Original Manuscript Submission to Annals of Surgery

Manuscript Title

"Contemplating The Next Manoeuvre"

Functional NeuroImaging Reveals Intra-Operative Decision Making Strategies

Manuscript Authors

Daniel Richard Leff FRCS (Gen Surg) PhD[1,2], Gabriella Yongue MBBS BSc[1,2], Ivo Vlaev PhD[2], Felipe Orihuela-Espina PhD[1,4], David James MRCS PhD[1,2], Michael Taylor PhD[2], Thanos Athanasiou PhD FECTS[2], Ray Dolan MRCPsych MD FRS[3], Guang-Zhong Yang FREng PhD[1], and Ara Darzi MD FRCS FACS FMedSci FRS[1,2]

Author Affiliations

1=Hamlyn Centre for Robotic Surgery, Imperial College London

2=Department of BioSurgery and Surgical Technology, Imperial College London

3=Welcome Trust Centre for NeuroImaging, University College London

4=Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Mexico

Corresponding Author

Professor Ara Darzi MD FACS FMedSci FRS

Hamlyn Chair of Surgery

Academic Surgical Unit, 10th floor, QEQM Wing

St Mary's Hospital

London

United Kingdom

W2 1NY

Tel: +44 (0)20 3312 1310

Fax: +44 (0)20 3312 6950

Email: a.darzi@imperial.ac.uk

AUTHOR CONTRIBUTIONS

The study design was conceived and developed by DRL, GY, IV, RD, GZY and AD. DRL, DJ, IV executed the experiment and collected the functional imaging data. Data pre-processing and statistical analysis was conducted by DRL, GY, FOE, MT and TA. Data interpretation was performed by DRL, FOE, IV in consultation with RD, GZY and AD. The manuscript was drafted by DRL, GY, and IV. Critical editing of the manuscript was performed by RD, GZY and AD.

MINI ABSTRACT

A functional neuroimaging study of intra-operative decision-making was conducted that suggests the transition from novice to experts is characterised by a switch from an effortful goal orientated system that relies on the prefrontal cortex to a recognition-primed system that is accompanied by prefrontal redundancy.

ABSTRACT

Objective: To investigate expertise related differences in the quality, certainty and consistency of intra-operative surgical decision-making (DM) and characterise the decision systems operators' employ using functional neuroimaging.

Background: Intra-operative DM strategy has been under-investigated and is more likely to be deduced from patterns of brain activation during DM rather than an analysis of operative manoeuvres per se. Novices are hypothesised to rely on effortful, goal orientated DM which is known to lead to activation across the dorsolateral prefrontal cortex (DLPFC), whereas experts are expected to utilise habitual DM in which decisions are ingrained, recognition-primed and PFC independent.

Methods: 22 subjects (10 medical student novices, 7 residents and 5 attendings) participated in a study that required them to review intra-operative videos (simulated laparoscopic cholecystectomy), determine the next safest operative manoeuvre upon video termination (10s), and report their decision certainty. Decision-making paradigms were classified according whether or not the operator's next technical move was declared ('primed') or hidden ('unprimed'). Simultaneously, changes in cortical oxygenated haemoglobin (HbO_2) and deoxygenated haemoglobin (HHb) inferring prefrontal activation (i.e. HbO_2 increase and HHb decrease), were recorded using Optical Topography (ETG-4000, Hitachi Medical Corp). Decision certainty, consistency (primed versus unprimed) and quality (script concordance) were assessed.

Results: Attendings and residents were significantly more certain of decisions ($p<0.001$) and decision quality was observed to be superior to novices (script concordance: attendings = 90%, residents = 78.3%, novices = 53.3%). Decision consistency (primed versus un-primed) was significantly superior in both experts ($p<0.001$) and residents ($p<0.05$) compared to novices ($p=0.183$). During un-primed DM, novices significantly activated the DLPFC whereas significant activation was not observed amongst residents and attendings. Medial PFC activation trends were observed only in novices. During primed DM, significant activation was not observed in any group.

Conclusion: Expert DM is characterised by improved quality, consistency and certainty. Attendings employ a habitual decision system that appears to be PFC independent, whereas novices utilise an effortful goal-orientated approach under uncertainty. In the presence of operative cues (primes) novices *disengage* the PFC and appear to accept the observed decision as correct.

Keywords: decision-making, functional, near-infrared spectroscopy, brain, prefrontal, simulation, surgery, training,

Main Manuscript

A surgeon's ability to make reasoned judgements under pressure during operative interventions influences surgical workflow and patient safety. Accurate perception and interpretation of the dynamic nature of the operative scene known as situational awareness (SA)[1] and appropriate decision-making (DM) to guide sequential operative manoeuvres should be considered safety-critical skills. Yet, whilst there has been a systematic focus on training and assessment of technical skills, research pertaining to surgical cognition in general[2, 3] and operative situational awareness[4] or DM more specifically[5] are scant, possibly due to the challenges associated with investigating complex executive functions[6].

Operative DM can be simplified as a continuous cycle of monitoring and SA, appropriate action taking and outcome evaluation to update and improve the operator's DM system[6]. As illustrated in Figure 1, within this model exist a range of DM strategies that can be actioned depending upon the available time, perceived risk to the patient and experience of the operator. For example, expert surgeons encountering a familiar operative scene are anticipated to engage a *recognition-primed* approach to select solutions from memory. Conversely, residents with limited domain experience are hypothesised to associate operative scenes with a set of action rules known as "habit learning" (or habitual DM which involves learning the value of actions in different states of the world), or to use analytical DM to compare and contrast the perceived risks, associated with a range of possible solutions (e.g. 'dissect' versus 'divide'), known as "goal-directed learning" (or goal-directed DM which involves explicit knowledge of the action-outcome contingencies)[7, 8]. Furthermore, for the expert trainer guiding a resident through an intervention, SA also involves assessments of the trainee's DM system, allowing the procedure to flow where trainer-resident DM appears congruent but importantly knowing when to veto incorrect decisions and take back control. The latter often relies on an incongruent behavioural trigger or cue such as the resident inserting a pair of scissors when the trainer perceives that more dissection is required. Experimentally, surgical simulation facilitates manipulation of behavioural cues, which can be covertly introduced as an "unconscious prime" to investigate the impact they may have on trainer DM.

Critically, expertise in operative DM is unlikely to be revealed in behavioural responses such as action selection or choice of operative manoeuvres *per se* since the internal rumination of "*what to do next*" in surgery does not have a behavioural correlate that can be linearly mapped. Instead, we anticipate that disparities in intra-operative DM manifest as differences in the internal decision systems and cognitive strategies operators' employ. Therefore, the scientific challenge is how to reliably interrogate surgeons to unveil operative DM strategy. Whilst post-event interviewing of surgeons provides a degree of insight[5, 9, 10] the approach is time-consuming, subjective and cannot be used to anchor residents' progress through training. An alternative strategy is to capitalise on developments in non-invasive functional neuroimaging technologies to monitor operator brain function during operative interventions on the basis that the magnitude or pattern of cortical response correlates with the decision system utilised.

The brain contains multiple, distinct decision systems[7, 8, 11-13] differentiated according to their engagement of the corticostriatolimbic circuits in the brain[14]. Each system assigns a 'value' to available actions, and thus compete with the actions favoured by other systems[15]. Recent evidence indicates competition between a cognitive, goal-directed planning system centred in the lateral prefrontal cortex and parietal cortex, and habitual decision system associated with dopamine and the basal ganglia[16, 17]. Decisions requiring effort, working memory and deductive reasoning have been shown to activate the dorsolateral prefrontal cortex (DLPFC)[18-20], while habitual decisions are stimulus-response associations learned through repeated practice and rewards in a stable environment (such mental habits are usually the consequence of past goal pursuits, but once acquired, habits are cued and performed without mediation of a goal)[13]. As one's experience accumulates, control over decisions gradually transfers from goal-directed process, which demand effort and time, to the habitual processes which are rapid and easy to execute[8]. Based on this evidence and DM theories already outlined, novice surgeons are expected to recruit the DLPFC to a greater extent than expert surgeons owing to escalated levels of uncertainty, need for internal cross-referencing and more detailed analysis of options during operative DM.

METHODS

Subjects

Following local regional ethical approval (LREC: 05/Q0403/142), 22 healthy individuals were recruited from Imperial College London and Imperial College Healthcare NHS Trust. Participants were subdivided into three groups according to prior operative expertise in laparoscopic cholecystectomy as follows: 10 medical students [mean age ± SD (years) = 22.40 ±0.97] with no prior experience of laparoscopy were classified as 'novices'. 7 participants were 'residents' enrolled in specialty training schemes [mean age ± SD (years) 32.14 ± 1.77] and had prior experience of assisting on laparoscopic cholecystectomy or performing the procedure under supervision. Finally, 5 attendings were classified as 'experts' [mean age ± SD (years) = 32.14 ± 1.77] on the basis of more than 100 independent laparoscopic cholecystectomies. A history of neuropsychiatric disorders was an exclusion criterion (n=0) and all participants were asked to refrain from alcohol and caffeine for 24hours given the known effects on cerebral haemodynamics[21].

Task and Training

Prior to the experiment, all subjects were provided with a training session that included an overview of the operative anatomy and principles of laparoscopic cholecystectomy as well as the operative steps (i.e. Calot's dissection, clipping of cystic artery and duct, etc). Following training, subjects' were asked to complete a short test that posed questions to test knowledge and understanding of the operative anatomy and procedural flow of laparoscopic cholecystectomy (see questionnaire supplementary appendix). Failure to achieve perfect score in the test led to exclusion (n=0). Following successful test completion, subjects proceeded to the DM experiment.

Operative Decision-Making Paradigm Experimental Set-up

The experiment focused on interrogating intra-operative DM during laparoscopic cholecystectomy. Subjects were asked to regard a monitor and observe a series of video clips (n=12) of high-fidelity simulated laparoscopic cholecystectomy (pre-recorded using LapMentor, Simbionix, Israel). Each video clip lasted 10s, revealed a sequence of operative manoeuvres at random (i.e. unpredictable), and terminated at a point at which an operative decision was required. Video clips were classified as either *"primed"* (n=5) in which the operator's next step was readily declared (e.g. scissors brought into view suggesting DM to cut), or *"un-primed"* (n=7) which terminated immediately after a given action without indication of what occurred next in the simulation (Fig. 2a.). The sequence in which subjects experienced primed and un-primed video clips was randomised. After each video clip subjects were asked to verbally report the recommended next operative manoeuvre from a list provided on the monitor. Each operative decision was recorded by the investigators (DRL, GY). Following the DM task, subjects were asked to state how certain they were of their decision on a scale of one to six (1=low certainty, 6= high certainty).

Experimental Set-up and Block Design Experiment

As illustrated in Figure 2b, a block design experiment was conducted comprising twelve sequential blocks, each comprising episodes of “rest”, and three *stimuli* identified as “video review”, “decision” and “certainty”. During rest periods (30s) subjects were seated asked to place their hands on a table and regard a fixation cross. During video review subjects were instructed to pay close attention to the operative video clip (10s) with a view to reporting the next operative manoeuvre upon video termination. During decision episodes a slide was presented to the subject as an aide memoire of the surgical options (e.g. dissect further, divide cystic artery, convert to open, etc) and subjects verbally reported their decision (10s). Finally, subjects reported decision certainty (10s). Before progression to the next video clip, a post trial rest period (30s) was introduced to enable cortical haemodynamics to return to baseline. Cortical activity was measured throughout using fNIRS-based Optical Topography (OT) which converts changes in light levels into changes in cortical haemodynamics[22] and therefore monitors the haemodynamic response to neuronal activation (“neurovascular coupling principle”[23]. The typical haemodynamic response to neuronal activation comprises a rise in oxygenated haemoglobin (HbO_2) and a decrease in deoxygenated haemoglobin (HHb).

Functional Neuroimaging

Subjects’ were neuro-monitored using a commercial OT system (ETG-4000, Hitachi Medical Corp., Japan). OT is a portable, non-invasive technique that is resistant to motion artefact and has been successfully used in the study of technical skills across the field of surgery[24 2, 25-27]. Multichannel OT is a technique that measures changes in light levels across multiple cortical locations simultaneously. Light is shone on to the subject's scalp (700-900nm) and attenuated light is detected by neighbouring photodiode detectors. The modified Beer-Lambert Law[28] was used to compute relative changes in haemoglobin concentration at multiple locations between emitters and detectors (referred to as 'channels'). Here, 15 optodes (emitters / detectors) were deployed 30mm apart in a 5 x 3 flexible plastic array positioned according to the 10-20 system of electrode placement[29] to monitor haemodynamic change across the PFC, as illustrated in Fig. 2b. NIR light at 695 and 830nm was emitted from 8 optical fibre sources and detected by 7 neighbouring avalanche photodiode detectors, resulting in 22 different measuring channels. Probes were fastened into C-shaped metallic holders and the entire array was secured to the operator's scalp using surgical bandage (Surgifix, Colorline, Italy) as highlighted in Fig. 2b.

Stress

Subjective levels of stress were monitored on the basis that stress related changes in systemic physiology might influence functional OT data[30]. Subjects' were asked to complete short form of the Spielberger State-Trait Anxiety Inventory (STAI) before, during and after the study.

Data Processing and Statistical Analysis

Decision Quality, Consistency and Certainty

The quality of DM responses was assessed using script concordance, which is a tool designed to assess clinical reasoning on the basis that judgement can be probed and concordance with a reference panel of experts measured[31]. Script concordance is calculated by scoring each decision by comparing it to the DM of a panel of experts. Here, we invited a panel of expert consultant surgeons not recruited to the study (n=10) to review each laparoscopic cholecystectomy video used in the experiment and record what was in their expert opinion the correct next operative move. In this regard, we obtained consensus as to the most appropriate next operative step and hence were able to award points for participant DM based on the expert responses (Table 1). Decision consistency was determined by correlating decisions for each 'primed' video with the 'un-primed' equivalent (10 videos) using Spearman correlation analysis. Decision certainty scores were tabulated according to operator expertise and decision type (i.e. 'un-primed' and 'primed'). The Chi square test was used to compare certainty between the three different experience groups and also within experience group between 'un-primed' and 'primed' conditions. For statistical analysis of decision quality, consistency and certainty $p<0.05$ was deemed statistically significant.

Functional Neuroimaging Data

Functional neuroimaging data was analysed using the Imperial College Neuroimaging Analysis (ICNA), a bespoke software package programmed using Matlab (Mathworks, USA). Raw optical data was subject to integrity checks to eliminate instrumentation noise, system drift, optode mirroring and apparent non-recording as well as to increase signal to noise ratio[21]. Data was decimated and linearly de-trended and relative changes in light intensities were converted into changes in haemoglobin concentration using the modified Beer-Lambert Law[28].

For a given experience group, haemodynamic time courses were produced for each of the 22 channels and visually inspected to identify areas consistent with activation *i.e.* increases in HbO_2 or decreases in HHb, and confirmed using a statistical channel-based analytical framework referred to as the "activation matrix". Activity matrices were constructed by assessing task-induced changes in both HbO_2 and HHb. For each channel, average baseline rest Hb data (5s of data prior to stimulus onset) was compared to average trial Hb data (17s of data, 2s following stimulus onset) using the Wilcoxon Sign Rank test. Channels displaying statistically significant ($p<0.05$) increases in HbO_2 coupled to statistically significant ($p<0.05$) decreases in HHb were considered *activated*. Conversely, channels displaying the opposing trend were considered *deactivated*. Channels in which directional changes in Hb species were commensurate with either activation or deactivation but for which only one Hb species reached statistical threshold were termed 'activation or deactivation trends'.

Regarding channels displaying activation or activation trends, a new variable termed "ΔHb" was computed to compare the magnitude of cortical haemodynamic changes between experience groups. For each channel and Hb species, ΔHb represented the difference between rest Hb data and stimulus Hb data (i.e. ΔHb = Δ stimulus Hb - Δ rest Hb). Here, rest data was calculated by averaging the last 5s of each rest period prior to the video presentation, whilst stimulus data represented the average of 17s epochs commencing 2s after the stimulus onset. For a given channel, ΔHb data was compared between novices and operators with either prior laparoscopic training or real operative experience (i.e. residents and attendings combined) using the Mann Whitney U test. ΔHb data were further grand averaged across DLPFC channels to obtain individual proxy indicators of brain activity (thus allowing one observation per-trial per-individual). Finally, a Generalized Linear Mixed Model (GLMM) was computed across and within each expertise group, using grand averaged ΔHb data, with Δ HbO_2 and Δ HHb – as the dependent variable; and priming condition (primed vs. unprimed) as fixed effects (within-subject factor); and subjects, trial number, and stimulus as random effects.

Stress Data

Within group comparisons in STAI responses before, during and after the experiment was analysed using the Wilcoxon Signed Rank test.

RESULTS

Cohort Demographics

7 female and 15 male subjects participated. No significant gender distribution differences (χ^2 = 1.45, p=0.483), or differences in handedness (χ^2 = 5.87, p=0.209) were identified between the groups. Participant's ages ranged from 21 to 51 years and experts were significantly older than residents [mean age ± SD (years): attendings = 36.20 ± 8.79, residents = 32.14 ± 1.77, p<0.05] and novices [mean age ± SD (years) = attendings = 36.20 ± 8.79, novices = 22.40 ±0.97, p<0.05].

Operative Decision Certainty

As depicted in Figure 3, DM certainty varied significantly with expertise (p<0.001). A greater proportion of attendings' were observed to be highly certain of operative decisions versus residents and novices (% reporting high certainty: attendings' = 73%, residents = 60%, novices = 11%). Both attendings and residents were significantly more certain of decisions than novices (mean certainty ± SD: novices = 3.95 ± 1.20, residents = 5.37 ± 0.94, experts = 5.68 ± 0.60; attendings vs novices χ^2 = 87.35, p<0.001, residents vs novices χ^2 = 71.22, p<0.001). However, there was no statistical difference in DM certainty between residents and attendings (χ^2 = 7.31, p=0.120). Priming had no significant impact on decision certainty regardless of operator experience (novices: χ^2 = 3.60 p=0.730, residents: χ^2 = 2.18, p=0.702, attendings: χ^2 = 1.84, p=0.606).

Operative Decision Quality, Decision Consistency and Stress

Script concordance confirmed that residents DM aligned more closely with expert panel DM than novices [script concordance % (score)= attendings = 90 (10.8), residents = 78.3 (9.4), novices = 53.3 (6.4), maximum score= 12)]. Attendings more frequently challenged the apparent next operative move in the primed video sequences, than did residents or novices [contradict prime decision: attendings = 85.0%, residents = 74.0%, novices = 44.0%]. The frequency with which primed cues were challenged varied significantly with expertise (χ^2 = 9.810, p=0.007). There was a lack of consistency in DM between matched unprimed and primed decision stimuli amongst novices (R^2 = 0.191, p=0.183) whereas residents' (R^2 = 0.445, p=0.007) and attendings' responses (R^2 = 0.524, p=0.001) were significantly more consistent across conditions. There was no statistically significant difference in STAI scores between groups (p=0.574). Table 2 summarises within-group analysis of STAI data. No significant changes in stress or anxiety were observed across the experiment amongst residents or attendings. However, comparing STAI scores during and after the experiment confirmed a significant decrease in anxiety amongst novices (p=0.011).

Cortical Haemodynamics

Un-Primed Decisions

Activation matrices for unprimed stimuli are illustrated by operator expertise in Figure 4 (panel a) (see supplementary material for full statistical analysis).

Regarding operative video review, a greater number of PFC channels displayed activation trends amongst novices than residents and attendings (activation trends: novices = 14/12, residents = 4/22, and attendings = 4/22). In addition, whilst amongst residents and attendings activation was observed across bilateral DLPFC, activation amongst novices was predominantly ventromedial. During decision-making trials, activated DLPFC channels (i.e. statistically significant changes in both HbO_2 and HHb species) were only observed amongst novices whereas activation trends were observed across bilateral DLPFC channels amongst residents and attendings (residents = right DLPFC= 4 channels, left DLPFC = 4 channels, attendings = right DLPFC = 2 channels, left DLPFC = 3 channels). Ventromedial activation trends were observed solely amongst novices during DM trials.

Table 3 highlights comparisons between operators in ΔHb data during DM stimuli for bilateral DLPFC channels. DM associated changes in cortical HbO_2 and HHb were substantially greater amongst novices versus operators with prior laparoscopic cholecystectomy experience. As illustrated in Figure 5 and Table 1, trends toward significantly greater activation responses in novices versus residents and attendings were observed in multiple bilateral DLPFC channels (ΔHbO_2: right DLPFC channel 22, ΔHHb: right DLPFC channel 5 and 13, and left DLPFC channel 10). However, a between-group GLMM model did not demonstrate an independent effect of expertise [ΔHbO_2: $F(2,786)=0.56$, $p=.569$; ΔHHb: $F(2,786)= 0.04$, $p=.957$].

Table 4 presents within-group GLMM results including the model's coefficients for the effect of the fixed factor (priming), which reveal the direction and significance of the effects. Overall, the priming effect was observed only for HbO_2 in novices – the significant negative coefficient implies that the priming reduced ΔHbO_2 across the DLPFC.

Primed Decisions

As highlighted in the averaged Hb time course curves in Figure 6, in general, PFC responses during operative DM were less apparent in the primed versus the un-primed condition. Indeed, regardless of expertise, priming did not lead to statistically significant activation either during video review or during DM stimuli. Rather during video review, an inverse relationship was identified between deactivation trends and operator expertise (deactivated channel trends: novices = 1/22, residents = 4/22, and attendings = 5/22). During DM trials, bilateral DLPFC activation trends were identified in novices and residents, whereas no significant cortical haemodynamic change was apparent amongst attendings.

Discussion

In this study, expertise related differences in intra-operative DM performance, consistency and certainty have been investigated, and DM strategies have been exposed using functional neuroimaging. As was hypothesised, expert DM was characterised by superior quality decisions, greater confidence in DM, and a willingness to challenge apparent decisions made by another operator. Furthermore, novice DM in the face of uncertainty (i.e. absence of the behavioural cue or prime) was manifest as greater dorsolateral, ventrolateral and medial PFC activations suggesting a need for greater attention, concentration and mental effort during DM. The introduction of a behavioural trigger that revealed the operator's next operative decision prompted attenuation of prefrontal activation amongst novices and deactivation trends amongst residents and attendings.

Emerging evidence has indicated neural interactions between habitual and goal-directed DM[32]. Transition from goal-orientated to habitual DM is likely to take place during the acquisition of expertise in surgical DM. This is because habits require extensive experience including schedules of reinforcement involving single actions and single outcomes, indicating that behaviour must be initially goal-directed before gradually becoming habitual over the course of experience. Therefore, the observed increased in confidence and quality of expert laparoscopists' DM likely reflects years of repeated exposure to similar operative scenes and reflection regarding the outcomes of their own DM, as well as observation of resident DM. Habitual DM presents stimulus-response associations learned through repeated practice and rewards in a stable environment[33]. Habits are implemented in the subcortical structures- the dorsolateral striatum and dopamine neurons into this area, arriving from substantia nigra and the ventral tegmental area, are important for learning the value of habitual actions and stimulus-response representations can also be encoded in cortico-thalamic loops and the infralimbic (medial) prefrontal cortex[32]. Hence the relative DLPFC and MPFC redundancy during expert DM reflects the establishment of patterns of habitual DM, which is stable and repetitive with similar cues, actions and rewards.

Conversely, the observed prefrontal activation response amongst novices suggests a goal-directed intra-operative DM approach. Goal-directed DM is implemented in different parts of the frontal lobe, concentrating on the anterior cingulate and orbitofrontal cortex, but also subsuming mechanisms localised in hippocampus and dorsomedial striatum[16]. Goal-directed decisions and actions are implemented predominantly in networks that mediate declarative expectations of future outcomes and conscious planning[11, 34]. Three main areas in the prefrontal cortex are used in DM processes, which interact with each other and with subcortical brain regions [18, 35-37]. The orbitofrontal cortex (OFC) has ramifications with the limbic system and plays an important role in reward-based and emotional DM[18, 36, 38, 39]. Effortful decisions depending on working memory and those that involve reasoning cause recruitment of the dorsolateral prefrontal cortex (DLPFC)[18-20] and the anterior cingulate cortex (ACC)[36, 40, 41]. Moreover, studies from patients with frontal lobe lesions confirm the importance of the DLPFC in planning, strategic development, cognitive flexibility and working memory[35, 36, 42]. Decisions requiring cross-reference to the decision maker's value system, incorporation of long-term or contextual information and decisions made under uncertainty are known to burden the DLPFC[18, 37, 43-45]. Finally, goal-directed PFC decision-making specifically involves the ACC during highly ambiguous situations in which the decision maker perceives several conflicting options and a high likelihood of error[36, 37], which also may explain the relative PFC redundancy amongst novices during primed intra-operative DM.

It is interesting to note that when faced with an apparent decision made by another operator (i.e. during surgical cues / behavioural primes), novices infrequently challenge the decision, possibly considering it to be the correct next operative move. Whilst subjects were not informed as to the operator's identity, novices may have assumed that operator was an expert attending. We speculate that in the minds of novices, this incorrectly reduces uncertainty and ambiguity and prompts them to accept the observed decision. This acceptance appears to manifest as a comparative prefrontal disengagement and lack of attention and concentration that was previously required for intra-operative DM under greater uncertainty, i.e. when what to do next was not obvious. Furthermore, expert surgeons primed with the salient cues during familiar operative scenes automatically make the associated decision without further thought, hence the lack of activation in goal-directed decision regions.

In summary, attendings' DM is characterised by greater certainty, improved alignment with an expert reference panel, and reduced reliance on the prefrontal lobe suggesting mature habitual responses. Prefrontal excitation observed in novices implies that the transition from trainee to expert is coupled with a switch from goal orientated to recognition based DM.

Limitations

A number of limitations of this study should be acknowledged. The nature of the experimental paradigm and time required for each subject (e.g. approximately one hour per subject for training, OT probe placement, task familiarisation and experiment) limited the recruitment of attendings. Whilst script concordance is a valid measure of agreement with panel consensus, it does not necessarily follow that the operative decisions made by attendings or indeed the expert panel were all "correct". Indeed, the concept of a single correct next operative decision is challenging to validate and it is more likely that for a given scenario one of several options are safe. This notwithstanding, the aim was to explore the internal cognitive process and cortical responses associated with operative DM and these are not influenced by the specific decision. Put simply, the study primarily sought to address how a decision was arrived at, as opposed to whether the decision was correct or not. It should be acknowledged that the time set aside for DM following video review is artificial, and the internal processing regarding operative decisions are likely to be made continually online. However, the experiment was designed to enable us to isolate DM associated cortical activations, which would not have been feasible in a less controlled experiment. Finally, we accept that given novices felt less stressed following the experiment, stress induced changes in haemodynamics may have contributed to our results.

References

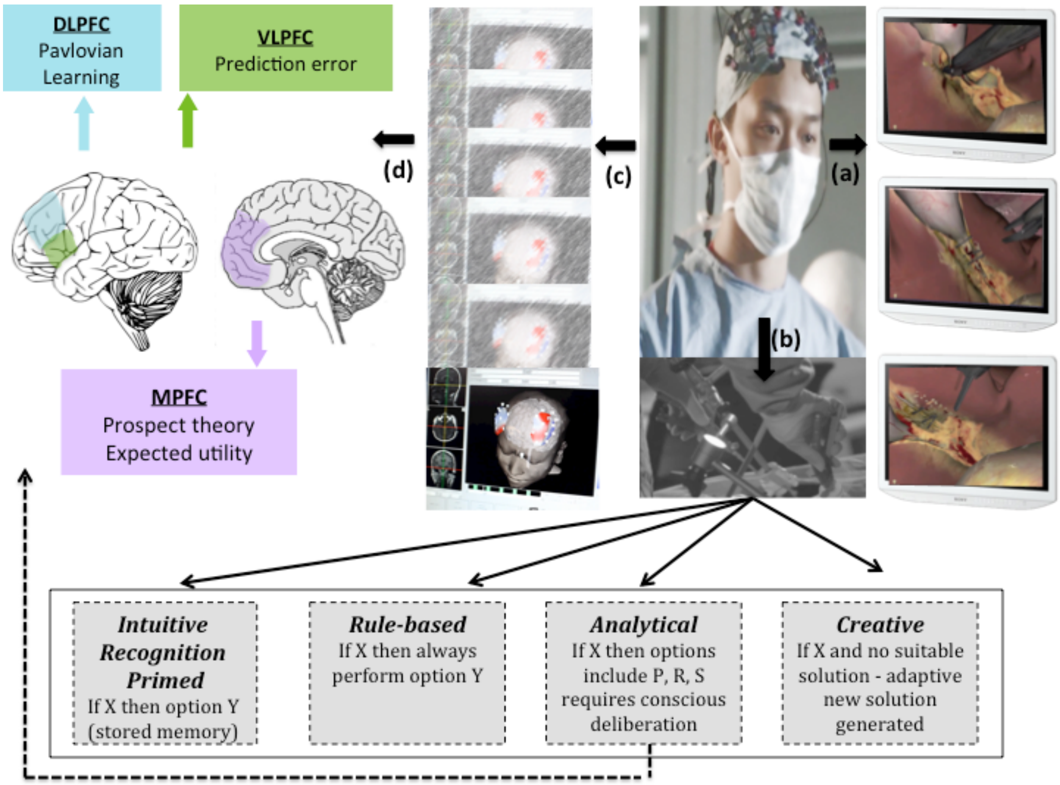
1. Endsley M. Toward a theory of situation awareness indynamic systems. *Hum Factors* 1995; 37:32-64.
2. Leff DR, Orihuela-Espina F, Athanasiou T, et al. "Circadian cortical compensation": a longitudinal study of brain function during technical and cognitive skills in acutely sleep-deprived surgical residents. *Ann Surg* 2010; 252(6):1082-90.
3. Bahrami P, Graham SJ, Grantcharov TP, et al. Neuroanatomical correlates of laparoscopic surgery training. *Surg Endosc* 2014.
4. Graafland M, Schraagen JM, Boermeester MA, et al. Training situational awareness to reduce surgical errors in the operating room. *Br J Surg* 2015; 102(1):16-23.
5. Pauley K, Flin R, Yule S, et al. Surgeons' intraoperative decision making and risk management. *Am J Surg* 2011; 202(4):375-81.
6. Flin R, Youngson G, Yule S. How do surgeons make intraoperative decisions? *Qual Saf Health Care* 2007; 16(3):235-9.
7. Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 2008; 9(7):545-556.
8. Dolan RJ, Dayan P. Goals and habits in the brain. *Neuron* 2013; 80(2):312-25.
9. Pauley K, Flin R, Azuara-Blanco A. Intra-operative decision making by ophthalmic surgeons. *Br J Ophthalmol* 2013; 97(10):1303-7.
10. Cristancho SM, Vanstone M, Lingard L, et al. When surgeons face intraoperative challenges: a naturalistic model of surgical decision making. *Am J Surg* 2013; 205(2):156-62.
11. Dolan Ray J, Dayan P. Goals and Habits in the Brain. *Neuron* 2013; 80(2):312-325.
12. Gottlieb J, Balan P. Attention as a decision in information space. *Trends in Cognitive Sciences* 2010; 14(6):240-248.
13. Vlaev I, Dolan, P. Action change theory: A reinforcement learning perspective on behaviour change. *General Pyschology* 2015; 19:69-95.
14. Seymour B, Singer T, Dolan R. The neurobiology of punishment. *Nat Rev Neurosci* 2007; 8(4):300-311.
15. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 2005; 8(12):1704-1711.
16. Glimcher PW. Choice: towards a standard back-pocket model. *Neuroeconomics: Decision making and the brain* 2009:503-521.
17. Rolls ET. Emotion explained: Oxford University Press, 2005.
18. Dixon ML, Christoff K. The lateral prefrontal cortex and complex value-based learning and decision making. *Neuroscience & Biobehavioral Reviews* (0).
19. Moll J, de Oliveira-Souza R. Moral judgments, emotions and the utilitarian brain. *Trends in Cognitive Sciences* 2007; 11(8):319-321.
20. Raine A, Yang Y. Neural foundations to moral reasoning and antisocial behavior. *Social Cognitive and Affective Neuroscience* 2006; 1(3):203-213.

21. Orihuela-Espina F, Leff DR, James DR, et al. Quality control and assurance in functional near infrared spectroscopy (fNIRS) experimentation. *Phys Med Biol* 2010; 55(13):3701-24.
22. Jobsis FF. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* 1977; 198(4323):1264-7.
23. Roy CS SC. On the regulation of blood-supply to the brain. *J Physiol.* 1890; (1-2): [85]-108:158-7-158-17.
24. Leff DR, Orihuela-Espina F, Atallah L, et al. Functional near infrared spectroscopy in novice and expert surgeons--a manifold embedding approach. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* 2007; 10(Pt 2):270-7.
25. Leff DR, Orihuela-Espina, F., Atallah, L., Athanasiou, T., Leong, J. J., Darzi, A. W., Yang, G. Z. Functional prefrontal reorganization accompanies learning-associated refinements in surgery: a manifold embedding approach. *Comput Aided Surg* 2008; 13(6):325-39.
26. 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.
27. 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.
28. Cope M, Delpy DT, Reynolds EO, et al. Methods of quantitating cerebral near infrared spectroscopy data. *Adv Exp Med Biol* 1988; 222:183-9.
29. Jasper HH. The ten-twenty electrode system of the International Federation. *Electroencephalogr. Clin. Neurophysiol.* 1958.; 10(371–375).
30. Tachtsidis I, Leung TS, Devoto L, et al. Measurement of frontal lobe functional activation and related systemic effects: a near-infrared spectroscopy investigation. *Adv Exp Med Biol* 2008; 614:397-403.
31. Charlin B, Roy L, Brailovsky C, et al. The Script Concordance test: a tool to assess the reflective clinician. *Teach Learn Med* 2000; 12(4):189-95.
32. Graybiel AM. Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* 2008; 31:359-387.
33. de Wit S, Watson P, Harsay HA, et al. Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *The Journal of Neuroscience* 2012; 32(35):12066-12075.
34. Berridge KC, Kringelbach ML. Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology* 2008; 199(3):457-480.
35. Volz KG, Schubotz RI, von Cramon DY. Decision-making and the frontal lobes. *Curr Opin Neurol* 2006; 19(4):401-6.
36. Rosenbloom MH, Schmahmann JD, Price BH. The functional neuroanatomy of decision-making. *J Neuropsychiatry Clin Neurosci* 2012; 24(3):266-77.
37. Krain AL, Wilson AM, Arbuckle R, et al. Distinct neural mechanisms of risk and ambiguity: a meta-analysis of decision-making. *Neuroimage* 2006; 32(1):477-84.

38. Sescousse G, Redoute J, Dreher JC. The architecture of reward value coding in the human orbitofrontal cortex. *J Neurosci* 2010; 30(39):13095-104.
39. Wallis JD. Orbitofrontal Cortex and Its Contribution to Decision-Making. *Annual Review of Neuroscience* 2007; 30(1):31-56.
40. Mulert C, Seifert C, Leicht G, et al. Single-trial coupling of EEG and fMRI reveals the involvement of early anterior cingulate cortex activation in effortful decision making. *NeuroImage* 2008; 42(1):158-168.
41. Esposito F, Mulert C, Goebel R. Combined distributed source and single-trial EEG–fMRI modeling: Application to effortful decision making processes. *NeuroImage* 2009; 47(1):112-121.
42. Gläscher J, Adolphs R, Damasio H, et al. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. *Proceedings of the National Academy of Sciences* 2012; 109(36):14681-14686.
43. Causse M, Peran P, Dehais F, et al. Affective decision making under uncertainty during a plausible aviation task: an fMRI study. *Neuroimage* 2013; 71:19-29.
44. Catena A, Perales JC, Megias A, et al. The brain network of expectancy and uncertainty processing. *PLoS One* 2012; 7(7):e40252.
45. Huettel SA, Song AW, McCarthy G. Decisions under Uncertainty: Probabilistic Context Influences Activation of Prefrontal and Parietal Cortices. *The Journal of Neuroscience* 2005; 25(13):3304-3311.

Figures and Figure Legends

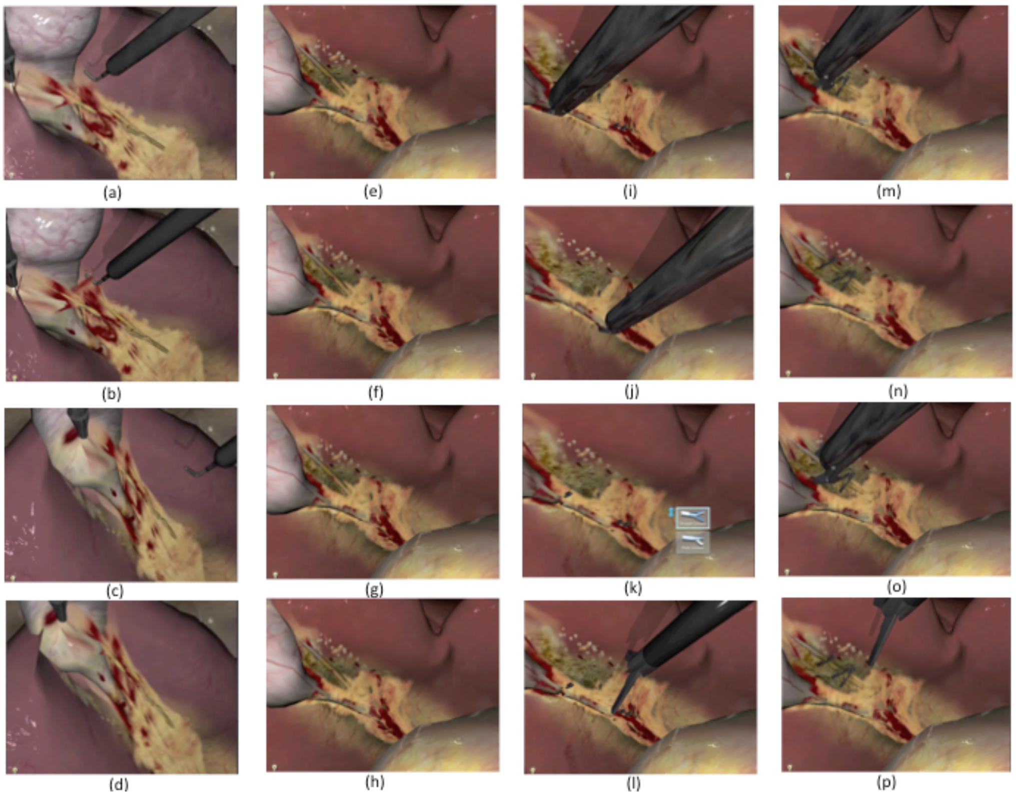
Figure 1



A proposed two-step model of surgeons' intra-operative decision-making, adapted from Flin et al[6] to incorporate a research hypothesis based on intra-operative neuro-monitoring. Surgeons closely monitor the operative scene (a), assess the operative anatomy, and use an appropriate DM strategy (b) to select the next safest operative manoeuvre. The strategy employed depends on available time, perceived risk and operator experience. The hypothesis is that experts employ a recognition-primed approach, whereas novices ruminate options using an analytical DM strategy. Within a neuroimaging framework, surgeons are monitored with multichannel OT such that at each DM phase optical brain data is acquired, and subsequently processed and analysed to determine the loci of greatest response from which the DM system employed can be elucidated (d). Analytical DM evokes dorsolateral prefrontal (DLPFC- operant learning), ventro-lateral prefrontal (VLPFC- prediction errors) and medial prefrontal activations (MPFC -prospect theory and expected utility).

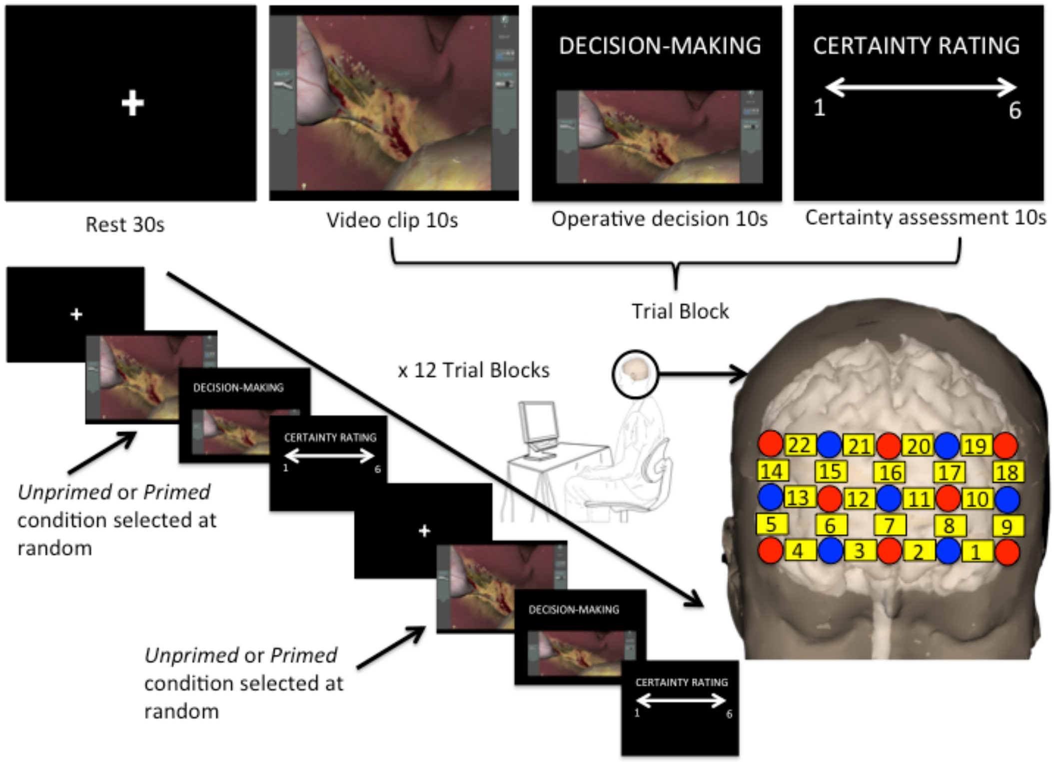
Figure 2 (panels a-b)

(a)



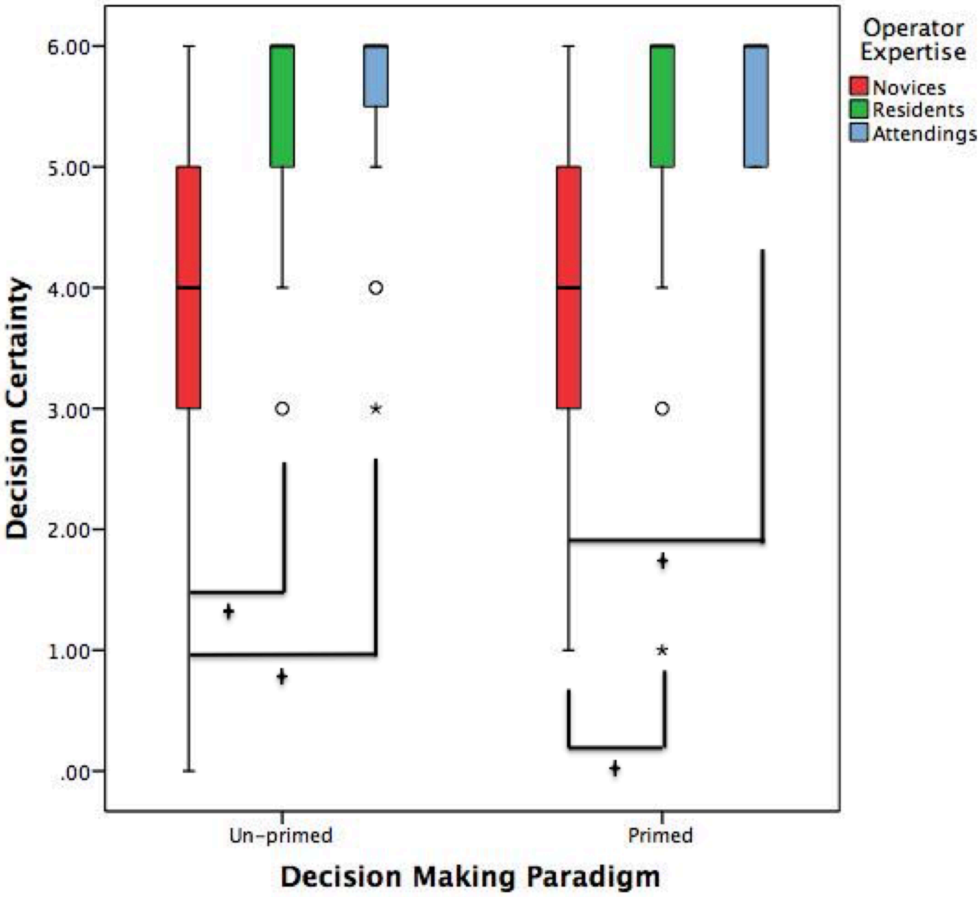
Images depicting different phases of simulated laparoscopic cholecystectomy Videos were classified as either *un-primed* (e.g. a-d and e-h) that terminated at a point where the operator's next manoeuvre was not apparent (d/h) or *primed* (e.g. i-l and n-p) which revealed the operator's intention, e.g. to clip or divide a structure (l/p). Examples of *un-primed* videos include episodes of Calot's triangle dissection (a-d) or gallbladder manipulation without dissection (e-h), following which further dissection would be required in both cases before cystic duct and artery could be safely clipped and ligated. Examples of *primed* videos include sequences of clipping and dividing the cystic duct (i-l) or the cystic artery (n-p). At termination of these primed video sequences, the operator's decision to divide the structure is both clear and incorrect (i.e. clips placed too low down near the common bile duct (i-l), and clipping of the cystic duct should proceed division of the cystic artery (n-p).

(b)



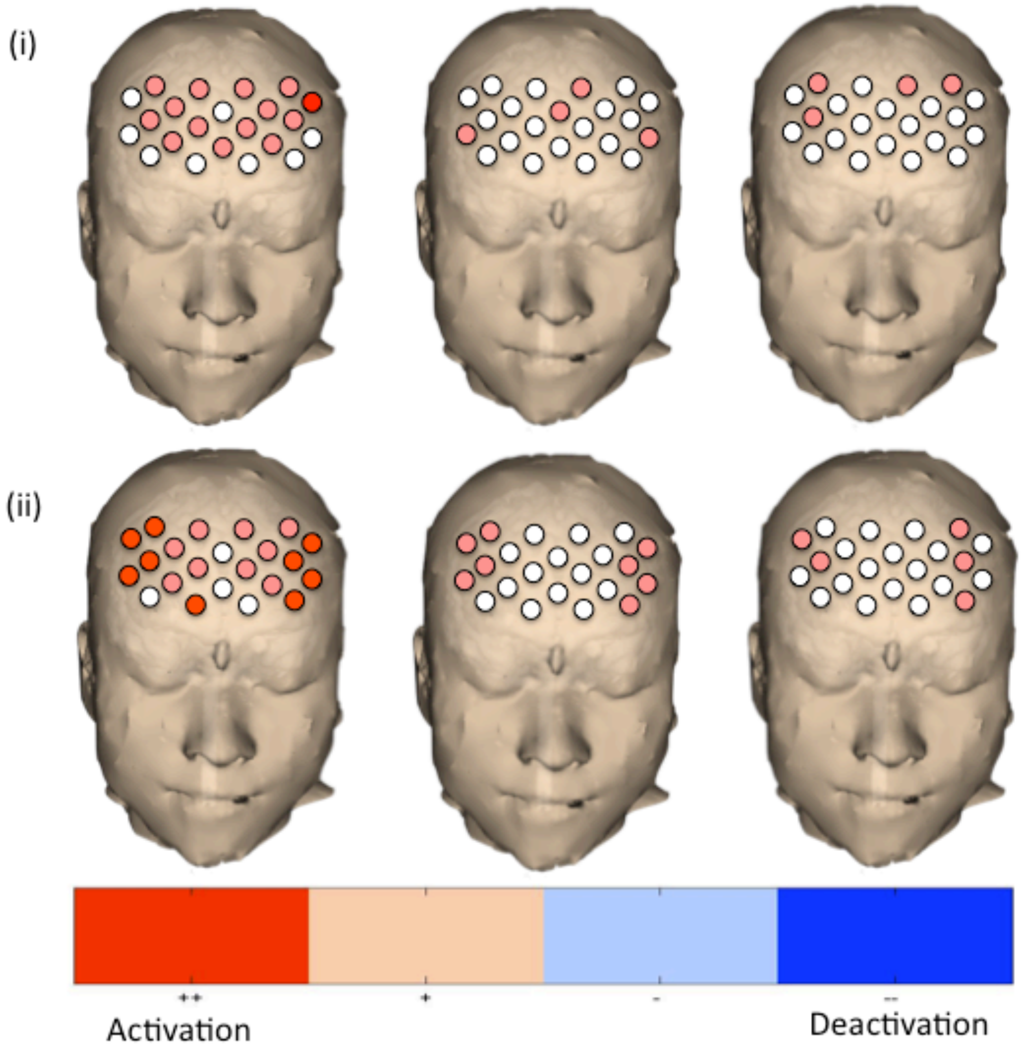
Experimental task set up. Subjects were seated at a table and observed video sequences of simulated laparoscopic cholecystectomy. The experiment was delivered as a block design, with repeated episodes of rest (30s) interspersed with trial blocks that were comprised of three sub-stimuli, namely: video clip review (10s), operative decision-making (10s) and certainty ratings (10s). During rest periods subjects observed the fixation cross, during video review they observed a certain phase of laparoscopic cholecystectomy and during decision making trials they viewed the video's final image and were asked to report the next safest operative maneouvre. Finally, they were asked to report the certainty or confidence in their decision-making. Video clips were classified either *primed* or *un-primed* as to whether the operator's next move was declarative or not. The sequence to which subjects were exposed to these two conditions was random. In total, subjects were exposed to 12 trial blocks whilst multichannel OT monitored changes in cortical haemodynamic change across 22 channels (yellow numbered squares) positioned across the dorsolateral, ventrolateral and medial prefrontal cortex.

Figure 3



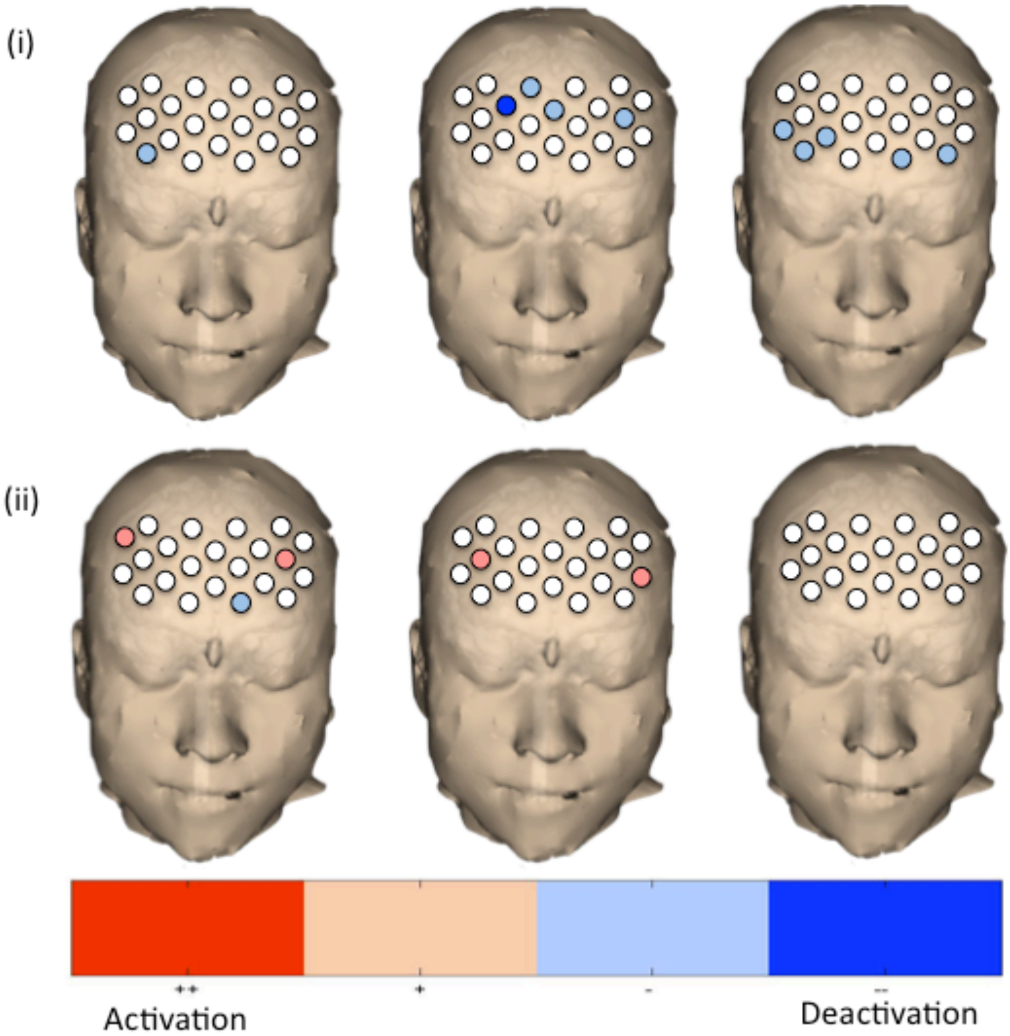
Box plots illustrating decision certainty according to condition (primed / un-primed) and operator expertise. Outliers (o) and extreme values (★) are highlighted. Statistically significant ($p<0.05$) differences in certainty between operators are depicted with a cross (✚).

Figure 4 (panels a –b)
(a)



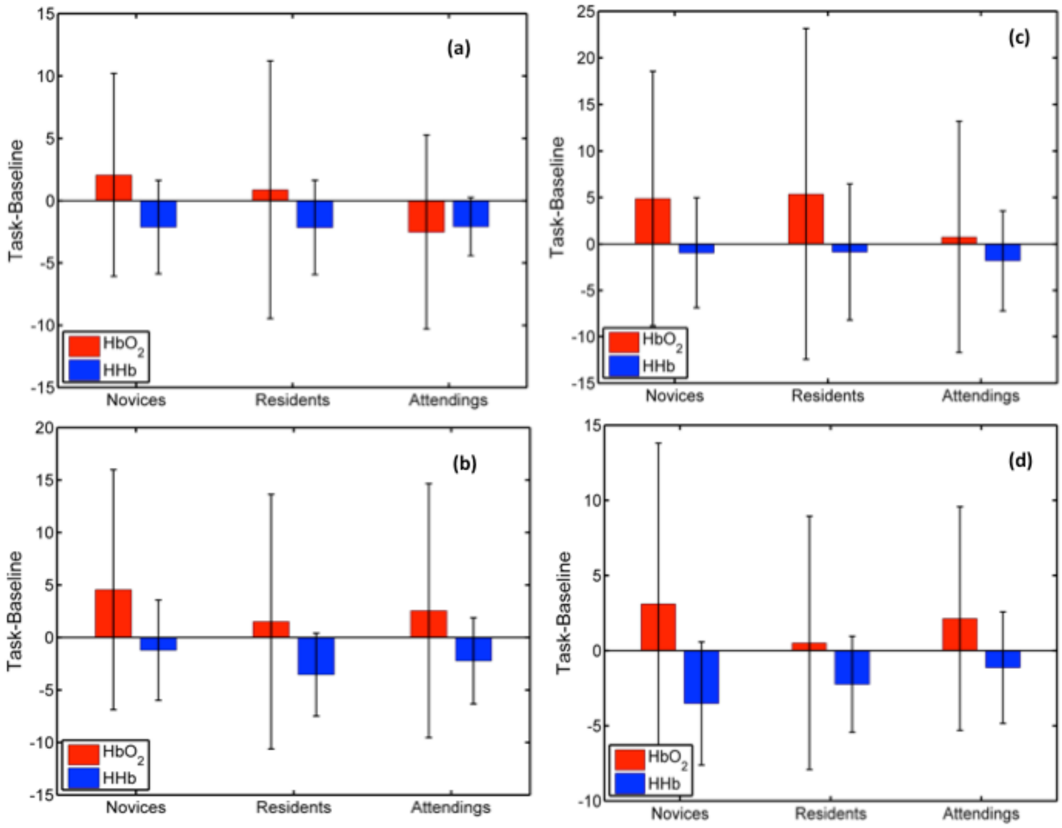
Charts summarise group averaged statistical analysis of HbO_2 and HHb and presented in the form of series of activation / deactivation matrices. Each plot represents an experience group (left column = novices, middle column = residents, right column = attendings) and the ***un-primed*** conditions either video review (i) or decision-making episodes (ii). 22 channels are highlighted (black circles) and colour coded to according to the magnitude of activation [both Hb species reach statistical threshold ($p<0.05$) = red, one Hb species reaching threshold ($p<0.05$) = pink], deactivation [both Hb species reach statistical threshold ($p<0.05$) = light blue, one Hb species reaching threshold ($p<0.05$) = dark blue], or an absence of significant cortical haemodynamic change (white circles).

(b)



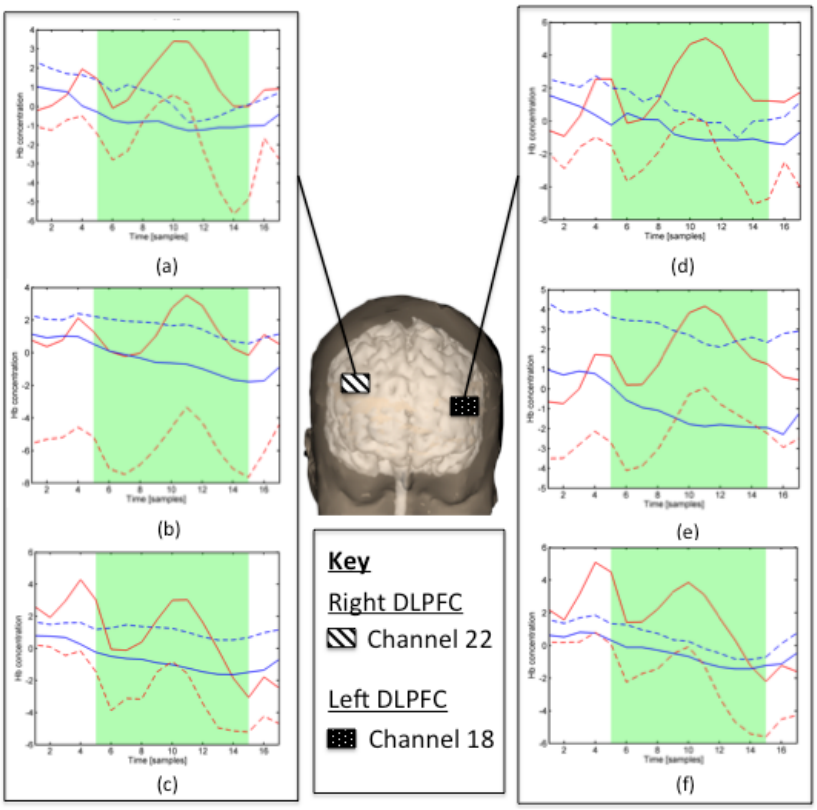
Charts summarise group averaged statistical analysis of HbO_2 and HHb and presented in the form of series of activation / deactivation matrices. Each plot represents an experience group (left column = novices, middle column = residents, right column = attendings) and the ***primed*** conditions either video review (i) or decision-making episodes (ii). 22 channels are highlighted (black circles) and colour coded to according to the magnitude of activation [both Hb species reach statistical threshold ($p<0.05$) = red, one Hb species reaching threshold ($p<0.05$) = pink], deactivation [both Hb species reach statistical threshold ($p<0.05$) = light blue, one Hb species reaching threshold ($p<0.05$) = dark blue], or an absence of significant cortical haemodynamic change (white circles).

Figure 5



Bar charts illustrating between-group differences in mean ΔHbO_2 (red bars) and ΔHHb (blue bars) for certain right dorsolateral prefrontal channels (a=ch22, b=ch5) and left dorsolateral prefrontal channels (c=ch1, d=ch10).

Figure 6



Group averaged Hb time course plots illustrating changes in HbO_2 and HHb during unprimed (bold red and bold blue) and primed (hashed red and hashed blue) DM trials for novices (a and d), residents (b and e) and attendings (c and f). Changes in Hb data during DM trials (green shaded) from rest (white) are shown for a typical channel in the right DLPFC (channel 22) and left DLPFC (channel 18).

Tables

Table 1

Video Clip	Intra-operative Options							
	Further dissection of Calot's triangle	Clip cystic duct	Clip cystic artery	Cut cystic artery	Cut cystic duct	Call a colleague to help	Convert to open	Abandon
1	-	1.000	-	-	-	-	-	-
2	1.000	-	-	-	-	-	-	-
3	0.111	-	-	-	1.000	-	-	-
4	1.000	-	0.667	-	-	-	-	-
5	1.000	0.111	-	-	-	-	-	-
6	-	-	-	1.000	0.429	-	-	-
7	1.000	-	-	0.286	0.143	-	-	-
8	-	1.000	-	-	0.111	-	-	-
9	1.000	-	-	-	-	0.25	-	-
10	1.000	-	-	0.667	-	-	-	-
11	1.000	-	-	-	-	-	-	-
12	1.000	0.167	0.5	-	-	-	-	-

Expert panel decision-making responses to each of the 12 operative videos. Each possible answer is allocated a score or credit according to the number of members on the expert panel who chose that particular answer. For example, video clip 7, from a panel of ten experts, seven chose to dissect Calot's triangle [score= 1(7/7)], two chose to cut the cystic artery [score =0.286 (2/7)], one chose to cut the cystic duct [score=0.143 (1/7)] and none chose options other options [score=0(0/7)]. Thus, the maximum credit for each answer is 1 and the minimum is 0. Each test subject is given a total score based on the total of number of credits gained divided by the total number of questions, subsequently be multiplied by 100 to give a percentage.

Table 2

Expertise	Phase			WSR		
	PRE	INTRA	POST	PRE vs INTRA	INTRA vs POST	PRE vs POST
NOVICES	22.5 (7.5)	19.0 (3.5)	22.5 (3.5)	.191	**.011**	.397
RESIDENTS	20.0 (7.0)	20.0 (7.0)	21.0 (5.0)	.786	.197	.680
ATTENDINGS	22.0 (14.5)	22.0 (8.5)	23.0 (5.5)	.180	.109	.564

Within group analysis of stress and anxiety as measured using the STAI instrument. STAI data are median (IQR) and are provided or each experience group and experimental phase i.e.: prior to (PRE), during (INTRA) and following the experiment (POST). Within group analysis (WSR) was conducted and significant p-values are highlighted in bold.

Table 3

DLPFC Region	DLPFC Channels	ΔHbO_2 Novices	ΔHbO_2 LC experience	MWU (p-value)
Right	22	2.86 (10.44)	-0.69 (10.53)	**0.030**
	14	3.41 (14.89)	0.48 (10.67)	0.181
	13	2.28 (12.18)	0.12 (8.65)	0.161
	5	3.27 (14.76)	1.75 (12.15)	0.152
Left	18	3.33 (12.44)	0.17 (10.52)	0.203
	10	3.66 (12.15)	0.87 (10.23)	0.172
	9	5.61 (15.32)	0.98 (10.14)	0.333
	1	4.49 (19.80)	2.35 (16.84)	0.379
DLPFC Region	**DLPFC Channels**	**ΔHHb Novices**	**ΔHHb LC experience**	**MWU (p-value)**
Right	22	-1.60 (4.33)	-2.59 (3.60)	0.535
	14	-1.96 (5.04)	-2.42 (3.54)	0.514
	13	-2.61 (5.23)	-1.63 (3.01)	**0.036**
	5	-1.52 (4.59)	-2.66 (4.22)	**0.030**
Left	18	-0.58 (2.60)	-0.25 (2.21)	0.701
	10	-3.10 (5.24)	-2.31 (3.91)	**0.020**
	9	-2.06 (5.52)	-2.02 (4.41)	0.752
	1	-0.75 (5.99)	-0.77 (4.86)	0.879

Between-group statistical analysis of ΔHb data for unprimed decision-making stimuli across bilateral DLPFC channels. ΔHb data are median (IQR). Significant p values ($p<0.05$) are highlighted (bold). Relevant pairwise comparisons (Mann Whitney U test) between operators with experience of real laparoscopic cholecystectomy (Attendings and Residents = A+R) and novices (N) are provided.

Table 4

Group	ΔHbO_2		ΔHHb	
	Coefficient	Sig.	Coefficient	Sig.
Novices	-2.995	***.002***	0.787	.075
Residents	-0.825	.485	0.588	.251
Attendings	-1.494	.223	0.840	.076

Results of GLMM analysis demonstrating the effect of priming on the dependent variables - DLPFC grand average ΔHbO_2 and ΔHHb for each experience group. Statistically significant p values are highlighted in bold italic font.

Supplementary Digital Content

Supplementary Table 1

Comparison between operators in the magnitude of decision certainty

Level of reported DM certainty	Proportion of responses		
	Novice	Trainee	Expert
1 – highly uncertain	1%	1%	0%
2	9%	0%	0%
3	23%	2%	2%
4	33%	13%	2%
5	22%	23%	23%
6-highly certain	11%	60%	73%

Supplementary Table 2

Participant demographics and volume of prior laparoscopic cholecystectomy experience.

Expertise	Age	Sex	Handedness	No. of Laparoscopic Cholecystectomies Performed		
				1st surgeon	2nd surgeon	Assistant
Novice	23	F	Right	0	0	0
Novice	21	M	Left	0	0	0
Novice	22	F	Right	0	0	0
Novice	22	F	Right	0	0	2
Novice	22	M	Left	0	0	2
Novice	23	M	Right	0	0	0
Novice	21	M	Right	0	0	0
Novice	24	M	Left	0	0	5
Novice	23	M	Ambidextrous	0	0	0
Novice	23	F	Right	0	0	0
Trainee	32	M	Right	0	2	20
Trainee	34	M	Right	15	100	-
Trainee	31	M	Right	35	80	-
Trainee	34	M	Right	15	50	-
Trainee	32	F	Right	3	70	-
Trainee	33	M	Right	3	30	-
Trainee	29	M	Right	6	80	-
Expert	33	F	Right	120	200	-
Expert	31	F	Right	50	200	-
Expert	37	M	Right	120	70	-
Expert	29	M	Right	85	60	-
Expert	51	M	Right	1000	-	-

Supplementary Appendix I

Laparoscopic Cholecystectomy Post-Training Assessment

1.) How many ports are traditionally used in a laparoscopic cholecystectomy?
2.) What structures make up Calot's triangle?
3.) Which two structures are clipped and divided in order to then be able to dissect the gall bladder of the liver?
4.) Once you have identified Calot's triangle what should you do next?
5.) How many clips do you put on the cystic artery and cystic duct?
6.) Indicate on the image below where would you cut the clipped structures?