operators' cognitions to be assessed. Non-invasive neuroimaging technologies such as electroencephalography (EEG)[272] and functional optical brain imaging have previously been applied to assess technical expertise[259] skills acquisition[29,259], cognitive burden[395] and fatigue[261]. These technologies can be applied to assess brain dynamics underlying cognitive processes even in actively moving participants[434,435].

Robotic MIS platforms can conceivably acquire and learn, in situ, operator-specific motor and cognitive behaviour through human-robot interaction. This novel concept termed 'perceptual docking'[436] can be realised from emerging multimodal sensing and feedback rather than one aspect of surgeon-robot interaction. Surgical robotic platforms with tremor filtration capabilities, image magnification and improved actuation offer a high degree of surgical precision and may offload a degree of cognitive burden placed on the operator during MIS. Critically, analyses of operator brain function when synergized with online learning algorithms may enable the robot to benefit from human surgical intelligence and learn to better assist the surgeon, preventing errors and enhancing patient safety. Brain activity may provide an additional seamless communication channel between the surgeon and robot, and thus improve the robot's understanding of the cognitive challenges faced by the operator. This secondary, implicit interaction loop would provide valuable information to the robot, without demanding extra cognitive load from the operator[437]. Technology that adapts to the users in this way based on their brain activity can be termed neuroadaptive technology, and the required form of communication between the brain and the machine can be implemented using passive brain-computer interface (BCI)[438].

In order to evaluate the potential value of passive BCI to robotic surgery we employed an EEG-based BCI system to detect tasks and assess task load over surrogate measures that have been employed in other studies. Traditionally, BCI is defined as a non-muscular communication and control channel for people suffering from diseases that disrupt the neural pathways through which the brain communicates with and controls its external environment[439]. Brain activity associated with the user's intent is measured and translated in real time into control signals for communication systems or other external devices. A recent development within the field of BCI broadens this general BCI approach by substituting the user's volitionally generated command (e.g. intentionally imagined hand movements) with passively conveyed implicit information[438,440]. Such a passive BCI derives its input from brain activity arising without the purpose of voluntary control, e.g. spontaneous activity indicative of task-induced cognitive or affective states[438,441]. This activity reflects covert aspects of the user's state. Therefore, it carries task-relevant information, and can thus be used to goal directedly enrich the human-machine interaction without requiring any overt communication from the user[438]. This extends the potential field of application of BCI technology to include users without disabilities[439], for example, in critical working environments where mental state measurements can provide tangible benefits. Indeed, passive BCIs have proven to be a valuable tool for detecting cognitive load[442] and working memory load[443]. These effects have been demonstrated under laboratory conditions[444] as well as for more complex and natural tasks such as simulated flight[445].

The vast majority of these studies evaluate spectral power differences in certain frequency bands that have shown to be sensitive to different cognitive load conditions. In particular, an increase of frontal theta activity is observed while parietal alpha power decreases as more cognitive resources are allocated to the task[444,446-448]. Hence, we hypothesize that passive BCI can be used to reliably differentiate between OHKT and MISKT, and that during performance of technically challenging tasks (MISKT, dual-tasking) an increased frontal theta activity with concurrent decreased parietal alpha activity will be observed in the EEG.

A.2. Methods

A.2.1. Participants

Nine participants were recruited from Imperial College London after seeking consent and screening for neuropsychiatric illnesses. Participants comprised of eight residents (PGY3-8) and one attendee aged 32.2 ± 3.3 years (mean ± S.D.). All participants underwent 15 minutes of mandatory practice for warm up of the relevant surgical tasks.

A.2.2. Experimental Tasks and Conditions

The primary task was to perform one set of OHKT and MISKT alone for a fixed period of 200 seconds. The OHKT task involved the formation of as many reef knots as possible within 200 seconds using suture material (2/0 Polysorb) on a bench knot-tying trainer (Ethicon Ltd, Sommerville, New Jersey, USA) as illustrated in Figure A.2. MISKT was performed in a laparoscopic box trainer (iSurgicals, UK) using 2/0 Polysorb suture material with laparoscopic needle holders (model 26173KC; Karl Storz GmbH and Co, Tuttlingen, Germany) as illustrated in Figure A.1.

In order to simulate a more realistic setting, a secondary auditory task was added to the experiment. In addition to performing OHKT/MISKT, participants were required to simultaneously count the number of high-toned beeps randomly introduced amongst a series of low tone beeps (secondary task). The auditory stimuli were generated at a standardized decibel level from a speaker placed at a set distance away, on the right side of the participant (simulating a stationary auxiliary device). A tone was presented every second. 20% of tones were high-toned target stimuli.

Figure A.1. A participant wearing the mobile EEG cap. Data recorded at each electrode is transferred wirelessly to the amplifier connected to a standard PC. The participant is performing the MISKT task in a laparoscopic box trainer.

There were five experimental conditions: Secondary (auditory) task only, primary (OHKT, MISKT) task only, and dual-task OHKT and MISKT plus auditory. Following the warm-up phase, the secondary task was performed once for a baseline measurement. The other four conditions were performed in three sets (i.e. 12 sets in total), the order of which was randomized to minimize learning and other temporal effects such as fatigue.

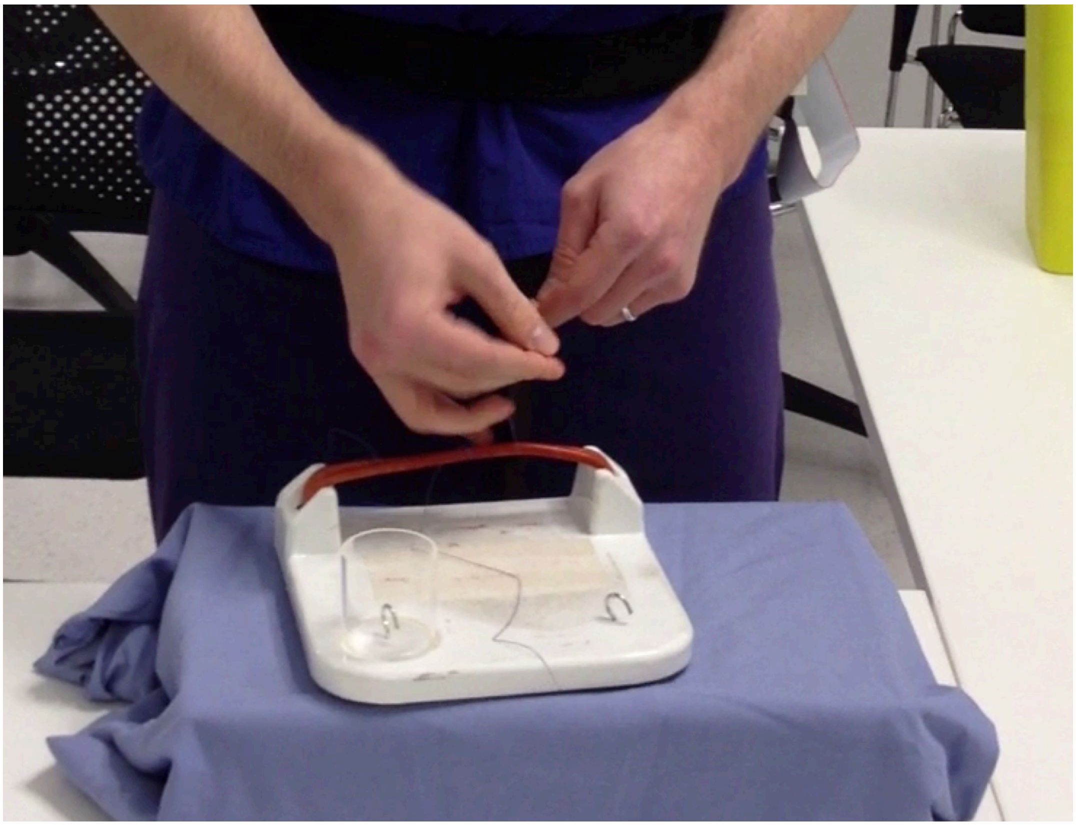

Figure A.2. An example of the OHKT task on the bench knot-tying trainer.

A.2.3. Subjective Measure: NAXA-TLX

Subjective cognitive load of the baseline, mono, and dual tasks were all assessed by the NASA-TLX questionnaire[432]. After each set, participants provided ratings on six sub- scales (mental demands, physical demands, temporal demands, own performance, effort, and frustration), which were then combined into a single task load index, ranging from 0 to 100.

A.2.4. Objective Measure: EEG Set-Up and Processing

32-channel EEG was recorded from all subjects with a BrainAmp system (Brain Products GmbH, Gilching, Germany). Data was transferred wirelessly between cap and amplifier by a BrainAmp MOVE system (Brain Products GmbH, Gilching, Germany), in order not to restrict the free movement of the participants (see figure A.1). Due to technical failure, EEG data from the first three participants was discarded, leaving six participants. This is the minimum required number for statistical significance in our tests.

A passive BCI was set up using the BCILAB toolbox[367] with the intention to discriminate continuously, based on the EEG, whether a participant is working in OHKT or MISKT mode. In brief, each second of the continuous EEG data was transformed into a set of features of lower dimensionality. The transformation was optimised to discriminate between the two primary conditions (OHKT/MISKT), allowing continuous task detection. We used data from the dual-task sessions, as these represent the most realistic condition. Specifically, features were extracted by the spectrally-weighted Common Spatial Patterns (SpecCSP) method[449] and classification was done with a Linear Discriminant Analysis (LDA)[450] regularized by shrinkage[451]. SpecCSP generated 16 pairs of spatiotemporal filters for each participant based on all data recorded for this participant in the dual-task session. The spatial and temporal parameters of these filters were optimized iteratively to optimally discriminate between the two classes (OHKT/MISKT) based on brain activity including the theta and alpha bands (5–18 Hz). 32-dimensional features were generated by projecting one second of data recorded after the onset of each tone with each of the filters generated by SpecCSP. This resulted in 512 normally distributed features per class. LDA is the optimal classifier for this decision problem, as it provides an optimized decision plane and suffices a very low Vapnik-Chervonenkis (VC) dimension[452].

An estimate of the online (real-time) accuracy of the resulting classifier was derived by calibrating the classifier on one part of the data, and applying it to the remaining part. Thus, for each second of the data that was not used for calibration, the classifier indicated whether or not the surgeon was at that time performing OHKT or MISKT. This was done a number of times, calibrating on and applying to different parts of the available data. The resulting accuracy indicates the percentage of correct indications. Specifically, we used a (5,5)-times nested cross-validation[450] with margins of 5. These margins were selected to guarantee the 2Dness of the features. The outcomes of the 5-fold outer runs, regularized by the one shrinkage parameter derived in the appropriate inner runs, gave the estimates for the reliability of each model. The overall reliability (Estimated Classification Accuracy, ECA) was then given by the mean of these single runs' reliability. The validity of this estimate is supported by the low probability of overfitting of classifiers with low VC dimension, by the fact that the sub-optimal ratio between feature dimensionality and number of trials can be counterbalanced by a well-chosen shrinkage regularization, and lastly by the fact that a nested cross- validation was applied properly.

For an additional accuracy measure, testing how well the classifier performed on data recorded in a different con- text, each participant's classifier, trained on dual-task EEG data, was then tested on the EEG data recorded for this respective participant during the mono-task session. Features were generated as above from each (non-overlapping) second of this session, and classified as being either OHKT or MISKT task periods. This was then compared to the real task labels, resulting in a pseudo-online classification accuracy (POCA), simulating an online application of the classifier on a separate data set. In essence, the SpecCSP patterns generated by the classifier can be interpreted as filters, isolating specific, maximally discriminative processes[449]. They thus identify different cognitive processes, which can be investigated further. To investigate these electrocortical processes underlying classification, EEGLAB[453] was used to cluster patterns resulting from each SpecCSP filter by multiplying with the inverse covariance matrix of the underlying data. These patterns represent the scalp projections of the under- lying generator sources. Clustering is a method to identify sources that consistently aided classification across participants. To this end, the patterns' spatial distributions were first reduced to 25 dimensions by means of PCA and sub- sequently clustered using k-means specifying 28 clusters[450]. Further analysis focused on both the clusters and the individual SpecCSP patterns. The time course of the event-related spectral perturbation (ERSP)[454] and the frequency power spectrum were calculated for 0.5 seconds before the onset of the auditory stimulus until 1.5 seconds after. The timing was chosen to demonstrate repeated stimulus related effects. Permutation tests, corrected for false discovery rate (FDR), were used to test for significant differences between the ERSPs, as implemented in EEGLAB[453].

A.3. Results

A.3.1. Questionnaire Results

The raw questionnaire results are listed in Table A.1. The scores were compared using Wilcoxon signed-rank tests. The alpha level was set at 0.05. The results of these comparisons are listed in Table A.2. It is clear that dual-task MISKT was found to be the most challenging task, followed by dual- task OHKT and mono-task MISKT. The two least challenging tasks were the secondary task in itself, and mono-task OHKT. These two did not differ significantly from each other.

Task	**NASA-TLX (Mean ± S.E.)**
Mono OHKT	16.99 ± 2.85
Mono MISKT	48.79 ± 6.17
Auditory	24.33 ± 6.07
Dual OHKT	51.04 ± 7.43
Dual MISKT	67.55 ± 6.92

Table A.1. Group averaged NASA-TLX scores for each task

Paired Comparison (Wilcoxon Signed-Rank test)	***p*-value**
Dual OHKT vs Mono OHKT	0.018
Dual MISKT vs Mono MISKT	0.012
Dual OHKT vs Auditory	0.017
Dual MISKT vs Auditory	0.011
Mono OHKT vs Mono MISKT	0.012
Mono MISKT vs Auditory	0.025
Mono OHKT vs Auditory	0.553

Table A.2. NASA-TLX scores for tasks were compared (Wilcoxon signed-rank test) and their corresponding p values are tabulated.

A.3.2. BCI Results

A.3.2.1. Classification Accuracy

Table A.3 provides the estimated classification accuracies from the calibration session and the pseudo-online classification accuracies from the test session for each participant. High accuracies in both measures attest to a high reliability of the passive BCI system applied here. As there is no significant difference (tested with a Student's t-test) between the ECA and the POCA measures, it is unlikely that the BCI definition overfitted on unknown factors during calibration. Together, these results provide a promising step towards robust, real-time task detection in the theatre.

Participant ID	ECA	POCA
1	82.4% (6.2%)	72.5%
2	77.8% (11.4%)	77.2%
3	95.5% (4.1%)	91.7%
4	90.7% (5.9%)	95.8%
6	99.7% (0.4%)	97.8%
Average	85.8% (6.8%)	86.6%

Table A.3. Estimated classification accuracy (ECA) and pseudo-online classification accuracy (POCA) for each participant.

A.3.2.2. Contributing EEG Activity

Figure A.3 highlights a number of findings from the neuro-physiological analysis of the EEG activity that contributed to classification. The SpecCSP patterns that were generated and used by the classifier, can be interpreted as isolating different discriminative processes[449]. Across participants, the classifier identified a number of consistent patterns reflecting task-relevant processes. Based on the topography of their scalp projections, these processes can be localised within the three-dimensional space covered by the electrodes. We highlight four clusters that illustrate the variety of cortical and non-cortical information that passive BCI systems can potentially make use of to enable real-time user state measurement.

The top row of each panel shows the mean spatial distribution of the cluster patterns. Below that, the cluster's mean 2-second ERSP up to 25 Hz, time-locked to the auditory stimuli, is shown for both the OHKT condition (blue) and the MISKT, higher task load condition (green), as well as a plot highlighting significant differences between the two computed by permutation statistics. Exemplary single- participant data (frequency spectra or ERSP) is illustrated on the bottom row. Panel a of Figure A.3 shows a cluster associated with ocular activity, as identifiable by the spatial distribution of the weights. The stimulus-locked increase in broad- band spectral power indicates that this cluster entails slow eye movements (horizontal movements associated with the task) as well as high frequency muscle activity related to muscular components of the eye movement in the MISKT condition. The increase in the low frequency range of the individual spectral plot clearly shows an increase of blink activity during higher task load.

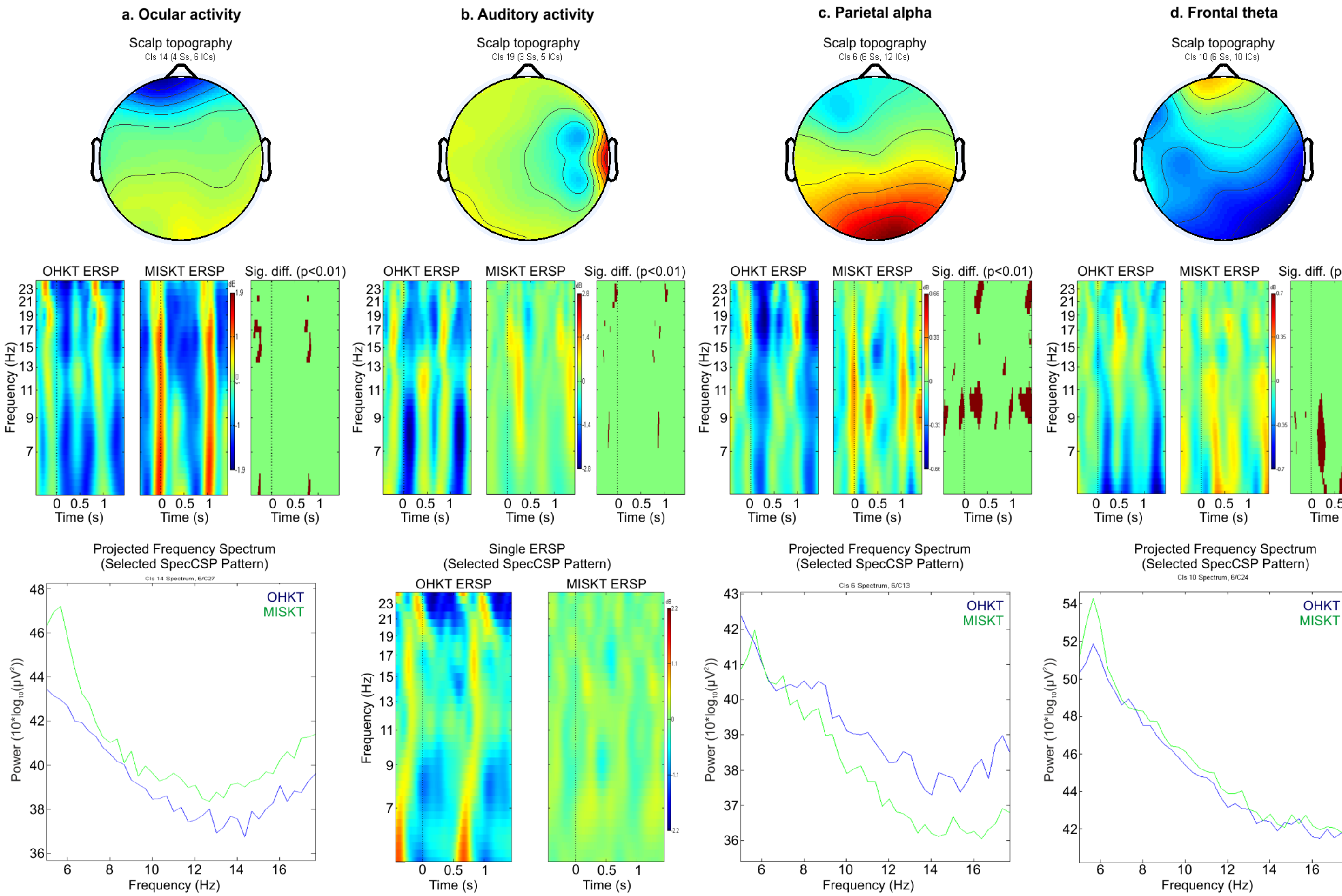

Figure A.3. Analysis of a number of patterns consistently identified by the BCI across participants show different activities that contributed to classification. Top row: Representation of clustered SpecCSP pattern means showing the topography of the identified influences. Middle row: Cluster mean ERSP time-locked to tone onset in both conditions, illustrating the spectral activity. Bottom row (a,c,d): Frequency spectrum projected from a single pattern illustrating non-time-locked differences between conditions. Bottom row (b): Selected component ERSP, illustrating individual differences within this cluster.

A cluster most likely representing the processing of auditory information, with high weights near the right auditory cortex, is presented in panel b of Figure A.3. This cluster's mean ERSP reveals stronger desynchronization in a wide frequency range up to 25 Hz for the lower task load condition (OHKT) time locked to the auditory stimulus onset. This is reflective of increased processing of auditory information (i.e. more resources available for auditory processing) during lower task load levels. However, these differences are only marginally significant and are present only for very restricted periods of time. The bottom part of this panel shows the ERSP projected by a single pattern for one participant, revealing strong differences for this individual.

Increased blink activity and modulated auditory processing can be well explained given the current conditions, but may not be specifically related to task load in general. For more generic, context-independent effects, we had hypothesised to find changes in parietal alpha and frontal theta activity. These effects of task load are further investigated in panels c and d of Figure A.3. Panel c represents a cluster of SpecCSP topographies with high positive weights over occipito-parietal areas. The mean ERSPs demonstrate, contrary to the hypothesis, increased alpha desynchronization in the low as compared to the high load condition. However, individual spectra of certain participants, as e.g. illustrated in the bottom row, do demonstrate the expected, opposite effect in the alpha band. Here one must note that the topographic plot shows two separate weight centres over frontal and posterior areas, indicating that this component may reflect two, probably related processes. The observed effects are time-locked to the onset of the auditory stimulus, and may thus partially represent a cognitive process related to the secondary task, rather than the expected more generic load effect.

Panel d, finally, shows three plots generated from a cluster representing theta activity, most likely the frontal theta component described in[444]. The mean topographic distribution of the patterns' SpecCSP weights shows a fronto-posterior polarity inversion with positive fronto-central weight and negative weights over the posterior cortex, with a slight right-lateralization. In line with our hypothesis, the ERSPs show a clear increase in the theta band during the MISKT condition. The frequency spectrum projected from a single SpecCSP pattern demonstrates how the energy in the theta band increases with higher task load. There also are supportive effects in the alpha band, but these fail to reach significance here.

A.4. Discussion and Conclusion

This study simulated a realistic operative environment using both a traditional and an MIS version of an established surgical task. These tasks were combined with an additional auditory task to increase load. A noninvasive EEG system that allowed free movement was used to record the participants' brain activity throughout the experiment. Based on this data, a passive BCI was calibrated in order to automatically discriminate between periods of OHKT (lower load) and MISKT (higher load). We had hypothesised that a) this discrimination was reliably possible, and b) that it would take parietal alpha and frontal theta effects into account.

In similar simulations of high-risk scenarios, such as military operations or air traffic control, using fNIRS optical brain imaging, an increase in task complexity was linked to an increase in prefrontal cortex activity[314,340]. The pre- frontal cortex is a region within the frontal lobe associated with e.g. executive control and error detection. When an auditory secondary task is introduced, cognitive load increases, leading to a reduction in alertness, as e.g. observed by prolonged brake reaction times in a driving task[455]. Similarly, the impact of auditory distractors on the surgeon whilst operating, as used here, may increase load especially during critical, technically challenging moments or near the end of the procedure as fatigue sets in.

The passive BCI used here was calibrated on data from these dual-task scenarios. It then was capable of dis- criminating between EEG data recorded during mono-task OHKT and MISKT sessions with a high reliability. The cross-validated accuracy estimations for each participant are not significantly different than the accuracies achieved on the test set, strongly supporting the BCI's ability to automatically discriminate between conditions even in real time. In practice, EEG does not only contain brain activity, but also first order artefacts related to bodily activity (e.g. muscle tension, eye movements) and second order artefacts (e.g. external electromagnetic fields, unstable electrodes, mechanical forces working on the cables or electrodes[456,457]). Because of this, a range of different types of activity is available for the BCI to aid the discrimination between conditions, both cortical and non-cortical. Further analysis of the features underlying the BCI's operation pro- vides information with respect to the selected dynamics that contributed to classification. Indeed, a number of different activity sources are reflected in the features.

The processes underlying the demonstrated successful discrimination include cortical activity in line with previous findings related to alpha and theta activity, as hypothesised, although some individual differences remain. We also found other processes that may not be uniquely related to generic task load, but may still be relevant to track in real time, as BCI technology allows. The analysis presented here does not definitively exclude the possibility that other processes played a role in classification, but the BCI's level of performance across training and testing conditions attests that the classifier was unlikely to be overfitted on circumstantial activity. It is clear that BCI enables a number of different task-relevant cortical and non-cortical measurements to take place simultaneously.

A focus on cortical as opposed to artefactual activity is advantageous as it is likely to be more robust across different situations and context. Figure A.3 shows that relevant cortical information was indeed available, and taken into account for classification. It is thus possible to monitor and detect task-relevant cortical activity from the EEG of a surgeon actively involved in a surgical procedure.

An ability to detect task load in real time enables a number of potentially important improvements to be implemented to aid the surgeon and the operative team. Regarding robotic surgery, this pBCI framework has implications that may help maximize patient safety. For example, at times of escalating load burden, the system may initiate active dampening which may change the ratio of instrument motion scaling, or engage dynamic active constraints to prevent instruments from entering critical 'no go' anatomical zones to restrict the chance of injury to vital organs. Mundane tasks may be relegated to the robot as a means of adaptive automation to enable the surgeon to focus solely on critical tasks requiring higher level of decision-making. Such an intelligent system may also be used for surgical training and assessment of technical and cognitive skills, and opens avenues for neural feedback training to improve performance.

A.5. Outlook

This study represents a first proof of principle that passive BCIs can identify task-relevant cognitive processes in established surgery scenarios, and differentiate between different tasks based on continuous EEG data. The estimated and pseudo-online classification accuracies support the idea that passive BCIs may be used for real-time task load detection in the theatre during robotic surgery scenarios. This opens up a variety of applications for BCI technology, and gives a perspective on new types of assistive technology in these and other environments. Aside from task load, other cognitive and affective processes can potentially be assessed using the same BCI methodology, providing richer, more detailed information about the current state of the operator. This information can then be used to automatically improve the interaction between the operator, machine, and operative assistants. Transferring knowledge from neuroscientific research into the field of passive BCI should enable many different cognitive and affective aspects to be detected automatically. Robotic surgery in particular could benefit from information about the current level of attention of a surgeon, or by detecting specific intentions (e.g. the need for a specific tool, which can then be prepared).

Future steps continuing the work presented here could include an investigation of the task specificity of the presented approach. Can it be transferred easily to tasks that are significantly different from OHKT and MISKT? Practical aspects of the set-up may also be investigated: Can the amount of training data be reduced without losing accuracy, and what is the minimum amount of electrodes needed for reliable results? A final, highly interesting step would be to investigate the application of other developments from the field of human-computer interaction and passive BCI to robotic surgery, in particular, passive BCI-based implicit control[458]. Implicit control could open up an additional, goal-oriented communication channel for the surgeon without placing any additional load on them. The fusion of two worlds, that of passive BCI research and robotic surgery shows the potential of leading us into a fascinating new world of brain-based interaction.

References

1. Jasper HH. The ten twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology* 1958; **10**: 371-5.

2. Dreyfus SED, H. L. A five stage model of the mental activities involved in directed skill acquisition.: University of California, Berkeley, 1980.

3. Censor N, Sagi D, Cohen LG. Common mechanisms of human perceptual and motor learning. *Nat Rev Neurosci* 2012; **13**(9): 658-64.

4. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med* 2013; **369**(15): 1434-42.

5. Hughes-Hallett A, Mayer EK, Marcus HJ, et al. Quantifying innovation in surgery. *Ann Surg* 2014; **260**(2): 205-11.

6. Veldkamp R, Kuhry E, Hop WC, et al. Laparoscopic surgery versus open surgery for colon cancer: short-term outcomes of a randomised trial. *Lancet Oncol* 2005; **6**(7): 477-84.

7. Eurostat. Surgical operation and procedure statistics. 2015. http://ec.europa.eu/eurostat/statistics-explained/index.php/Surgical_operations_and_procedures_statistics#Number_of_surgical_operations_and_procedures (accessed 01/07/2015.

8. Bege T, Lelong B, Esterni B, et al. The learning curve for the laparoscopic approach to conservative mesorectal excision for rectal cancer: lessons drawn from a single institution's experience. *Ann Surg* 2010; **251**(2): 249-53.

9. Dagash H, Chowdhury M, Pierro A. When can I be proficient in laparoscopic surgery? A systematic review of the evidence. *J Pediatr Surg* 2003; **38**(5): 720-4.

10. Berguer R, Smith WD, Chung YH. Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery. *Surg Endosc* 2001; **15**(10): 1204-7.

11. Podnos YD, Jimenez JC, Wilson SE, Stevens CM, Nguyen NT. Complications after laparoscopic gastric bypass: a review of 3464 cases. *Arch Surg* 2003; **138**(9): 957-61.

12. Thompson JE, Egger S, Bohm M, et al. Superior quality of life and improved surgical margins are achievable with robotic radical prostatectomy after a long learning curve: a prospective single-surgeon study of 1552 consecutive cases. *Eur Urol* 2014; **65**(3): 521-31.

13. Purcell Jackson G, Tarpley JL. How long does it take to train a surgeon? *BMJ* 2009; **339**: b4260.

14. Berry W. Surgical malpractice: myths and realities. *J Med Pract Manage* 2006; **22**(1): 50-1.

15. Smith R. All changed, changed utterly. British medicine will be transformed by the Bristol case. *BMJ* 1998; **316**(7149): 1917-8.

16. Szasz P, Louridas M, Harris KA, Aggarwal R, Grantcharov TP. Assessing Technical Competence in Surgical Trainees: A Systematic Review. *Ann Surg* 2015; **261**(6): 1046-55.

17. Kneebone RL. Practice, rehearsal, and performance: an approach for simulation-based surgical and procedure training. *JAMA* 2009; **302**(12): 1336-8.

18. Reznick RK, MacRae H. Teaching surgical skills--changes in the wind. *N Engl J Med* 2006; **355**(25): 2664-9.

19. Datta V, Chang A, Mackay S, Darzi A. The relationship between motion analysis and surgical technical assessments. *Am J Surg* 2002; **184**(1): 70-3.

20. Van Sickle KR, Ritter EM, Baghai M, et al. Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *J Am Coll Surg* 2008; **207**(4): 560-8.

21. Gallagher AG, Satava RM, O'Sullivan G C. Attentional Capacity: An Essential Aspect of Surgeon Performance. *Ann Surg* 2013.

22. Schmidt RAL, T. D. . Motor Control and Learning 5th ed: Human Kinetics 2011.

23. Milton J, Solodkin A, Hlustik P, Small SL. The mind of expert motor performance is cool and focused. *Neuroimage* 2007; **35**(2): 804-13.

24. Friston KJ. Modalities, modes, and models in functional neuroimaging. *Science* 2009; **326**(5951): 399-403.

25. Dayan E, Cohen LG. Neuroplasticity subserving motor skill learning. *Neuron* 2011; **72**(3): 443-54.

26. Fundamentals of Laparoscopic Surgery http://www.flsprogram.org/ (accessed 01/07/2016.

27. Korndorffer JR, Jr., Dunne JB, Sierra R, Stefanidis D, Touchard CL, Scott DJ. Simulator training for laparoscopic suturing using performance goals translates to the operating room. *J Am Coll Surg* 2005; **201**(1): 23-9.

28. Friston KJ. Functional and effective connectivity: a review. *Brain Connect* 2011; **1**(1): 13-36.

29. Ohuchida K, Kenmotsu H, Yamamoto A, et al. The frontal cortex is activated during learning of endoscopic procedures. *Surg Endosc* 2009; **23**(10): 2296-301.

30. Ericsson KA. An expert-performance perspective of research on medical expertise: the study of clinical performance. *Med Educ* 2007; **41**(12): 1124-30.

31. Carter BN. The fruition of Halsted's concept of surgical training. *Surgery* 1952; **32**(3): 518-27.

32. Ericcson KAK, R. Th.; Tesch-Romer, C. The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 1993; **100**(3): 363-406.

33. Cocks M, Moulton CA, Luu S, Cil T. What surgeons can learn from athletes: mental practice in sports and surgery. *J Surg Educ* 2014; **71**(2): 262-9.

34. Ericsson KA. Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. *Academic Medicine* 2003; **79**(10): s70-s81.

35. Intercollegiate Surgical Curriculum Programme. 2016. https://www.iscp.ac.uk (accessed 01/07/2016.

36. O'Grady G, Loveday B, Harper S, Adams B, Civil ID, Peters M. Working hours and roster structures of surgical trainees in Australia and New Zealand. *ANZ J Surg* 2010; **80**(12): 890-5.

37. Willingham DB. A neuropsychological theory of motor skill learning. *Psychol Rev* 1998; **105**(3): 558-84.

38. Bryan WLH, N. Studies on the telegraphic language: The acquisition of a hierachy of habits. *Psychologocal Review* 1897; **6**: 345-75.

39. Snoddy GS. Learning and stability: A psychophysical analysis of a case of motor learning with clinical applications. *Journal of Applied Psychology* 1926; **10**: 1-36.

40. Adams JA. A closed-loop theory of motor learning. *J Mot Behav* 1971; **3**(2): 111-49.

41. Gentile AM. A working model of skill acquisition with application to teaching. *Quest* 1972; **17**: 3-23.

42. Fitts PM, Posner MI. Human performance. Belmont, Calif.,: Brooks/Cole Pub. Co.; 1967.

43. Anderson JR. Acquisition of cognitive skill. *Psychological review* 1982; **89**: 369-406.

44. Anson G, Elliott D, Davids K. Information processing and constraints-based views of skill acquisition: divergent or complementary? *Motor Control* 2005; **9**(3): 217-41.

45. Posner MIS, C. R. R. Attention and cognitive control. In: Information processing and cognition. The Loyola Symposium; 1975; Hillsdale, NJ: Erlbaum.: Solso RL, ed; 1975. p. 55-85.

46. Schneider WS, R. M. Controlled and automatic human information processing: Detection, search and attention. *Psychological Review* 1997; **84**(1): 1-66.

47. Dreyfus SE. The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science, Technology and Society* 2004; **24**(3): 177-81.

48. Dreyfus SED, H. L. Mind over machine: the power of human intuitive expertise in the era of the computer. New York: Free Press; 1986.

49. Mitchell EL, Arora S. How educational theory can inform the training and practice of vascular surgeons. *J Vasc Surg* 2012; **56**(2): 530-7.

50. Moe VF. How to Understand Skill Acquisition in Sport. *Bulletin of Science, Technology and Society* 2004; **24**(3): 213-24.

51. Dreyfus SE. The Five-Stage Model of Skill Acquisition. *Bulletin of Science, Technology and Society* 2004; **24**(3): 177-81.

52. Schmidt RAW, C. A. Motor Learning and Performance A Situation-Based Learning Approach. 4th ed: Human Kinetics; 2007.

53. Crossman ERFW. A theory of the acquisition of speed skill. . *Ergonomics* 1959; **2**: 153-66.

54. Ericsson KAK, R. Th.; Tesch-Romer, C. The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 1993; **100**: 363-406.

55. Stefanidis D, Heniford BT. The formula for a successful laparoscopic skills curriculum. *Arch Surg* 2009; **144**(1): 77-82; discussion

56. Gallagher AG, Jordan-Black JA, O'Sullivan GC. Prospective, randomized assessment of the acquisition, maintenance, and loss of laparoscopic skills. *Ann Surg* 2012; **256**(2): 387-93.

57. Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice distribution in procedural skills training: a randomized controlled trial. *Surg Endosc* 2002; **16**(6): 957-61.

58. Lee TDG, E. D. Distribution of Practice in Motor Skill Acquisition: Learning and Performance Effects Reconsidered. *Research Quaterly for Excercise and Sport* 1988; **59**(4): 277-87.

59. Shea JBM, R. L. Contextual interference effects on the acquisition, retention and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory* 1979; **5**: 179-87.

60. Batting WF. Intratask interference as a source of facilitation on transfer and retention. In: Thompson EFV, J. F., ed. Topics in learning and performance. New York: Academic Press; 1972: 131-59.

61. Naylor JC, Briggs GE. Effects of task complexity and task organization on the relative efficiency of part and whole training methods. *J Exp Psychol* 1963; **65**: 217-24.

62. Dubrowski A, Backstein D, Abughaduma R, Leidl D, Carnahan H. The influence of practice schedules in the learning of a complex bone-plating surgical task. *Am J Surg* 2005; **190**(3): 359-63.

63. Heuer H, Klimmer F, Luttmann A, Bolbach U. Specificity of motor learning in simulator training of endoscopic-surgery skills. *Ergonomics* 2012; **55**(10): 1157-65.

64. Sulzenbruck S, Heuer H. Type of visual feedback during practice influences the precision of the acquired internal model of a complex visuo-motor transformation. *Ergonomics* 2011; **54**(1): 34-46.

65. Solley WH. The effects of verbal instruction of speed and accuraccy upon the learning of a motor skill. *Research Quarterly* 1952; **23**: 231-40.

66. Porte MC, Xeroulis G, Reznick RK, Dubrowski A. Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learning new surgical skills. *Am J Surg* 2007; **193**(1): 105-10.

67. Britt RC, Scerbo MW, Montano M, Kennedy RA, Prytz E, Stefanidis D. Intracorporeal suturing: Transfer from Fundamentals of Laparoscopic Surgery to cadavers results in substantial increase in mental workload. *Surgery* 2015; **158**(5): 1428-33.

68. Gallagher AG, Seymour NE, Jordan-Black JA, Bunting BP, McGlade K, Satava RM. Prospective, randomized assessment of transfer of training (ToT) and transfer effectiveness ratio (TER) of

virtual reality simulation training for laparoscopic skill acquisition. *Ann Surg* 2013; **257**(6): 1025-31.

69. Diesen DL, Erhunmwunsee L, Bennett KM, et al. Effectiveness of laparoscopic computer simulator versus usage of box trainer for endoscopic surgery training of novices. *J Surg Educ* 2011; **68**(4): 282-9.

70. Tanoue K, Ieiri S, Konishi K, et al. Effectiveness of endoscopic surgery training for medical students using a virtual reality simulator versus a box trainer: a randomized controlled trial. *Surg Endosc* 2008; **22**(4): 985-90.

71. Muresan C, 3rd, Lee TH, Seagull J, Park AE. Transfer of training in the development of intracorporeal suturing skill in medical student novices: a prospective randomized trial. *Am J Surg* 2010; **200**(4): 537-41.

72. Borghini G, Astolfi L, Vecchiato G, Mattia D, Babiloni F. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci Biobehav Rev* 2014; **44**: 58-75.

73. Causse M, Peysakhovich V, Fabre EF. High Working Memory Load Impairs Language Processing during a Simulated Piloting Task: An ERP and Pupillometry Study. *Front Hum Neurosci* 2016; **10**: 240.

74. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surg Endosc* 2003; **17**(10): 1525-9.

75. Simons DJC, C. F. Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception* 1999; **28**(9): 1059-74.

76. Beilock SL, Bertenthal BI, McCoy AM, Carr TH. Haste does not always make waste: expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychon Bull Rev* 2004; **11**(2): 373-9.

77. Perkins-Ceccato N, Passmore SR, Lee TD. Effects of focus of attention depend on golfers' skill. *J Sports Sci* 2003; **21**(8): 593-600.

78. Masters RSW. Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology* 1992; **83**: 343-58.

79. Hsu KE, Man FY, Gizicki RA, Feldman LS, Fried GM. Experienced surgeons can do more than one thing at a time: effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons. *Surg Endosc* 2008; **22**(1): 196-201.

80. Stefanidis D, Scerbo MW, Korndorffer JR, Jr., Scott DJ. Redefining simulator proficiency using automaticity theory. *Am J Surg* 2007; **193**(4): 502-6.

81. Hart SG. NASA TLX: Task Load Index. 1988. http://humansystems.arc.nasa.gov/groups/tlx/tlxpublications.html (accessed 09/07/2014.

82. Stefanidis D, Korndorffer JR, Jr., Markley S, Sierra R, Heniford BT, Scott DJ. Closing the gap in operative performance between novices and experts: does harder mean better for laparoscopic simulator training? *J Am Coll Surg* 2007; **205**(2): 307-13.

83. Zheng B, Jiang X, Tien G, Meneghetti A, Panton ON, Atkins MS. Workload assessment of surgeons: correlation between NASA TLX and blinks. *Surg Endosc* 2012; **26**(10): 2746-50.

84. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. *BMJ* 2003; **327**(7422): 1032-7.

85. Reznick RK. Teaching and testing technical skills. *Am J Surg* 1993; **165**(3): 358-61.

86. Gerrig RJ. Psychology and Life: Pearson; 2013.

87. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997; **84**(2): 273-8.

88. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative "bench station" examination. *Am J Surg* 1997; **173**(3): 226-30.

89. Dath D, Regehr G, Birch D, et al. Toward reliable operative assessment: the reliability and feasibility of videotaped assessment of laparoscopic technical skills. *Surg Endosc* 2004; **18**(12): 1800-4.

90. Joice P, Hanna GB, Cuschieri A. Errors enacted during endoscopic surgery--a human reliability analysis. *Appl Ergon* 1998; **29**(6): 409-14.

91. Tang B, Hanna GB, Joice P, Cuschieri A. Identification and categorization of technical errors by Observational Clinical Human Reliability Assessment (OCHRA) during laparoscopic cholecystectomy. *Arch Surg* 2004; **139**(11): 1215-20.

92. Foster JD, Miskovic D, Allison AS, et al. Application of objective clinical human reliability analysis (OCHRA) in assessment of technical performance in laparoscopic rectal cancer surgery. *Tech Coloproctol* 2016; **20**(6): 361-7.

93. Miskovic D, Ni M, Wyles SM, Parvaiz A, Hanna GB. Observational clinical human reliability analysis (OCHRA) for competency assessment in laparoscopic colorectal surgery at the specialist level. *Surg Endosc* 2012; **26**(3): 796-803.

94. Vassiliou MC, Feldman LS, Andrew CG, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg* 2005; **190**(1): 107-13.

95. Kramp KH, van Det MJ, Hoff C, Lamme B, Veeger NJ, Pierie JP. Validity and reliability of global operative assessment of laparoscopic skills (GOALS) in novice trainees performing a laparoscopic cholecystectomy. *J Surg Educ* 2015; **72**(2): 351-8.

96. Gumbs AA, Hogle NJ, Fowler DL. Evaluation of resident laparoscopic performance using global operative assessment of laparoscopic skills. *J Am Coll Surg* 2007; **204**(2): 308-13.

97. Bellorin O, Kundel A, Sharma S, Ramirez-Valderrama A, Lee P. Training model for laparoscopic Heller and Dor fundoplication: a tool for laparoscopic skills training and assessment-construct validity using the GOALS score. *Surg Endosc* 2016; **30**(8): 3654-60.

98. Kurashima Y, Feldman LS, Al-Sabah S, Kaneva PA, Fried GM, Vassiliou MC. A tool for training and evaluation of laparoscopic inguinal hernia repair: the Global Operative Assessment Of Laparoscopic Skills-Groin Hernia (GOALS-GH). *Am J Surg* 2011; **201**(1): 54-61.

99. Zevin B, Bonrath EM, Aggarwal R, et al. Development, feasibility, validity, and reliability of a scale for objective assessment of operative performance in laparoscopic gastric bypass surgery. *J Am Coll Surg* 2013; **216**(5): 955-65 e8; quiz 1029-31, 33.

100. Fraser SA, Klassen DR, Feldman LS, Ghitulescu GA, Stanbridge D, Fried GM. Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system. *Surg Endosc* 2003; **17**(6): 964-7.

101. Fried GMF, L. S.; Ghitilescu, G. A.; Derossis, A. M; Stanbridge, D. . A model for evaluation of laparoscopic skills: is there correlation to level of training? *Surg Endosc* 2001; **15**: S127.

102. Fried GMF, L. S.; Ghitilescu, G. A.; Derossis, A. M; Stanbridge, D. A model for evaluation of laparoscopic skills: is there external validity? *Surg Endosc* 2001; **15**: S128.

103. Vassiliou MC, Feldman LS, Fraser SA, et al. Evaluating intraoperative laparoscopic skill: direct observation versus blinded videotaped performances. *Surg Innov* 2007; **14**(3): 211-6.

104. Peters JH, Fried GM, Swanstrom LL, et al. Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. *Surgery* 2004; **135**(1): 21-7.

105. McCluney AL, Vassiliou MC, Kaneva PA, et al. FLS simulator performance predicts intraoperative laparoscopic skill. *Surg Endosc* 2007; **21**(11): 1991-5.

106. Feldman LS, Hagarty SE, Ghitulescu G, Stanbridge D, Fried GM. Relationship between objective assessment of technical skills and subjective in-training evaluations in surgical residents. *J Am Coll Surg* 2004; **198**(1): 105-10.

107. Shiffer C. American Board of Surgery to Require ACLS, ATLS and FLS for General Surgery Certification. 2008. https://www.absurgery.org/default.jsp?news_newreqs (accessed 01/07/2016.

108. Sanchez-Margallo JA, Sanchez-Margallo FM, Oropesa I, Gomez EJ. Systems and technologies for objective evaluation of technical skills in laparoscopic surgery. *Minim Invasive Ther Allied Technol* 2014; **23**(1): 40-51.

109. Chmarra MK, Grimbergen CA, Jansen FW, Dankelman J. How to objectively classify residents based on their psychomotor laparoscopic skills? *Minim Invasive Ther Allied Technol* 2010; **19**(1): 2-11.

110. Bann SD, Khan MS, Darzi AW. Measurement of surgical dexterity using motion analysis of simple bench tasks. *World J Surg* 2003; **27**(4): 390-4.

111. Smith SG, Torkington J, Brown TJ, Taffinder NJ, Darzi A. Motion analysis. *Surg Endosc* 2002; **16**(4): 640-5.

112. Taffinder N, Smith SG, Huber J, Russell RC, Darzi A. The effect of a second-generation 3D endoscope on the laparoscopic precision of novices and experienced surgeons. *Surg Endosc* 1999; **13**(11): 1087-92.

113. Kwasnicki RM, Aggarwal R, Lewis TM, Purkayastha S, Darzi A, Paraskeva PA. A comparison of skill acquisition and transfer in single incision and multi-port laparoscopic surgery. *J Surg Educ* 2013; **70**(2): 172-9.

114. Yamaguchi S, Yoshida D, Kenmotsu H, et al. Objective assessment of laparoscopic suturing skills using a motion-tracking system. *Surg Endosc* 2011; **25**(3): 771-5.

115. Hanna GB, Drew T, Clinch P, Hunter B, Cuschieri A. Computer-controlled endoscopic performance assessment system. *Surg Endosc* 1998; **12**(7): 997-1000.

116. Francis NK, Hanna GB, Cuschieri A. The performance of master surgeons on the Advanced Dundee Endoscopic Psychomotor Tester: contrast validity study. *Arch Surg* 2002; **137**(7): 841-4.

117. Francis NK, Hanna GB, Cuschieri A. Reliability of the Advanced Dundee Endoscopic Psychomotor Tester for bimanual tasks. *Arch Surg* 2001; **136**(1): 40-3.

118. Hwang H, Lim J, Kinnaird C, et al. Correlating motor performance with surgical error in laparoscopic cholecystectomy. *Surg Endosc* 2006; **20**(4): 651-5.

119. Egi H, Okajima M, Yoshimitsu M, et al. Objective assessment of endoscopic surgical skills by analyzing direction-dependent dexterity using the Hiroshima University Endoscopic Surgical Assessment Device (HUESAD). *Surg Today* 2008; **38**(8): 705-10.

120. Egi H, Okajima M, Kawahara T, et al. Scientific assessment of endoscopic surgical skills. *Minim Invasive Ther Allied Technol* 2010; **19**(1): 30-4.

121. Cundy TP, Thangaraj E, Rafii-Tari H, et al. Force-Sensing Enhanced Simulation Environment (ForSense) for laparoscopic surgery training and assessment. *Surgery* 2015; **157**(4): 723-31.

122. Horeman T, Dankelman J, Jansen FW, van den Dobbelsteen JJ. Assessment of laparoscopic skills based on force and motion parameters. *IEEE Trans Biomed Eng* 2014; **61**(3): 805-13.

123. Bouarfa L, Akman O, Schneider A, Jonker PP, Dankelman J. In-vivo real-time tracking of surgical instruments in endoscopic video. *Minim Invasive Ther Allied Technol* 2012; **21**(3): 129-34.

124. Reiley CE, Lin HC, Yuh DD, Hager GD. Review of methods for objective surgical skill evaluation. *Surg Endosc* 2011; **25**(2): 356-66.

125. Leong JJ, Nicolaou M, Atallah L, Mylonas GP, Darzi AW, Yang GZ. HMM assessment of quality of movement trajectory in laparoscopic surgery. *Med Image Comput Comput Assist Interv* 2006; **9**(Pt 1): 752-9.

126. Mackel T, Rosen J, Pugh C. Application of hidden markov modeling to objective medical skill evaluation. *Stud Health Technol Inform* 2007; **125**: 316-8.

127. Murphy TEV, C. M.; Yuh, D. D.; Okamura, A. Automatic motion recognition skill evaluation for dynamic tasks. *Eurohaptics* 2003: 363-73.

128. Lin HC, Shafran I, Murphy TE, Okamura AM, Yuh DD, Hager GD. Automatic detection and segmentation of robot-assisted surgical motions. *Med Image Comput Comput Assist Interv* 2005; **8**(Pt 1): 802-10.

129. Castellani AB, D.; Bicego, M.; Fiorini, P. Hybrid HMM/SVM model for the analysis and segmentation of teleoperation tasks. IEEE International Conference on Robotics and Automation (ICRA); 2004: IEEE; 2004. p. 2918-23.

130. Szalay D, MacRae H, Regehr G, Reznick R. Using operative outcome to assess technical skill. *Am J Surg* 2000; **180**(3): 234-7.

131. Datta V, Mandalia M, Mackay S, Chang A, Cheshire N, Darzi A. Relationship between skill and outcome in the laboratory-based model. *Surgery* 2002; **131**(3): 318-23.

132. Hanna GB, Frank TG, Cuschieri A. Objective assessment of endoscopic knot quality. *Am J Surg* 1997; **174**(4): 410-3.

133. McCloy R, Stone R. Science, medicine, and the future. Virtual reality in surgery. *BMJ* 2001; **323**(7318): 912-5.

134. Ericsson KA, Williams AM. Capturing naturally occurring superior performance in the laboratory: translational research on expert performance. *J Exp Psychol Appl* 2007; **13**(3): 115-23.

135. Janelle CH, B. Visual attention and brain processes that underlie expert performance: implications for sport and military psychology. *Mil Psychol* 2008; **20**: S39-S69.

136. Cline D. The angle velocity of eye movements. *Psychological Review* 1901; **8**: 145-57.

137. Fitts PMJREM, J. L. Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review* 1950; **9**(2): 24-9.

138. Duchowski AT. Eye Tracking Methodology. 2nd ed: Springer-Verlag London; 2007.

139. Hermen FF, R.; Ahmed, I. Eye movements in surgery: A literature review. *Journal of Eye Movement Research* 2013; **6**(4): 1-11.

140. Law BA, M. S.; Kirkpatrick, A. E.; Lomax, A. J. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. Symposium on eye tracking research and applications; 2004; 2004. p. 41-8.

141. Tien G, Zheng B, Atkins MS. Quantifying surgeons' vigilance during laparoscopic operations using eyegaze tracking. *Stud Health Technol Inform* 2011; **163**: 658-62.

142. Wilson M, McGrath J, Vine S, Brewer J, Defriend D, Masters R. Psychomotor control in a virtual laparoscopic surgery training environment: gaze control parameters differentiate novices from experts. *Surg Endosc* 2010; **24**(10): 2458-64.

143. Richstone L, Schwartz MJ, Seideman C, Cadeddu J, Marshall S, Kavoussi LR. Eye metrics as an objective assessment of surgical skill. *Ann Surg* 2010; **252**(1): 177-82.

144. James A, Vieira D, Lo B, Darzi A, Yang GZ. Eye-gaze driven surgical workflow segmentation. *Med Image Comput Comput Assist Interv* 2007; **10**(Pt 2): 110-7.

145. Brodmann K. Brodmann's: Localisation in the Cerebral Cortex. New York: Springr US; 2006.

146. Penfield WBE. Somatic motor and sensory representation in the cerebral cortex of man as studied by electric stimulation. *Brain* 1937; **60**(1): 389-443.

147. Leisman G, Melillo R. The basal ganglia: motor and cognitive relationships in a clinical neurobehavioral context. *Rev Neurosci* 2013; **24**(1): 9-25.

148. Leisman GM, R.; Carrick, F. R. Clinical Motor and Cognitive Neurobehavioral Relationships in the Basal Ganglia. In: Barrios FAB, C., ed. Basal Ganglia- An Integrative View: INTECH; 2013: 30.

149. Sommer MA. The role of the thalamus in motor control. *Curr Opin Neurobiol* 2003; **13**(6): 663-70.

150. Middleton FA, Strick PL. Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Res Brain Res Rev* 2000; **31**(2-3): 236-50.

151. Mink JW. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol* 1996; **50**(4): 381-425.

152. Shin C. Neurophysiologic basis of functional neuroimaging: animal studies. *J Clin Neurophysiol* 2000; **17**(1): 2-9.

153. Jackson AF, Bolger DJ. The neurophysiological bases of EEG and EEG measurement: a review for the rest of us. *Psychophysiology* 2014; **51**(11): 1061-71.

154. Collins RC. Basic aspects of functional brain metabolism. *Ciba Found Symp* 1991; **163**: 6-16; discussion -22.

155. Sokoloff L. The physiological and biochemical bases of functional brain imaging. *Cogn Neurodyn* 2008; **2**(1): 1-5.

156. Attwell D, Iadecola C. The neural basis of functional brain imaging signals. *Trends Neurosci* 2002; **25**(12): 621-5.

157. Roy CS, Sherrington CS. On the Regulation of the Blood-supply of the Brain. *J Physiol* 1890; **11**(1-2): 85-158 17.

158. Dirnagl U. Metabolic aspects of neurovascular coupling. *Adv Exp Med Biol* 1997; **413**: 155-9.

159. Kuschinsky W. Neuronal-vascular coupling. A unifying hypothesis. *Adv Exp Med Biol* 1997; **413**: 167-76.

160. Shibasaki H. Human brain mapping: hemodynamic response and electrophysiology. *Clin Neurophysiol* 2008; **119**(4): 731-43.

161. Devor A, Ulbert I, Dunn AK, et al. Coupling of the cortical hemodynamic response to cortical and thalamic neuronal activity. *Proc Natl Acad Sci U S A* 2005; **102**(10): 3822-7.

162. Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 2001; **412**(6843): 150-7.

163. Arthurs OJ, Boniface SJ. What aspect of the fMRI BOLD signal best reflects the underlying electrophysiology in human somatosensory cortex? *Clin Neurophysiol* 2003; **114**(7): 1203-9.

164. Arthurs OJ, Donovan T, Spiegelhalter DJ, Pickard JD, Boniface SJ. Intracortically distributed neurovascular coupling relationships within and between human somatosensory cortices. *Cereb Cortex* 2007; **17**(3): 661-8.

165. Caesar K, Akgoren N, Mathiesen C, Lauritzen M. Modification of activity-dependent increases in cerebellar blood flow by extracellular potassium in anaesthetized rats. *J Physiol* 1999; **520 Pt 1**: 281-92.

166. Haas LF. Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography. *J Neurol Neurosurg Psychiatry* 2003; **74**(1): 9.

167. Nunez PLS, R. Electric fields of the brain: The neurophysics of EEG. 2nd ed. New York: Oxford University Press; 2006.

168. Oostenveld R, Praamstra P. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin Neurophysiol* 2001; **112**(4): 713-9.

169. Lopez-Gordo MA, Sanchez-Morillo D, Pelayo Valle F. Dry EEG electrodes. *Sensors (Basel)* 2014; **14**(7): 12847-70.

170. Crosson B, Ford A, McGregor KM, et al. Functional imaging and related techniques: an introduction for rehabilitation researchers. *J Rehabil Res Dev* 2010; **47**(2): vii-xxxiv.

171. Raichle ME. A brief history of human brain mapping. *Trends Neurosci* 2009; **32**(2): 118-26.

172. Otte A, Halsband U. Brain imaging tools in neurosciences. *J Physiol Paris* 2006; **99**(4-6): 281-92.

173. Ogawa S, Lee TM, Kay AR, Tank DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A* 1990; **87**(24): 9868-72.

174. Lauritzen M. Reading vascular changes in brain imaging: is dendritic calcium the key? *Nat Rev Neurosci* 2005; **6**(1): 77-85.

175. Ferrari M, Quaresima V. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage* 2012; **63**(2): 921-35.

176. Strangman G, Boas DA, Sutton JP. Non-invasive neuroimaging using near-infrared light. *Biol Psychiatry* 2002; **52**(7): 679-93.

177. Chance B, Luo Q, Nioka S, Alsop DC, Detre JA. Optical investigations of physiology: a study of intrinsic and extrinsic biomedical contrast. *Philos Trans R Soc Lond B Biol Sci* 1997; **352**(1354): 707-16.

178. Jobsis FF. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* 1977; **198**(4323): 1264-7.

179. Delpy DTC, M. Quantification in tissue near-infrared spectroscopy. *Philos Trans R Soc Lond B Biol Sci* 1997; **357**(1354): 649-59.

180. Fox PT, Raichle ME. Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proc Natl Acad Sci U S A* 1986; **83**(4): 1140-4.

181. Hoshi Y, Kobayashi N, Tamura M. Interpretation of near-infrared spectroscopy signals: a study with a newly developed perfused rat brain model. *J Appl Physiol (1985)* 2001; **90**(5): 1657-62.

182. Hoshi Y. Functional near-infrared spectroscopy: current status and future prospects. *J Biomed Opt* 2007; **12**(6): 062106.

183. Lloyd-Fox S, Blasi A, Elwell CE. Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci Biobehav Rev* 2010; **34**(3): 269-84.

184. Huppert TJ, Hoge RD, Diamond SG, Franceschini MA, Boas DA. A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *Neuroimage* 2006; **29**(2): 368-82.

185. Delpy DT, Cope M, van der Zee P, Arridge S, Wray S, Wyatt J. Estimation of optical pathlength through tissue from direct time of flight measurement. *Phys Med Biol* 1988; **33**(12): 1433-42.

186. Scholkmann F, Kleiser S, Metz AJ, et al. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage* 2014; **85 Pt 1**: 6-27.

187. Kocsis L, Herman P, Eke A. The modified Beer-Lambert law revisited. *Phys Med Biol* 2006; **51**(5): N91-8.

188. Cope M, Delpy DT. System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Med Biol Eng Comput* 1988; **26**(3): 289-94.

189. Watanabe EY, Y.; Maki, A.; Ito, Y.; Koizumi, H.; Mayanagi, Y. Infrared spectroscopy as noninvasive functional mapping of human brain. *Neurosci Res Suppl* 1994; **19**: 2501.

190. Minagawa-Kawai Y, Mori K, Hebden JC, Dupoux E. Optical imaging of infants' neurocognitive development: recent advances and perspectives. *Dev Neurobiol* 2008; **68**(6): 712-28.

191. Hoshi Y, Chen SJ. Regional cerebral blood flow changes associated with emotions in children. *Pediatr Neurol* 2002; **27**(4): 275-81.

192. Colier WQ, V.; Barattelli, G.; Cavallari, P.; van der Sluijs, M.; Ferrari, M. Detailed evidence of cerebral hemoglobin oxygenation in response to motor cortical activation revealed by a continuous wave spectrometer with 10 Hz temporal resolution. *Proc SPIE* 1997; **2979**: 390-6.

193. Okamoto M, Dan H, Shimizu K, et al. Multimodal assessment of cortical activation during apple peeling by NIRS and fMRI. *Neuroimage* 2004; **21**(4): 1275-88.

194. Leff DR, Orihuela-Espina F, Elwell CE, et al. Assessment of the cerebral cortex during motor task behaviours in adults: a systematic review of functional near infrared spectroscopy (fNIRS) studies. *Neuroimage* 2011; **54**(4): 2922-36.

195. Lloyd-Fox S, Richards JE, Blasi A, Murphy DG, Elwell CE, Johnson MH. Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. *Neurophotonics* 2014; **1**(2): 025006.

196. Walsh V, Cowey A. Transcranial magnetic stimulation and cognitive neuroscience. *Nat Rev Neurosci* 2000; **1**(1): 73-9.

197. Bakker AS, B.; Ainslie, P.; Smith, K. Near-Infrared Spectroscopy. In: Ainslie P, ed. Applied Aspects of Ultrasonography in Humans: InTech; 2012.

198. Choi J, Wolf M, Toronov V, et al. Noninvasive determination of the optical properties of adult brain: near-infrared spectroscopy approach. *J Biomed Opt* 2004; **9**(1): 221-9.

199. Orihuela-Espina F, Leff DR, James DR, Darzi AW, Yang GZ. Quality control and assurance in functional near infrared spectroscopy (fNIRS) experimentation. *Phys Med Biol* 2010; **55**(13): 3701-24.

200. Jurcak V, Tsuzuki D, Dan I. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. *Neuroimage* 2007; **34**(4): 1600-11.

201. Okamoto M, Dan H, Sakamoto K, et al. Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10-20 system oriented for transcranial functional brain mapping. *Neuroimage* 2004; **21**(1): 99-111.

202. Singh AK, Okamoto M, Dan H, Jurcak V, Dan I. Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. *Neuroimage* 2005; **27**(4): 842-51.

203. Tsuzuki D, Dan I. Spatial registration for functional near-infrared spectroscopy: from channel position on the scalp to cortical location in individual and group analyses. *Neuroimage* 2014; **85 Pt 1**: 92-103.

204. Goltz F. In: MacCormac W, editor. Transactions of the 7th International Medical Congress. London: JW Kolkmann; 1881. p. 218-28.

205. Catani M, Thiebaut de Schotten M. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex* 2008; **44**(8): 1105-32.

206. Bozzali M, Parker GJ, Serra L, et al. Anatomical connectivity mapping: a new tool to assess brain disconnection in Alzheimer's disease. *Neuroimage* 2011; **54**(3): 2045-51.

207. Fitzsimmons J, Kubicki M, Shenton ME. Review of functional and anatomical brain connectivity findings in schizophrenia. *Curr Opin Psychiatry* 2013; **26**(2): 172-87.

208. Jarbo K, Verstynen TD. Converging structural and functional connectivity of orbitofrontal, dorsolateral prefrontal, and posterior parietal cortex in the human striatum. *J Neurosci* 2015; **35**(9): 3865-78.

209. Friston KJ, Frith CD, Liddle PF, Frackowiak RS. Functional connectivity: the principal-component analysis of large (PET) data sets. *J Cereb Blood Flow Metab* 1993; **13**(1): 5-14.

210. Gerstein GL, Perkel DH. Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 1969; **164**(3881): 828-30.

211. Homae F, Watanabe H, Otobe T, et al. Development of global cortical networks in early infancy. *J Neurosci* 2010; **30**(14): 4877-82.

212. Mesquita RC, Franceschini MA, Boas DA. Resting state functional connectivity of the whole head with near-infrared spectroscopy. *Biomed Opt Express* 2010; **1**(1): 324-36.

213. Sasai S, Homae F, Watanabe H, et al. A NIRS-fMRI study of resting state network. *Neuroimage* 2012; **63**(1): 179-93.

214. White BR, Snyder AZ, Cohen AL, et al. Resting-state functional connectivity in the human brain revealed with diffuse optical tomography. *Neuroimage* 2009; **47**(1): 148-56.

215. Xu J, Liu X, Zhang J, et al. FC-NIRS: A Functional Connectivity Analysis Tool for Near-Infrared Spectroscopy Data. *Biomed Res Int* 2015; **2015**: 248724.

216. Zhang H, Zhang YJ, Lu CM, Ma SY, Zang YF, Zhu CZ. Functional connectivity as revealed by independent component analysis of resting-state fNIRS measurements. *Neuroimage* 2010; **51**(3): 1150-61.

217. Zhu H, Fan Y, Guo H, Huang D, He S. Reduced interhemispheric functional connectivity of children with autism spectrum disorder: evidence from functional near infrared spectroscopy studies. *Biomed Opt Express* 2014; **5**(4): 1262-74.

218. Tak S, Ye JC. Statistical analysis of fNIRS data: a comprehensive review. *Neuroimage* 2014; **85 Pt 1**: 72-91.

219. Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 2007; **8**(9): 700-11.

220. Ma L, Wang B, Narayana S, et al. Changes in regional activity are accompanied with changes in inter-regional connectivity during 4 weeks motor learning. *Brain Res* 2010; **1318**: 64-76.

221. Sun FT, Miller LM, Rao AA, D'Esposito M. Functional connectivity of cortical networks involved in bimanual motor sequence learning. *Cereb Cortex* 2007; **17**(5): 1227-34.

222. Bajaj S, Drake D, Butler AJ, Dhamala M. Oscillatory motor network activity during rest and movement: an fNIRS study. *Front Syst Neurosci* 2014; **8**: 13.

223. Friston KJF, C. D.; Frackowiak, R. S. J. Time-dependent changes in effective connectivity measured with PET. *Human Brain Mapping* 1993; **1**: 69-80.

224. Friston KJF, C.; Frackowiak, R. S. J. Time-dependent changes in effective connectivity measured with PET. *Human Brain Mapping* 1993; **1**: 69-79.

225. Friston K, Moran R, Seth AK. Analysing connectivity with Granger causality and dynamic causal modelling. *Curr Opin Neurobiol* 2013; **23**(2): 172-8.

226. Tak S, Kempny AM, Friston KJ, Leff AP, Penny WD. Dynamic causal modelling for functional near-infrared spectroscopy. *Neuroimage* 2015; **111**: 338-49.

227. Hatakenaka M, Miyai I, Mihara M, Sakoda S, Kubota K. Frontal regions involved in learning of motor skill--A functional NIRS study. *Neuroimage* 2007; **34**(1): 109-16.

228. Shetty K, Leff DR, Orihuela-Espina F, Darzi A, Yang GZ. A study of executive control during intracorporeal minimally invasive suturing (ICS). fNIRS Meeting. London, United Kingdom; 2012.

229. Sachdeva AK, Bell RH, Jr., Britt LD, Tarpley JL, Blair PG, Tarpley MJ. National efforts to reform residency education in surgery. *Acad Med* 2007; **82**(12): 1200-10.

230. Darzi A. Quality and the NHS next stage review. *Lancet* 2008; **371**(9624): 1563-4.

231. Sharma B, Mishra A, Aggarwal R, Grantcharov TP. Non-technical skills assessment in surgery. *Surg Oncol* 2011; **20**(3): 169-77.

232. Tachtsidis I, Leung TS, Devoto L, Delpy DT, Elwell CE. Measurement of frontal lobe functional activation and related systemic effects: a near-infrared spectroscopy investigation. *Adv Exp Med Biol* 2008; **614**: 397-403.

233. Datta V, Mackay S, Mandalia M, Darzi A. The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *J Am Coll Surg* 2001; **193**(5): 479-85.

234. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Learning curves and impact of previous operative experience on performance on a virtual reality simulator to test laparoscopic surgical skills. *Am J Surg* 2003; **185**(2): 146-9.

235. Leff DR, Leong JJ, Aggarwal R, Yang GZ, Darzi A. Could variations in technical skills acquisition in surgery be explained by differences in cortical plasticity? *Ann Surg* 2008; **247**(3): 540-3.

236. Richards D, Stedmon A. To delegate or not to delegate: A review of control frameworks for autonomous cars. *Appl Ergon* 2016; **53 Pt B**: 383-8.

237. Salas EM, D. Human Factors in Aviation. 2nd ed: Elsevier; 2010.

238. Walker GHS, N. A. Human Factors in Automotive Engineering and Technology: CRC Press; 2015.

239. Salvendy G. Handbook of Human Factors and Ergonomics. 4th ed; 2012.

240. Janelle CH, B. Visual Attention and Brain Processes that Underlie Expert Performance: Implications for Sport and Military Psychology. *Mil Psych* 2008; **20**: S39-S69.

241. Zheng B, Cassera MA, Martinec DV, Spaun GO, Swanstrom LL. Measuring mental workload during the performance of advanced laparoscopic tasks. *Surg Endosc* 2010; **24**(1): 45-50.

242. Tien T, Pucher PH, Sodergren MH, Sriskandarajah K, Yang GZ, Darzi A. Eye tracking for skills assessment and training: a systematic review. *J Surg Res* 2014; **191**(1): 169-78.

243. Parasuraman R, Christensen J, Grafton S. Neuroergonomics: the brain in action and at work. *Neuroimage* 2012; **59**(1): 1-3.

244. Parasuraman R, Wilson GF. Putting the brain to work: neuroergonomics past, present, and future. *Hum Factors* 2008; **50**(3): 468-74.

245. Doyon J, Benali H. Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol* 2005; **15**(2): 161-7.

246. Halsband U, Lange RK. Motor learning in man: a review of functional and clinical studies. *J Physiol Paris* 2006; **99**(4-6): 414-24.

247. Huneau C, Benali H, Chabriat H. Investigating Human Neurovascular Coupling Using Functional Neuroimaging: A Critical Review of Dynamic Models. *Front Neurosci* 2015; **9**: 467.

248. Strangman G, Culver JP, Thompson JH, Boas DA. A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation. *Neuroimage* 2002; **17**(2): 719-31.

249. Sporns O, Tononi G, Edelman GM. Connectivity and complexity: the relationship between neuroanatomy and brain dynamics. *Neural Netw* 2000; **13**(8-9): 909-22.

250. Tononi G, Edelman GM, Sporns O. Complexity and coherency: integrating information in the brain. *Trends Cogn Sci* 1998; **2**(12): 474-84.

251. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151**(4): 264-9, W64.

252. Oxford Centre for Evidence-based Medicine – Levels of Evidence. 2009. http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/ (accessed 01/07/2016.

253. James DR, Leff DR, Orihuela-Espina F, et al. Enhanced frontoparietal network architectures following "gaze-contingent" versus "free-hand" motor learning. *Neuroimage* 2013; **64**: 267-76.

254. Leff DR, James DR, Orihuela-Espina F, et al. The impact of expert visual guidance on trainee visual search strategy, visual attention and motor skills. *Front Hum Neurosci* 2015; **9**: 526.

255. Ros T, Moseley MJ, Bloom PA, Benjamin L, Parkinson LA, Gruzelier JH. Optimizing microsurgical skills with EEG neurofeedback. *BMC Neurosci* 2009; **10**: 87.

256. Wanzel KR, Anastakis DJ, McAndrews MP, et al. Visual-spatial ability and fMRI cortical activation in surgery residents. *Am J Surg* 2007; **193**(4): 507-10.

257. Crewther BT, Shetty K, Jarchi D, et al. Skill acquisition and stress adaptations following laparoscopic surgery training and detraining in novice surgeons. *Surg Endosc* 2015.

258. James DR, Orihuela-Espina F, Leff DR, et al. The ergonomics of natural orifice translumenal endoscopic surgery (NOTES) navigation in terms of performance, stress, and cognitive behavior. *Surgery* 2011; **149**(4): 525-33.

259. Leff DR, Elwell CE, Orihuela-Espina F, et al. Changes in prefrontal cortical behaviour depend upon familiarity on a bimanual co-ordination task: an fNIRS study. *Neuroimage* 2008; **39**(2): 805-13.

260. Leff DR, Orihuela-Espina F, Atallah L, et al. Functional prefrontal reorganization accompanies learning-associated refinements in surgery: a manifold embedding approach. *Comput Aided Surg* 2008; **13**(6): 325-39.

261. Leff DR, Orihuela-Espina F, Athanasiou T, et al. "Circadian cortical compensation": a longitudinal study of brain function during technical and cognitive skills in acutely sleep-deprived surgical residents. *Ann Surg* 2010; **252**(6): 1082-90.

262. Leff DR, Orihuela-Espina F, Leong J, Darzi A, Yang GZ. Modelling dynamic fronto-parietal behaviour during minimally invasive surgery--a Markovian trip distribution approach. *Med Image Comput Comput Assist Interv* 2008; **11**(Pt 2): 595-602.

263. Leff DR, Yongue G, Vlaev I, et al. "Contemplating the Next Maneuver": Functional Neuroimaging Reveals Intraoperative Decision-making Strategies. *Ann Surg* 2016.

264. Bahrami P, Graham SJ, Grantcharov TP, et al. Neuroanatomical correlates of laparoscopic surgery training. *Surg Endosc* 2014; **28**(7): 2189-98.

265. Bahrami P, Schweizer TA, Tam F, Grantcharov TP, Cusimano MD, Graham SJ. Functional MRI-compatible laparoscopic surgery training simulator. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 2011; **65**(3): 873-81.

266. Duty B, Andonian S, Ma Y, et al. Correlation of laparoscopic experience with differential functional brain activation: a positron emission tomography study with oxygen 15-labeled water. *Arch Surg* 2012; **147**(7): 627-32.

267. Guru KA, Esfahani ET, Raza SJ, et al. Cognitive skills assessment during robot-assisted surgery: separating the wheat from the chaff. *BJU Int* 2015; **115**(1): 166-74.

268. Hussein AA, Shafiei SB, Sharif M, et al. Technical Mentorship during Robot-Assisted Surgery: A Cognitive Analysis? *BJU Int* 2016.

269. Miura S, Kobayashi Y, Kawamura K, Nakashima Y, Fujie MG. Brain activation in parietal area during manipulation with a surgical robot simulator. *Int J Comput Assist Radiol Surg* 2015; **10**(6): 783-90.

270. Bocci T, Moretto C, Tognazzi S, et al. How does a surgeon's brain buzz? An EEG coherence study on the interaction between humans and robot. *Behav Brain Funct* 2013; **9**: 14.

271. Morris MC, Frodl T, D'Souza A, Fagan AJ, Ridgway PF. Assessment of competence in surgical skills using functional magnetic resonance imaging: a feasibility study. *J Surg Educ* 2015; **72**(2): 198-204.

272. Zhu FF, Poolton JM, Wilson MR, Hu Y, Maxwell JP, Masters RS. Implicit motor learning promotes neural efficiency during laparoscopy. *Surg Endosc* 2011; **25**(9): 2950-5.

273. Shibusawa M, Takeda T, Nakajima K, Ishigami K, Sakatani K. Functional near-infrared spectroscopy study on primary motor and sensory cortex response to clenching. *Neurosci Lett* 2009; **449**(2): 98-102.

274. Darzi A, Smith S, Taffinder N. Assessing operative skill. Needs to become more objective. *BMJ* 1999; **318**(7188): 887-8.

275. Tenenbaum JB, de Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science* 2000; **290**(5500): 2319-23.

276. Rubner YT, C.; Guibas, L.J. A metric for distributions with applications to image databases. Proceedings of the 6th International Conference on Computer Vision; 1998; Bombay, India; 1998.

277. Kelly AM, Garavan H. Human functional neuroimaging of brain changes associated with practice. *Cereb Cortex* 2005; **15**(8): 1089-102.

278. Rowe J, Friston K, Frackowiak R, Passingham R. Attention to action: specific modulation of corticocortical interactions in humans. *Neuroimage* 2002; **17**(2): 988-98.

279. Petersen SE, van Mier H, Fiez JA, Raichle ME. The effects of practice on the functional anatomy of task performance. *Proc Natl Acad Sci U S A* 1998; **95**(3): 853-60.

280. Ikkai A, Curtis CE. Common neural mechanisms supporting spatial working memory, attention and motor intention. *Neuropsychologia* 2011; **49**(6): 1428-34.

281. Swinnen SP, Wenderoth N. Two hands, one brain: cognitive neuroscience of bimanual skill. *Trends Cogn Sci* 2004; **8**(1): 18-25.

282. Gupta R, Cathelineau X, Rozet F, Vallancien G. Feedback from operative performance to improve training program of laparoscopic radical prostatectomy. *J Endourol* 2004; **18**(9): 836-9.

283. Nieboer TE, Sari V, Kluivers KB, Weinans MJ, Vierhout ME, Stegeman DF. A randomized trial of training the non-dominant upper extremity to enhance laparoscopic performance. *Minim Invasive Ther Allied Technol* 2012; **21**(4): 259-64.

284. Chang L, Satava RM, Pellegrini CA, Sinanan MN. Robotic surgery: identifying the learning curve through objective measurement of skill. *Surg Endosc* 2003; **17**(11): 1744-8.

285. Moorthy K, Munz Y, Dosis A, et al. Dexterity enhancement with robotic surgery. *Surg Endosc* 2004; **18**(5): 790-5.

286. Sarle R, Tewari A, Shrivastava A, Peabody J, Menon M. Surgical robotics and laparoscopic training drills. *J Endourol* 2004; **18**(1): 63-6; discussion 6-7.

287. Stefanidis D, Wang F, Korndorffer JR, Jr., Dunne JB, Scott DJ. Robotic assistance improves intracorporeal suturing performance and safety in the operating room while decreasing operator workload. *Surg Endosc* 2010; **24**(2): 377-82.

288. Duque J, Davare M, Delaunay L, et al. Monitoring coordination during bimanual movements: where is the mastermind? *J Cogn Neurosci* 2010; **22**(3): 526-42.

289. Liu J, Morel A, Wannier T, Rouiller EM. Origins of callosal projections to the supplementary motor area (SMA): a direct comparison between pre-SMA and SMA-proper in macaque monkeys. *J Comp Neurol* 2002; **443**(1): 71-85.

290. Andres FG, Mima T, Schulman AE, Dichgans J, Hallett M, Gerloff C. Functional coupling of human cortical sensorimotor areas during bimanual skill acquisition. *Brain* 1999; **122 (Pt 5)**: 855-70.

291. Wilson TW, Kurz MJ, Arpin DJ. Functional specialization within the supplementary motor area: a fNIRS study of bimanual coordination. *Neuroimage* 2014; **85 Pt 1**: 445-50.

292. Maravita A, Iriki A. Tools for the body (schema). *Trends Cogn Sci* 2004; **8**(2): 79-86.

293. Leff DR, Aggarwal R, Rana M, et al. Laparoscopic skills suffer on the first shift of sequential night shifts: program directors beware and residents prepare. *Ann Surg* 2008; **247**(3): 530-9.

294. Lockley SW, Cronin JW, Evans EE, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. *N Engl J Med* 2004; **351**(18): 1829-37.

295. Wesnes KA, Walker MB, Walker LG, et al. Cognitive performance and mood after a weekend on call in a surgical unit. *Br J Surg* 1997; **84**(4): 493-5.

296. Shepard RN, Metzler J. Mental rotation of three-dimensional objects. *Science* 1971; **171**(3972): 701-3.

297. Wanzel KR, Hamstra SJ, Anastakis DJ, Matsumoto ED, Cusimano MD. Effect of visual-spatial ability on learning of spatially-complex surgical skills. *Lancet* 2002; **359**(9302): 230-1.

298. Leff DR, Leong J, Yang GZ, Darzi AW. Visuo-spatial ability and fMRI cortical activation in surgery residents. *Am J Surg* 2008; **195**(1): 138.

299. Coynel D, Marrelec G, Perlbarg V, et al. Dynamics of motor-related functional integration during motor sequence learning. *Neuroimage* 2010; **49**(1): 759-66.

300. Shadmehr R, Holcomb HH. Neural correlates of motor memory consolidation. *Science* 1997; **277**(5327): 821-5.

301. Chetwood AS, Kwok KW, Sun LW, et al. Collaborative eye tracking: a potential training tool in laparoscopic surgery. *Surg Endosc* 2012; **26**(7): 2003-9.

302. Theeuwes J. Top-down and bottom-up control of visual selection. *Acta Psychol (Amst)* 2010; **135**(2): 77-99.

303. Kojima H, Suzuki T. Hemodynamic change in occipital lobe during visual search: visual attention allocation measured with NIRS. *Neuropsychologia* 2010; **48**(1): 349-52.

304. Dixon ML, Christoff K. The lateral prefrontal cortex and complex value-based learning and decision making. *Neurosci Biobehav Rev* 2014; **45**: 9-18.

305. Friston KJ. Functional Integration. In: Frackowiak RSJ, ed. Human Brain Function. 2nd Edition ed: Elsevier; 2004: 971-8.

306. Hardwick RM, Rottschy C, Miall RC, Eickhoff SB. A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage* 2013; **67**: 283-97.

307. Hikosaka O, Nakamura K, Sakai K, Nakahara H. Central mechanisms of motor skill learning. *Curr Opin Neurobiol* 2002; **12**(2): 217-22.

308. Abernethy B, Poolton JM, Masters RS, Patil NG. Implications of an expertise model for surgical skills training. *ANZ J Surg* 2008; **78**(12): 1092-5.

309. Tun JK, Kneebone R. Bridging worlds: applying the science of motor learning to clinical education. *Med Educ* 2011; **45**(2): 111-4.

310. Fitts PMP, M. I. Human performance. Belmont, Calif.,: Brooks/Cole Pub. Co.; 1967.

311. Debaere F, Wenderoth N, Sunaert S, Van Hecke P, Swinnen SP. Changes in brain activation during the acquisition of a new bimanual coodination task. *Neuropsychologia* 2004; **42**(7): 855-67.

312. Puttemans V, Wenderoth N, Swinnen SP. Changes in brain activation during the acquisition of a multifrequency bimanual coordination task: from the cognitive stage to advanced levels of automaticity. *J Neurosci* 2005; **25**(17): 4270-8.

313. Tracy J, Flanders A, Madi S, et al. Regional brain activation associated with different performance patterns during learning of a complex motor skill. *Cereb Cortex* 2003; **13**(9): 904-10.

314. Ayaz H, Shewokis PA, Bunce S, Izzetoglu K, Willems B, Onaral B. Optical brain monitoring for operator training and mental workload assessment. *Neuroimage* 2012; **59**(1): 36-47.

315. Harada H, Nashihara H, Morozumi K, Ota H, Hatakeyama E. A comparison of cerebral activity in the prefrontal region between young adults and the elderly while driving. *J Physiol Anthropol* 2007; **26**(3): 409-14.

316. Picton TW, Stuss DT, Alexander MP, Shallice T, Binns MA, Gillingham S. Effects of focal frontal lesions on response inhibition. *Cereb Cortex* 2007; **17**(4): 826-38.

317. Karni A, Meyer G, Jezzard P, Adams MM, Turner R, Ungerleider LG. Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature* 1995; **377**(6545): 155-8.

318. Karni A, Meyer G, Rey-Hipolito C, et al. The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci U S A* 1998; **95**(3): 861-8.

319. Kroeze SG, Mayer EK, Chopra S, Aggarwal R, Darzi A, Patel A. Assessment of laparoscopic suturing skills of urology residents: a pan-European study. *Eur Urol* 2009; **56**(5): 865-72.

320. Moorthy K, Munz Y, Dosis A, Bello F, Chang A, Darzi A. Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity. *Surg Endosc* 2004; **18**(11): 1608-12.

321. Bell EC, Willson MC, Wilman AH, Dave S, Silverstone PH. Males and females differ in brain activation during cognitive tasks. *Neuroimage* 2006; **30**(2): 529-38.

322. Tchantchaleishvili V, Myers PO. Left-handedness--a handicap for training in surgery? *J Surg Educ* 2010; **67**(4): 233-6.

323. The American Board of Surgery to require ACLS, ATLS and FLS for General Surgery Certification. 2008. http://www.absurgery.org/default.jsp?news_newreqs (accessed 01/07/2016.

324. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. *Ann Surg* 2004; **240**(3): 518-25; discussion 25-8.

325. Floyer-Lea A, Matthews PM. Distinguishable brain activation networks for short- and long-term motor skill learning. *Journal of neurophysiology* 2005; **94**(1): 512-8.

326. Leff DRL, J.; Atallah, L.; Athanasiou, T.; Koh, P. H.; Elwell, C.; Delpy, D.; Yang, G. Z.; Darzi, A. Open Surgical Knot-Tying Induces a Lateralised Prefrontal Brain Response in Surgical Novices. Organization for Human Brain Mapping. Chicago, Illinois, USA; 2007.

327. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998; **393**(6684): 440-2.

328. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 2009; **10**(3): 186-98.

329. Langer N, von Bastian CC, Wirz H, Oberauer K, Jancke L. The effects of working memory training on functional brain network efficiency. *Cortex* 2013; **49**(9): 2424-38.

330. Swinnen SP. Intermanual coordination: from behavioural principles to neural-network interactions. *Nat Rev Neurosci* 2002; **3**(5): 348-59.

331. Pagador JB, Sanchez-Margallo FM, Sanchez-Peralta LF, et al. Decomposition and analysis of laparoscopic suturing task using tool-motion analysis (TMA): improving the objective assessment. *Int J Comput Assist Radiol Surg* 2012; **7**(2): 305-13.

332. Hart SG. NASA-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting; 2006 October; 2006. p. 904-8.

333. Tsuzuki D, Dan I. Spatial registration for functional near-infrared spectroscopy: From channel position on the scalp to cortical location in individual and group analyses. *Neuroimage* 2013.

334. Orihuela-Espina F, Leff DR, James DRC, Darzi A, Yang GZ. A software tool for manifold embedded based analysis of functional near infrared spectroscopy data. Organization for Human Brain Mapping. San Francisco, California, USA; 2009.

335. Yang HK, Lee JW, Lee KH, et al. Application for the wearable heart activity monitoring system: analysis of the autonomic function of HRV. *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference* 2008; **2008**: 1258-61.

336. Bolanos M, Bernat EM, He B, Aviyente S. A weighted small world network measure for assessing functional connectivity. *J Neurosci Methods* 2013; **212**(1): 133-42.

337. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 2010; **52**(3): 1059-69.

338. Bassett DS, Bullmore E. Small-world brain networks. *Neuroscientist* 2006; **12**(6): 512-23.

339. Dulan G, Rege RV, Hogg DC, et al. Proficiency-based training for robotic surgery: construct validity, workload, and expert levels for nine inanimate exercises. *Surg Endosc* 2012; **26**(6): 1516-21.

340. Izzetoglu K, Bunce S, Onaral B, Pourrezaei K, Chance B. Functional optical brain imaging using near-infrared during cognitive tasks. *Int J Hum-Comput Int* 2004; **17**(2): 211-27.

341. Iacoboni M, Woods RP, Mazziotta JC. Brain-behavior relationships: evidence from practice effects in spatial stimulus-response compatibility. *Journal of neurophysiology* 1996; **76**(1): 321-31.

342. Lefebvre S, Dricot L, Gradkowski W, Laloux P, Vandermeeren Y. Brain activations underlying different patterns of performance improvement during early motor skill learning. *Neuroimage* 2012; **62**(1): 290-9.

343. Swinnen SP, Jardin K, Meulenbroek R, Dounskaia N, Den Brandt MH. Egocentric and allocentric constraints in the expression of patterns of interlimb coordination. *J Cogn Neurosci* 1997; **9**(3): 348-77.

344. Swinnen SP, Jardin K, Verschueren S, et al. Exploring interlimb constraints during bimanual graphic performance: effects of muscle grouping and direction. *Behavioural brain research* 1998; **90**(1): 79-87.

345. Debaere F, Swinnen SP, Beatse E, Sunaert S, Van Hecke P, Duysens J. Brain areas involved in interlimb coordination: a distributed network. *Neuroimage* 2001; **14**(5): 947-58.

346. Kermadi I, Liu Y, Rouiller EM. Do bimanual motor actions involve the dorsal premotor (PMd), cingulate (CMA) and posterior parietal (PPC) cortices? Comparison with primary and supplementary motor cortical areas. *Somatosens Mot Res* 2000; **17**(3): 255-71.

347. Stephan KM, Binkofski F, Halsband U, et al. The role of ventral medial wall motor areas in bimanual co-ordination. A combined lesion and activation study. *Brain* 1999; **122 (Pt 2)**: 351-68.

348. Stephan KM, Binkofski F, Posse S, Seitz RJ, Freund HJ. Cerebral midline structures in bimanual coordination. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale* 1999; **128**(1-2): 243-9.

349. Garbarini F, D'Agata F, Piedimonte A, et al. Drawing lines while imagining circles: Neural basis of the bimanual coupling effect during motor execution and motor imagery. *Neuroimage* 2013; **88C**: 100-12.

350. Immisch I, Waldvogel D, van Gelderen P, Hallett M. The role of the medial wall and its anatomical variations for bimanual antiphase and in-phase movements. *Neuroimage* 2001; **14**(3): 674-84.

351. Jancke L, Shah NJ, Peters M. Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. *Brain Res Cogn Brain Res* 2000; **10**(1-2): 177-83.

352. Holper L, Biallas M, Wolf M. Task complexity relates to activation of cortical motor areas during uni- and bimanual performance: a functional NIRS study. *Neuroimage* 2009; **46**(4): 1105-13.

353. Voss MW, Prakash RS, Erickson KI, et al. Effects of training strategies implemented in a complex videogame on functional connectivity of attentional networks. *Neuroimage* 2012; **59**(1): 138-48.

354. Heitger MH, Ronsse R, Dhollander T, Dupont P, Caeyenberghs K, Swinnen SP. Motor learning-induced changes in functional brain connectivity as revealed by means of graph-theoretical network analysis. *Neuroimage* 2012; **61**(3): 633-50.

355. Taubert M, Lohmann G, Margulies DS, Villringer A, Ragert P. Long-term effects of motor training on resting-state networks and underlying brain structure. *Neuroimage* 2011; **57**(4): 1492-8.

356. Abe M, Schambra H, Wassermann EM, Luckenbaugh D, Schweighofer N, Cohen LG. Reward improves long-term retention of a motor memory through induction of offline memory gains. *Current biology : CB* 2011; **21**(7): 557-62.

357. Lehericy S, Benali H, Van de Moortele PF, et al. Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *Proc Natl Acad Sci U S A* 2005; **102**(35): 12566-71.

358. Debas K, Carrier J, Orban P, et al. Brain plasticity related to the consolidation of motor sequence learning and motor adaptation. *Proc Natl Acad Sci U S A* 2010; **107**(41): 17839-44.

359. Krakauer JW, Mazzoni P. Human sensorimotor learning: adaptation, skill, and beyond. *Curr Opin Neurobiol* 2011; **21**(4): 636-44.

360. Shmuelof L, Krakauer JW. Are we ready for a natural history of motor learning? *Neuron* 2011; **72**(3): 469-76.

361. Leff DR, Orihuela-Espina F, Atallah L, Darzi A, Yang GZ. Functional near infrared spectroscopy in novice and expert surgeons--a manifold embedding approach. *Med Image Comput Comput Assist Interv* 2007; **10**(Pt 2): 270-7.

362. Tu FC, Wu WY, Lin HH, Hsiao SM. Laparoscopic intracorporeal knot tying using a novel device. *Surg Endosc* 2012; **26**(3): 872-6.

363. Rosser JC, Jr., Colsant BJ, Lynch PJ, Herman B, Klonsky J, Young SM. The use of a "hybrid" trainer in an established laparoscopic skills program. *JSLS* 2006; **10**(1): 4-10.

364. Leff DR, James DR, Orihuela-Espina F, Yang GZ, Darzi AW. The frontal cortex is activated during learning of endoscopic procedures (Ohuchida et al., Surgical Endoscopy, January 2009). *Surg Endosc* 2010; **24**(4): 968-9.

365. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971; **9**(1): 97-113.

366. Schlaug G, Knorr U, Seitz R. Inter-subject variability of cerebral activations in acquiring a motor skill: a study with positron emission tomography. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale* 1994; **98**(3): 523-34.

367. Taskforce. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *European Heart Journal* 1996; **17**(3): 354-81.

368. Poldrack RA. Imaging brain plasticity: conceptual and methodological issues--a theoretical review. *Neuroimage* 2000; **12**(1): 1-13.

369. Dreyfus LHD, S. E. . Mind over machine: the power of human intuitive expertise in the era of the computer. New York, United States of America: Free Press; 1988.

370. Gobel EW, Parrish TB, Reber PJ. Neural correlates of skill acquisition: decreased cortical activity during a serial interception sequence learning task. *Neuroimage* 2011; **58**(4): 1150-7.

371. Durantin G, Gagnon JF, Tremblay S, Dehais F. Using near infrared spectroscopy and heart rate variability to detect mental overload. *Behavioural brain research* 2014; **259**: 16-23.

372. Herff C, Heger D, Fortmann O, Hennrich J, Putze F, Schultz T. Mental workload during n-back task-quantified in the prefrontal cortex using fNIRS. *Front Hum Neurosci* 2013; **7**: 935.

373. Mandrick K, Derosiere G, Dray G, Coulon D, Micallef JP, Perrey S. Prefrontal cortex activity during motor tasks with additional mental load requiring attentional demand: a near-infrared spectroscopy study. *Neuroscience research* 2013; **76**(3): 156-62.

374. Doi H, Nishitani S, Shinohara K. NIRS as a tool for assaying emotional function in the prefrontal cortex. *Front Hum Neurosci* 2013; **7**: 770.

375. Glascher J, Hampton AN, O'Doherty JP. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex* 2009; **19**(2): 483-95.

376. Robertson EM, Pascual-Leone A, Miall RC. Current concepts in procedural consolidation. *Nat Rev Neurosci* 2004; **5**(7): 576-82.

377. Buxton RB, Uludag K, Dubowitz DJ, Liu TT. Modeling the hemodynamic response to brain activation. *Neuroimage* 2004; **23 Suppl 1**: S220-33.

378. Puckett AM, Mathis JR, DeYoe EA. An investigation of positive and inverted hemodynamic response functions across multiple visual areas. *Hum Brain Mapp* 2014; **35**(11): 5550-64.

379. Watanabe H, Homae F, Taga G. Activation and deactivation in response to visual stimulation in the occipital cortex of 6-month-old human infants. *Dev Psychobiol* 2012; **54**(1): 1-15.

380. Kozberg MG, Chen BR, DeLeo SE, Bouchard MB, Hillman EM. Resolving the transition from negative to positive blood oxygen level-dependent responses in the developing brain. *Proc Natl Acad Sci U S A* 2013; **110**(11): 4380-5.

381. Curtin A, Izzetoglu K, Reynolds J, et al. Functional near-infrared spectroscopy for the measurement of propofol effects in conscious sedation during outpatient elective colonoscopy. *Neuroimage* 2014; **85 Pt 1**: 626-36.

382. Southgate V, Begus K, Lloyd-Fox S, di Gangi V, Hamilton A. Goal representation in the infant brain. *Neuroimage* 2014; **85 Pt 1**: 294-301.

383. Banaji M, Mallet A, Elwell CE, Nicholls P, Cooper CE. A model of brain circulation and metabolism: NIRS signal changes during physiological challenges. *PLoS computational biology* 2008; **4**(11): e1000212.

384. Kolyva C, Tachtsidis I, Ghosh A, et al. Systematic investigation of changes in oxidized cerebral cytochrome c oxidase concentration during frontal lobe activation in healthy adults. *Biomed Opt Express* 2012; **3**(10): 2550-66.

385. Seiyama A, Higaki K, Takeuchi N, Uehara M, Takayama N. Estimation of Skin Blood Flow Artefacts in NIRS Signals During a Verbal Fluency Task. *Adv Exp Med Biol* 2016; **876**: 327-34.

386. Takahashi T, Takikawa Y, Kawagoe R, Shibuya S, Iwano T, Kitazawa S. Influence of skin blood flow on near-infrared spectroscopy signals measured on the forehead during a verbal fluency task. *Neuroimage* 2011; **57**(3): 991-1002.

387. Raghu Prasad MS, Manivannan M, Chandramohan SM. Effects of laparoscopic instrument and finger on force perception: a first step towards laparoscopic force-skills training. *Surg Endosc* 2015; **29**(7): 1927-43.

388. Shibuya K, Sadamoto T, Sato K, Moriyama M, Iwadate M. Quantification of delayed oxygenation in ipsilateral primary motor cortex compared with contralateral side during a unimanual dominant-hand motor task using near-infrared spectroscopy. *Brain Res* 2008; **1210**: 142-7.

389. Waldvogel D, van Gelderen P, Ishii K, Hallett M. The effect of movement amplitude on activation in functional magnetic resonance imaging studies. *J Cereb Blood Flow Metab* 1999; **19**(11): 1209-12.

390. Kuboyama N, Nabetani T, Shibuya K, Machida K, Ogaki T. The effect of maximal finger tapping on cerebral activation. *Journal of physiological anthropology and applied human science* 2004; **23**(4): 105-10.

391. Kuboyama N, Nabetani T, Shibuya K, Machida K, Ogaki T. Relationship between cerebral activity and movement frequency of maximal finger tapping. *Journal of physiological anthropology and applied human science* 2005; **24**(3): 201-8.

392. Obrig H, Hirth C, Junge-Hulsing JG, et al. Cerebral oxygenation changes in response to motor stimulation. *J Appl Physiol (1985)* 1996; **81**(3): 1174-83.

393. Sato H, Fuchino Y, Kiguchi M, et al. Intersubject variability of near-infrared spectroscopy signals during sensorimotor cortex activation. *J Biomed Opt* 2005; **10**(4): 44001.

394. Parasuraman R, Rizzo, M. Neuroergonomics: The Brain At Work. 198 Madison Avenue, New York, United States: Oxford University Press; 2003.

395. James DR, Orihuela-Espina F, Leff DR, et al. Cognitive burden estimation for visuomotor learning with fNIRS. *Med Image Comput Comput Assist Interv* 2010; **13**(Pt 3): 319-26.

396. A HR, Domenico SI, Ayaz H, Gulrajani S, Lam J, Ruocco AC. Differentiating functions of the lateral and medial prefrontal cortex in motor response inhibition. *Neuroimage* 2014; **85 Pt 1**: 423-31.

397. Ikegami T, Taga G. Decrease in cortical activation during learning of a multi-joint discrete motor task. *Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale* 2008; **191**(2): 221-36.

398. Draganski B, May A. Training-induced structural changes in the adult human brain. *Behavioural brain research* 2008; **192**(1): 137-42.

399. Heitger MH, Goble DJ, Dhollander T, et al. Bimanual motor coordination in older adults is associated with increased functional brain connectivity--a graph-theoretical analysis. *PloS one* 2013; **8**(4): e62133.

400. Nan W, Rodrigues JP, Ma J, et al. Individual alpha neurofeedback training effect on short term memory. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology* 2012; **86**(1): 83-7.

401. Wang JR, Hsieh S. Neurofeedback training improves attention and working memory performance. *Clin Neurophysiol* 2013; **124**(12): 2406-20.

402. Zoefel B, Huster RJ, Herrmann CS. Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance. *Neuroimage* 2011; **54**(2): 1427-31.

403. Coyle SM, Ward TE, Markham CM. Brain-computer interface using a simplified functional near-infrared spectroscopy system. *J Neural Eng* 2007; **4**(3): 219-26.

404. Power SD, Kushki A, Chau T. Automatic single-trial discrimination of mental arithmetic, mental singing and the no-control state from prefrontal activity: toward a three-state NIRS-BCI. *BMC research notes* 2012; **5**: 141.

405. Naito MM, Y.; Ozawa, K.; Kiguchi, M.; Kanazawa, T. . A communication means for totally locked-in ALS patients based on changes in cerebral blood volume measured with near-infrared light. *IEICE - Transactions on Information and Systems* 2007; **E90-D**(7): 1028-37.

406. Jurman G, Riccadonna S, Furlanello C. A comparison of MCC and CEN error measures in multi-class prediction. *PloS one* 2012; **7**(8): e41882.

407. Tsuzuki D, Cai DS, Dan H, et al. Stable and convenient spatial registration of stand-alone NIRS data through anchor-based probabilistic registration. *Neuroscience research* 2012; **72**(2): 163-71.

408. Elwell CE, Cope M, Edwards AD, Wyatt JS, Delpy DT, Reynolds EO. Quantification of adult cerebral hemodynamics by near-infrared spectroscopy. *J Appl Physiol (1985)* 1994; **77**(6): 2753-60.

409. Obrig H, Wenzel R, Kohl M, et al. Near-infrared spectroscopy: does it function in functional activation studies of the adult brain? *International journal of psychophysiology : official journal of the International Organization of Psychophysiology* 2000; **35**(2-3): 125-42.

410. Abdi H. RV Coefficient and Congruence Coefficient. In: Salkind N, ed. Encyclopedia of Measurement and Statisitics. 1st ed. New York: Sage Publications; 2007: 849.

411. Josse JP, J.; Husson, F. Testing the significance of the RV Coefficient. *Computational Statistics & Data Analysis* 2008; **53**(1): 82-91.

412. Gaonkar B, Davatzikos C. Analytic estimation of statistical significance maps for support vector machine based multi-variate image analysis and classification. *Neuroimage* 2013; **78**: 270-83.

413. Suykens JAVG, T.; De Brabanter, J.; De Moor, B.; Vandewalle, J.; . Least squares support vector machines. Singapore: World Scientific Pub. Co.; 2002.

414. Press WHT, S. A.; Vetterling, W. T.; Flannery, B. P. Simulated Annealing Methods. Numerical Recipes: The Art of Scientific Computting. 3rd ed. New York: Cambridge University Press; 2007: 549.

415. Gorodkin J. Comparing two K-category assignments by a K-category correlation coefficient. *Comput Biol Chem* 2004; **28**(5-6): 367-74.

416. Bullmore ET, Bassett DS. Brain graphs: graphical models of the human brain connectome. *Annual review of clinical psychology* 2011; **7**: 113-40.

417. Rissman J, Gazzaley A, D'Esposito M. Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage* 2004; **23**(2): 752-63.

418. Bernardi G, Ricciardi E, Sani L, et al. How skill expertise shapes the brain functional architecture: an fMRI study of visuo-spatial and motor processing in professional racing-car and naive drivers. *PloS one* 2013; **8**(10): e77764.

419. Leff DR, Leong, J., F., Atallah, L., Athansiou, T., Hui Koh, P., Elwell, C., Delpy, D., Yang, G.Z., Darzi, A. Open Surgical Knot-Tying Induces a Lateralised Prefrontal Brain Response in Surgical Novices. 13th meeting of the Organization for Human Brain Mapping. Chicago, Illinois, United States; 2007.

420. Jenkins IH, Brooks DJ, Nixon PD, Frackowiak RS, Passingham RE. Motor sequence learning: a study with positron emission tomography. *J Neurosci* 1994; **14**(6): 3775-90.

421. Nachev P, Kennard C, Husain M. Functional role of the supplementary and pre-supplementary motor areas. *Nat Rev Neurosci* 2008; **9**(11): 856-69.

422. Wymbs NF, Grafton ST. The Human Motor System Supports Sequence-Specific Representations over Multiple Training-Dependent Timescales. *Cereb Cortex* 2014.

423. Hong KS, Naseer N, Kim YH. Classification of prefrontal and motor cortex signals for three-class fNIRS-BCI. *Neurosci Lett* 2015; **587**: 87-92.

424. Zimmermann R, Marchal-Crespo L, Lambercy O, et al. Towards a BCI for sensorimotor training: initial results from simultaneous fNIRS and biosignal recordings. *Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference* 2011; **2011**: 6339-43.

425. Power SD, Kushki A, Chau T. Towards a system-paced near-infrared spectroscopy brain-computer interface: differentiating prefrontal activity due to mental arithmetic and mental singing from the no-control state. *J Neural Eng* 2011; **8**(6): 066004.

426. Naseer N, Hong KS. fNIRS-based brain-computer interfaces: a review. *Front Hum Neurosci* 2015; **9**: 3.

427. Guillou PJ, Quirke P, Thorpe H, et al. Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial): multicentre, randomised controlled trial. *Lancet* 2005; **365**(9472): 1718-26.

428. Healey AN, Sevdalis N, Vincent CA. Measuring intra-operative interference from distraction and interruption observed in the operating theatre. *Ergonomics* 2006; **49**(5-6): 589-604.

429. Zheng B, Tien G, Atkins SM, et al. Surgeon's vigilance in the operating room. *Am J Surg* 2011; **201**(5): 673-7.

430. Staal MA. Stress, cognition, and human perfor- mance: A literature review and conceptual framework, NASA Technical Memorandum 2004.

431. Fafrowicz M, Marek T. Quo vadis, neuroergonomics? *Ergonomics* 2007; **50**(11): 1941-9.

432. Hart SGS, L. E. Development of NASA-TLX (task load index): Results of empirical and theoretical research In: Hancock PAM, N, ed. Human Mental Workload. North-Holland, Amsterdam; 1988: 139-83.

433. Just MAC, P. A.; Miyake, A. Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science* 2003; **4**(1-2): 56-88.

434. Gramann K, Ferris DP, Gwin J, Makeig S. Imaging natural cognition in action. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology* 2014; **91**(1): 22-9.

435. Makeig S, Gramann K, Jung TP, Sejnowski TJ, Poizner H. Linking brain, mind and behavior. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology* 2009; **73**(2): 95-100.

436. Yang G-ZM, G. P.; Kwok, K.-W.; Chung, A. Perceptual Docking for Robotic Control. In: Dohi TS, I.; Liao, H., editor. Medical Imaging and Augmented Reality. Berlin: Springer-Verlag; 2008. p. 21-30.

437. Rotting MZ, T. O.; Trosterer, S.; Dzaack, J. Implicit Interaction in Multimodal Human-Machine Systems. In: Schlick CM, ed. Industrial Engineering and Ergonomics. Berlin: Springer; 2009: 523-36.

438. Zander TO, Kothe C. Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *J Neural Eng* 2011; **8**(2): 025005.

439. Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. *Clin Neurophysiol* 2002; **113**(6): 767-91.

440. Zander TOK, C.; Jatzev, S.; Gaertner, M. Enhancing human-computer interaction with in- put from active and passive brain-computer inter- faces. In: Tan DSN, A., ed. Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction. London: Springer-Verlag; 2010: 181-99.

441. Zander TO, Jatzev S. Context-aware brain-computer interfaces: exploring the information space of user, technical system and environment. *J Neural Eng* 2012; **9**(1): 016003.

442. Kohlmorgen JD, G.; Braun, M.; Blankertz, B.; Muller, K. R.; Curio, G.; Hagemann, K.; Bruns, A.; Schrauf, M.; Kincses, W. Improving human performance in a real operating environment through real-time mental workload detection. In: Dornhege GM, J. R.; Hinterberger, T.; McFarland, D. J.; Muller, K.-R., ed. Toward Brain-Computer Interfacing. Cambridge: MIT Press; 2007: 409-22.

443. Grimes DT, D.S.; Hudson, S. E.; Pradeep, S.; Rao, R. P. N. Feasibility and pragmatics of classifying working memory load with an electroencephalograph. In: Czerwinski M, editor. Pro SIGCHI conf on Human factors in computing systems: ACM; 2008. p. 835-44.

444. Gevins A, Smith ME. Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cereb Cortex* 2000; **10**(9): 829-39.

445. Wickens CDM, J. S. Applied Attention Theory: CRC Press; 2007.

446. Gevins AS, M. E. Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science* 2003; **4**(1-2): 113-31.

447. Holm A, Lukander K, Korpela J, Sallinen M, Muller KM. Estimating brain load from the EEG. *ScientificWorldJournal* 2009; **9**: 639-51.

448. Smith ME, Gevins A, Brown H, Karnik A, Du R. Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Hum Factors* 2001; **43**(3): 366-80.

449. Tomioka RD, G.; Nolte, G.; Blankertz, B.; Aihara, K.; Muller, K.-R. Spectrally weighted common spatial pattern algorithm for single trial EEG classification. Tokyo, Japan: University of Tokyo, 2006.

450. Duda ROH, P. E.; Stork, D. G. Pattern Classification: Wiley; 2000.

451. Blankertz B, Lemm S, Treder M, Haufe S, Muller KR. Single-trial analysis and classification of ERP components--a tutorial. *Neuroimage* 2011; **56**(2): 814-25.

452. Vapnik VNC, A. Y. On theuniform convergence of relative frequencies of events to their probabilities. *Theory of Probability & Its Applications* 1971; **16**(2): 264-80.

453. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 2004; **134**(1): 9-21.

454. Makeig S. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalogr Clin Neurophysiol* 1993; **86**(4): 283-93.

455. Sonnleitner A, Treder MS, Simon M, et al. EEG alpha spindles and prolonged brake reaction times during auditory distraction in an on-road driving study. *Accid Anal Prev* 2014; **62**: 110-8.

456. Gramann K, Gwin JT, Bigdely-Shamlo N, Ferris DP, Makeig S. Visual evoked responses during standing and walking. *Front Hum Neurosci* 2010; **4**: 202.

457. Gwin JT, Gramann K, Makeig S, Ferris DP. Electrocortical activity is coupled to gait cycle phase during treadmill walking. *Neuroimage* 2011; **54**(2): 1289-96.

458. Zander TOB, J.; Lorenz, R.; Krol, L. R. Towards bci-based implicit control in human-computer interaction. In: Fairclough SG, K., ed. Advances in Physiological Computing. London: Springer_Verlag; 2014: 67-90.