Skill acquisition and stress adaptations following laparoscopic surgery training and detraining in novice surgeons

Blair T. Crewther[1] · Kunal Shetty[1,2] · Delaram Jarchi[1] · Shaun Selvadurai[3] · Christian J. Cook[1,4] · Daniel R. Leff[1,2] · Ara Darzi[2] · Guang-Zhong Yang[1]

Abstract

Background Surgical training and practice is stressful, but adaptive changes in the stress circuitry (e.g. perceptual, physiological, hormonal, neural) could support skill development. This work examined skill acquisition and stress adaptations in novice surgeons during laparoscopic surgery (LS) training and detraining.

Methods Twelve medical students were assessed for skill performance after 2 h (BASE), 5 h (MID) and 8 h (POST) of LS training in weeks 1–3, and then after 4 weeks of no training (RETEST). The stress outcomes included state anxiety, perceived stress and workload, heart rate (HR), heart rate variability (HRV), and salivary testosterone and cortisol concentrations. Functional near-infrared spectroscopy was used to assess cortical oxygenation change, as a marker of prefrontal cortex (PFC) activity.

Results Skill performance improved in every session from BASE ($p < 0.01$), with corresponding decreases in state anxiety, stress, workload, low- and high-frequency HRV in the MID, POST and/or RETEST sessions ($p < 0.05$). Left and right PFC were symmetrically activated within each testing session ($p < 0.01$). The stress and workload measures predicted skill performance and changes over time ($p < 0.05$), with state anxiety, mean HR and

the HRV measures also showing some predictive potential ($p < 0.10$).

Conclusions A 3-week LS training programme promoted stress-related adaptations likely to directly, or indirectly, support the acquisition of new surgical skills, and many outcomes were retained after a 4-week period without further LS training. These results have implications for medical training and education (e.g. distributed training for skill development and maintenance, stress resource and management training) and highlighted possible areas for new research (e.g. longitudinal stress and skill profiling).

Keywords Prefrontal cortex · Testosterone · Cortisol · Saliva · Motor learning · Neural

✉ Blair T. Crewther
blair.crewther@gmail.com

1 The Hamlyn Centre for Robotic Surgery, Imperial College, South Kensington Campus, London, UK

2 Department of Surgery and Cancer, Faculty of Medicine, Imperial College, London, UK

3 Faculty of Medicine, Imperial College, London, UK

4 School of Sport, Health and Exercise Science, Bangor University, Bangor, UK

Surgery is a stressful profession with physical, emotional and intellectual demands in both practice and training [1, 2]. Whilst low levels of stress can facilitate surgical performance, excessive stress can lead to impairments in technical and/or non-technical performance, especially for novice surgeons who lack the experience and resources to cope [1]. Indeed, surgeons who cope successfully with stress can perform technically better than those who do not experience stress, or who have failed to cope with it [1]. Thus, for the effective selection, training and management of novice surgeons, which has important implications for patient safety, the stress outcomes of surgical training need to be better understood.

Human stress reactivity is characterised by a central (neural) response linked to other peripheral systems. The prefrontal cortex (PFC) regulates the stress response by integrating cognitive and affective behaviour to modulate other autonomic and endocrine functions [3]. In novice surgeons, PFC activity is high at the naive stage of learning

and attenuates with the continued practice as the skill is performed automatically with time [4], a finding that is consistent with other motor learning studies [5]. Surgical training can also lower perceived (e.g. state anxiety) and physiological stress (e.g. heart rate [HR]), thereby improving skill performance [1]. In novices, these stress outcomes correlate to skill performance [6]. These data confirm that the stress circuitry is highly adaptive to surgical training and that such changes may support, either directly or indirectly, skill development.

Laparoscopic surgery (LS) is one example of a stressful activity. The complexity of instrument manoeuvres and hand–eye coordination during LS task testing or training can promote an array of perceptual (e.g. greater anxiety), physiological (e.g. skin resistance modifications, elevated HR), hormonal (e.g. higher cortisol [C]), and/or neural (e.g. PFC) modifications [4, 6–8]. Studies also indicate that LS produces greater mental strain and stress than conventional [9] and robotic surgery [10]. Therefore, LS would provide a robust model to examine adaptations in the stress circuitry and possible links to skill performance with training in novice surgeons. To our knowledge, this type of research has not been performed. Corresponding information on adaptive changes occurring with LS training cessation (i.e. detraining) is also scarce.

Functional near-infrared spectroscopy (fNIRS) is an established neuroimaging technique for monitoring PFC activity [11], with applications in a surgical domain [4, 12, 13]. It involves the use of fibre optics to measure changes in cerebral cortical blood flow, derived from changes in haemoglobin species such as oxyhaemoglobin (HbO_2) and deoxyhaemoglobin (HHb). A change in regional blood flow is inferred as local neuronal activation, based on the property of neurovascular coupling [14]. This technique offers several advantages over existing neuroimaging methods, in terms of portability, non-invasiveness, relative robustness to motion artefacts and non-restrictiveness, and it permits the use of the ferromagnetic instruments [11, 13]. Thus, fNIRS allows individuals to perform a task using actual surgical instruments within a realistic and ecologically valid environment.

This study examined skill acquisition and stress adaptations in a group of novice surgeons during LS training and detraining. Skill performance was monitored across an intensive training programme along with other perceptual (i.e. anxiety, stress and workload), physiological [i.e. HR, HR variability (HRV)], hormonal (i.e. T, C) and neural (i.e. left and right PFC activity) measures of stress. We hypothesised that LS training would improve skill performance and promote complimentary stress adaptations (e.g. lower anxiety, stress and workload perception, reduced HR and C responses, PFC attenuation), with these adaptations maintained after a short detraining period. We further hypothesised that one or more stress outcomes would be associated with individual changes in skill performance.

Materials and methods

Participants

Thirteen healthy males (mean age 20.1 ± 0.8 years) were recruited for this study, but one dropped out after the first training session. We calculated that ten subjects would be needed to detect a large effect size using a repeated measures design (alpha level = 0.05, power = 0.80). The number of subjects is also consistent with other studies (i.e. 9–15 subjects per group) examining stress physiology and/or skill performance in surgical or medical populations [8, 10, 15–19]. The participants were second- and third-year medical students enrolled at Imperial College, London, with no live or simulated LS experience. Pre-screening involved recruiting a homogenous cohort comprising right-handed young adult male participants with no neuropsychiatric illness and were required to abstain from caffeine or alcohol consumption from the night before testing, given the effects of these substances on cortical responses [13]. Written consent was sought prior to commencement of the study, and ethical approval was provided by the National Research Ethics Service, UK (Reference 10/H0808/124 and 05/Q0403/142).

Training procedures

A single-cohort, experimental design with repeated measures was employed to address the study hypotheses. The participants completed a training programme that focused on intracorporeal laparoscopic suturing and knot-tying in a set sequence. On deconstruction, the sequence was broadly divided into three tasks: first, driving a needle across two edges of a simulated laceration, second, creating and tying a double knot to oppose the two stitched edges, and third, forming and tying a pair of single throw knots to secure the suture in place (http://www.flsprogram.org/). This task is considered one of the most challenging generic surgical manoeuvres, and executing it successfully is a prerequisite to performing complex, minimally invasive surgery [15, 20] and board certification in the USA (http://www.flsprogram.org/). As seen in Table 1, 8 h of distributed training was performed over a 3-week period using an intermittent format: week 1 = 1 × 2 h block, week 2 = 3 × 1 h blocks and week 3 = 3 × 1 h blocks. All training and testing was performed in a standard laparoscopic box trainer (iSim, iSurgicals, UK) under supervision from a surgeon with advanced LS training.

Table 1 Laparoscopic skills training and testing schedule

Week 1		Week 2		Week 3		Week 4–7	Week 8
Training 1 × 2 h block	Testing	Training 3 × 1 h blocks	Testing	Training 3 × 1 h blocks	Testing	Training nil	Testing
	BASE		MID		POST		RETEST

Testing procedures

Testing was conducted immediately after 2 h of training (BASE) in week 1, after 5 training hours (MID) in week 2, and after 8 h of training (POST) in week 3 (Table 1). A final session was performed in week 8 after 4 weeks of no training (RETEST). Each session was completed on a different day from training, except the BASE session, so participants rested for 30 min in a quiet room afterwards to ensure homeostatic restoration of the stress parameters. Each session began with the assessment of state anxiety, and a saliva sample was taken before subjects were equipped with a chest belt sensor (Zephyr Bioharness 2.0, NZ) and fNIRS head sensor. Next, the LS tasks outlined above were repeated using a block design involving 3 × 3 repetitions with 30-s rest before commencement of the first task (at baseline) and then 40-s rest between each subtask. After task completion, perceived stress and workload were assessed and another saliva sample was taken. Task completion ranged from 10 to 25 min with faster times recorded across each successive session. To account for diurnal variation, the sessions for each participant were scheduled to start at a similar time of day (5 pm ± 2 h). The testing sessions were conducted in a temperature-regulated room in the presence of two researchers (BC and KS).

Skill performance

For each suturing task, a performance score was calculated using a validated formula [20] adopted by the fundamentals of LS curriculum (http://www.flsprogram.org/). The score accounted for efficiency (time) and precision (errors). Precision was measured by the quality of the knot, opposition of sutured edges, accuracy of the entry and exit of the needle. The results were averaged across the 3 tasks (and 3 blocks) per session to derive a mean skill performance score for each participant.

Perceptual assessments

Pre-session state anxiety was measured using the 6-item State-Trait Anxiety Inventory (STAI) [21]. Post-session stress was assessed using a single-item score by asking the participants, "How stressful was the overall assessment?"

ranging from 1 (not stressful at all) to 5 (very stressful). Session workload was quantified using the NASA task load index [22]. Each task was assessed on 6 subscales (i.e. mental demands, physical demands, time demands, own performance, effort and frustration) on a 20-point scale, and paired comparisons were used to calculate a weighted score. The task scores (rating × weighted score) were aggregated to derive a pooled measure of session workload.

Physiological assessments

Electrocardiographic (ECG) data from the chest sensor were collected at 100 Hz and transmitted wirelessly to a computer for analysis using customised MATLAB software [23]. The key measures were mean and maximum HR, from which we calculated delta HR changes (maximum–mean), along with HRV in the high-frequency (HF 0.15–0.4 Hz) and low-frequency (LF 0.04–0.15 Hz) spectrums [9]. A LF/HF ratio was also calculated from the individual LF and HF values. The HR and HRV values were derived from all data collected across each session.

Hormonal assessments

Saliva provides a valid and non-invasive alternative to blood collections for steroid determination [24]. Saliva samples (1 ml) were collected by passive drool 5 min before and 10 min after each session and stored at −80 °C until assay. The samples were analysed in duplicate for T and C concentrations using commercial enzyme immunoassay kits (Salimetrics LLC, USA). The detection limits were 6.1 pg/ml for the T assay and 0.12 ng/ml for the C assay, with inter-assay variability (based on low and high controls) of 12.4 and 7.6 %, respectively. The samples from each participant were tested within the same assay to eliminate inter-assay variance. Pre- and post-session T and C concentrations were analysed, along with the delta T and C changes (post–pre session).

Assessment of the prefrontal cortex

Cerebral activity was monitored using a 44-channel fNIRS system (Hitachi ETG 4000, Japan), of which 22 channels were positioned on the scalp overlying the PFC to record

haemodynamics at 8 Hz. Attenuation of light data was converted to relative changes in HbO_2 and HHb (vs. pre-session values) using the modified Beer–Lambert law and decimated to 1 Hz and linearly detrended prior to undergoing data integrity checks [13]. The optode data were pooled across all LS tasks to derive a single-session value for left and right PFC activity. We included another measure of PFC asymmetry, a laterality index (LI), which was calculated from the individual optode data ((right–left)/(right + left)) [25] and pooled, as described above. Although HbO_2 and HHb were both assessed in this work, we focused on changes in HbO_2 levels to quantify stress reactivity and adaptation.

Statistical analyses

Within-session changes in the HR, hormonal and PFC variables were examined using paired T tests. Between-session differences in all variables were tested using a generalised estimation equation (GEE) model with an autoregressive correlational structure [26]. Main effects and interactions were determined by significance testing of the Wald Chi-square statistic (χ^2) with testing week as a covariate. Where appropriate, the LSD test was used as the post hoc procedure. Relationships between the performance outcomes (raw scores and percentage change) in each session were assessed using Pearson correlations. Within-individual testing was used to assess the relationship between the stress (predictor) and skill performance (predicted) variables as a pooled data set, based on individual slope patterns and T test analysis of the group mean from zero. Data were analysed using SPSS version 21.0 for Windows (IBM Corp., USA). Significance was set at an alpha level of $p \leq 0.05$.

Results

To analyse the HR, T and C variables, the GEE model included session, sample and the session × sample interaction as categorical factors. This procedure was repeated to examine PFC activity, with hemisphere (left, right) replacing the sample category. To compare the between-session changes in HR, T, C and PFC activity, along with the remaining variables (i.e. skill performance, STAI, stress, workload, LF, HF, LF/HF, LI), the GEE model included session as a single categorical factor. Analysis of skill performance (Fig. 1) revealed a session effect, χ^2 (3) = 120.9, $p < 0.001$, with performance improving from BASE to MID and from MID to the POST and RETEST sessions ($p < 0.001$).

As seen in Table 2, the STAI scores differed by testing session, χ^2 (3) = 27.2, $p < 0.001$, being lower at POST
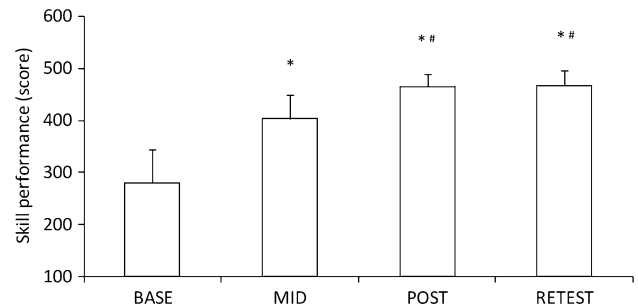
Fig. 1 Estimated marginal means (±SD) for skill performance in each testing session. *Significant from BASE $p < 0.05$, #significant from MID $p < 0.05$

than MID ($p < 0.001$). The changes in stress were also significant, χ^2 (3) = 18.3, $p < 0.001$, with lower scores in all sessions from BASE and a lower POST result from MID ($p < 0.029$). Similar changes in workload were identified, χ^2 (3) = 11.6, $p = 0.009$, with lower POST and RETEST scores than the BASE and MID sessions ($p < 0.018$). We found a main session effect on the HR data, χ^2 (3) = 8.46, $p = 0.037$, with lower MID and POST session HR values than BASE. Maximum HR was higher than mean HR across all sessions, χ^2 (1) = 70.6, $p < 0.001$, matching the within-session HR changes ($p < 0.001$), but the HR responses were stable across each session, χ^2 (3) = 2.79, $p = 0.426$.

Session effects on HF, χ^2 (3) = 16.4, $p = 0.001$, and LF, χ^2 (3) = 13.3, $p = 0.004$, HRV were observed with these outcomes decreasing in the MID, POST and/or RETEST session from BASE ($p < 0.009$). The LF/HF ratio was stable across all sessions, χ^2 (3) = 6.58, $p = 0.087$. Testosterone analysis revealed no session (χ^2 (3) = 1.71, $p = 0.635$), sample (χ^2 (1) = 0.503, $p = 0.478$) or interactive effects (χ^2 (3) = 6.27, $p = 0.099$). Cortisol testing also revealed no effect by session (χ^2 (3) = 4.19, $p = 0.241$), sample (χ^2 (1) = 0.014, $p = 0.907$) or any interaction (χ^2 (3) = 2.16, $p = 0.540$). Subsequently, we found no significant within-session T and C changes and no significant between-session differences in these responses.

Left and right PFC activity were elevated in all sessions ($p < 0.035$, Fig. 2A). Further testing revealed a main session effect on PFC activity, χ^2 (3) = 8.13, $p = 0.043$, but post hoc analysis revealed no significant between-session differences. There was no main hemisphere effect on PFC activity (χ^2 (1) = 0.138, $p = 0.710$) and no session × hemisphere interaction (χ^2 (3) = 3.41, $p = 0.332$). The LI did exhibit a session effect, χ^2 (3) = 8.29, $p = 0.040$, but no significant differences were seen post hoc (Fig. 2B).

Correlational testing revealed that BASE skill performance was related to MID performance ($r = 0.60$, $p = 0.038$) and to the percentage changes in skill

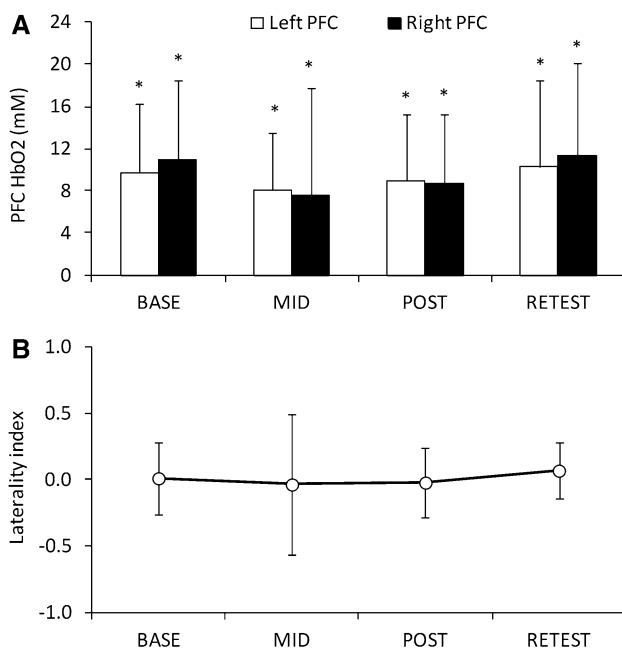**Table 2** Estimated marginal means (±SD) for the stress outcomes in each testing session

	BASE	MID	POST	RETEST
STAI (score)	10.0 ± 2.42	10.0 ± 2.31	8.92 ± 2.18[#]	8.83 ± 2.07
Stress (score)	3.75 ± 1.01	3.00 ± 1.00*	2.08 ± 1.26*[#]	2.75 ± 1.23*
Workload (score)	59.6 ± 6.74	53.4 ± 10.4	43.1 ± 15.8*[#]	43.5 ± 20.1*[#]
HR (bpm)				
Mean	91.0 ± 20.0	87.0 ± 17.4	85.3 ± 19.0	93.0 ± 18.2
Maximum	131 ± 27.5	122 ± 23.9	116 ± 28.4	124 ± 25.4
Change	40.1 ± 19.6[†]	35.0 ± 16.8[†]	30.2 ± 16.9[†]	31.3 ± 21.7[†]
HRV				
HF (ms^2)	0.97 ± 0.35	0.75 ± 0.30*	0.65 ± 0.16*	0.57 ± 0.19*
LF (ms^2)	0.97 ± 0.28	0.86 ± 0.29	0.78 ± 0.17*	0.73 ± 0.23*
LF/HF	1.07 ± 0.29	1.20 ± 0.18	1.22 ± 0.19	1.30 ± 0.15
Testosterone (pg/ml)				
Pre	133 ± 50.7	132 ± 40.9	138 ± 39.5	118 ± 44.7
Post	133 ± 52.8	141 ± 43.5	132 ± 30.3	127 ± 45.9
Change	−0.71 ± 21.6	9.72 ± 20.4	−5.82 ± 28.4	9.35 ± 20.0
Cortisol (ng/ml)				
Pre	1.26 ± 0.63	1.66 ± 0.81	1.53 ± 0.74	1.37 ± 0.38
Post	1.53 ± 0.94	1.56 ± 0.82	1.48 ± 1.30	1.31 ± 0.66
Change	0.27 ± 0.82	−0.10 ± 0.42	−0.05 ± 1.06	−0.06 ± 0.54

STAI State-Trait Anxiety Inventory, *HR* heart rate, *HRV* heart rate variability, *HF* high frequency, *LF* low frequency, *LF/HF* low frequency-to-high frequency ratio

* Significant from BASE $p < 0.05$

\# Significant from MID $p < 0.05$

† Significant within-session change $p < 0.05$

Fig. 2 Estimated marginal means (±SD) for left and right prefrontal cortex (PFC) activity (**A**) and the laterality index (**B**) in each testing session. *Significant within-session change $p < 0.05$

performance across all subsequent sessions ($r > −0.83$, $p < 0.009$). Slope testing identified several predictors of individual performance (Table 3). The stress and workload measures were negatively related to absolute skill performance ($p < 0.015$) and to the percentage changes over time ($p < 0.058$), whilst the STAI, mean HR, HF and LF measures were also negatively related to the performance outcomes at a higher significance threshold ($p < 0.10$).

Discussion

The study was undertaken to examine changes in technical skills and stress adaptations following LS training and detraining in novice surgeons. Intensive LS training enhanced skill performance and reduced session anxiety, stress and workload, in addition to lowering HRV (HF and LF), but without apparent changes in hormone activity. The adaptive changes remained after a short period with no task practice. The left and right PFC were similarly activated during testing, but there were no session differences in the neural outcomes.

One of the hallmarks of learning new surgical skills, as we identified, is a large initial performance gain followed

Table 3 Mean slopes (±SD) between the stress and skill performance outcomes across all testing sessions

Predictor	Skill performance (score)		Skill performance (% change)	
	Slope	p value	Slope	p value
STAI (score)	-97 ± 185	0.097	-50 ± 95	0.097
Stress (score)	-114 ± 120	0.007	-37 ± 44	0.015
Workload (score)	-116 ± 137	0.014	-53 ± 86	0.057
HR mean (bpm)	-85 ± 173	0.116	-36 ± 66	0.089
HR max (bpm)	-57 ± 174	0.283	-20 ± 65	0.315
HR change (bpm)	119 ± 426	0.352	43 ± 155	0.358
HF HRV (ms^2)	-104 ± 177	0.066	-44 ± 72	0.058
LF HRV (ms^2)	-66 ± 116	0.073	-24 ± 41	0.074
LF/HF	31 ± 224	0.639	22 ± 104	0.487
Testosterone pre (pg/ml)	-16 ± 166	0.740	-18 ± 76	0.433
Testosterone post (pg/ml)	-4.5 ± 187	0.935	-21 ± 104	0.507
Testosterone change (pg/ml)	9.4 ± 102	0.754	10 ± 52	0.521
Cortisol pre (ng/ml)	-46 ± 182	0.404	-38 ± 136	0.360
Cortisol post (ng/ml)	141 ± 435	0.285	62 ± 181	0.261
Cortisol change (ng/ml)	-38 ± 136	0.358	-1.4 ± 74	0.948
Left PFC (mM)	13 ± 284	0.874	27 ± 158	0.565
Right PFC (mM)	-37 ± 141	0.377	-25 ± 71	0.250
LI	35 ± 299	0.695	26 ± 140	0.532

STAI State-Trait Anxiety Inventory, *HR* heart rate, *HF* high frequency, *LF* low frequency, *HRV* heart rate variability, *LF/HF* low frequency-to-high frequency ratio, *PFC* prefrontal cortex, *LI* laterality index

by a period of stabilisation where skill refinement occurs [16, 20, 27, 28]. Proficiency in different LS tasks can be accomplished with 50–70 repeated trials (on average) [16, 20, 29, 30] and achieved in 1–2 training sessions [16]. In this work, the participants completed 8 h of training (i.e. 108 trials over 7 distributed training sessions). We also found that skill acquisition was maintained at a higher proficiency level after a 4-week detraining period. Other work identified a loss in LS skills after 2 weeks without training [28], but a shorter period of training and practice times were performed. Distributed (intermittent) programmes are thought to be superior to massed programmes in that they aid learning by allowing time to consolidate cognitively between practice sessions [28]. Our results confirm that intermittent training facilitates improvement and stabilisation of technical skills. Further research is needed to compare the two training approaches.

The corresponding reduction in perceived anxiety, stress levels and workload with LS training is consistent with the literature [1, 29], with the added novelty of parameter retention 4 weeks after training cessation. The corresponding decline in HF and LF power indicates a reduction in the autonomic inputs to cardiovascular stress reactivity [9]. These adaptations could support skill development by increasing stress-coping resources and reducing threat perception, thereby optimising attentional demands and subsequent learning. As supporting evidence, experienced surgeons often exhibit reduced stress responses and/or better surgical performance than lesser trained or novice surgeons [1, 6, 17]. In addition to skills training, the adaptive changes observed herein could be explained by habituation to the testing conditions, habituation to stress, and/or increased confidence in performance.

The testing of LS skill performance was also deemed to be stressful, as indicated by the large HR and PFC responses within each session and the subjective ratings of stress. Conversely, we found no within-session changes in the T or C biomarkers and no hormonal differences over time. Different surgical scenarios can consistently induce elevated perceptual and physiological stress outcomes, whereas hormonal parameters often vary in their responsiveness to these challenges [6, 8, 18, 19]. These findings could be attributed to individual differences in stress perception and subsequent hormone reactivity [31], or simply differences in baseline hormone concentrations [8]. We do recognise that the central and autonomic stress systems respond more quickly to stress stimuli than the hormonal systems [3], which highlights the importance of the sampling procedures to capture these events.

The left and right PFC were engaged during testing, but we found no long-term PFC changes that indicate a need for executive control over time. There are reports of attenuated PFC activity when new surgical [4] and motor skills are learned [32–35], so as the skill is performed

automatically, areas of the brain that 'scaffold' novel demands become redundant. The greater complexity of the LS tasks, than earlier employed motor tasks, could explain the consistent PFC response observed. Despite evidence also favouring the right PFC when learning [34], the left and right PFC were symmetrically activated across this study. The PFC represents only one area that subserves the brain's attentional system [34], and reduced cognitive demands may be indexed from patterns of recruitment in other areas (i.e. functional redistribution) [5]. Additional PFC roles in motor behaviour (e.g. executive function, judgement, working memory) might also explain the sustained global response and lack of PFC lateralisation despite longitudinal improvements in skills and stress adaptation.

Cross-sectional data have identified various physiological and perceptual correlates of skill performance in novice surgeons [6]. Similarly, we found that the stress and workload measures were predictive (negatively related) of individual changes in skill performance, with other measures (i.e. STAI, HR mean, HRV) showing some predictive potential at a higher significance threshold. In addition, the skill performance scores in the BASE session were positively related to the MID session performance scores, similar to other reports in trainee surgeons [27], and negatively related to the percentage changes in skill performance across all subsequent sessions. These findings add to the growing evidence of the predictive psychomotor ability on surgical performance [36]. Further longitudinal studies are needed to reliably profile the stress systems, in terms of their engagement during surgical training and assessment, along with moderating factors (e.g. coping styles, control, social support) [2], and ultimately, what it means for actual skill performance and development.

Overall, it appears that repeated exposure to a learning task can promote positive stress adaptations that are directly, or indirectly, related to skill improvements in novice surgeons. Thus, trainee surgeons could benefit from structured training to increase their perceived stress resources and/or stress management abilities [1, 37]. Introducing stressful assessments (e.g. simulations with social evaluation) at appropriate intervals might also reduce threat perception and the stress response, thereby aiding skill performance [30, 38]. Our data further highlighted large individual differences when performing a relatively unfamiliar technical skill, which might impact upon learning rates and training capacity for skill acquisition. Individual reductions in certain markers of stress also facilitated improvements in technical performance. A robust assessment of these outcomes may therefore provide a useful adjunct to selection procedures and guidance for a surgical career [1].

Data interpretation is limited by differences in the timing and frequency of the stress outcomes. Some measures were also pooled to facilitate session comparisons (e.g. task performance, workload, HR), preventing us from examining the individual tasks (with repeats) performed. Despite the small sample size ($n = 12$), the participants each completed 8 training sessions ($n = 96$) and 4 testing sessions ($n = 48$); therefore, a total of 144 separate trials were conducted. The lack of a control group and BASE testing immediately after 2 h of training are other limitations, but unavoidable given the nature of this study and practical constraints. Finally, we focused only on the technical aspect of performance and in a simulated environment that did not consider other factors (e.g. situation awareness, decision-making, leadership, human interactions) in the operating theatre that may influence surgeon stress levels and the training response. Still, examining these tasks in isolation allowed us to disentangle the results from the myriad of factors that come into play.

In summary, a 3-week LS training programme promoted stress adaptations likely to support the acquisition of new performance skills in novice surgeons. Moreover, many of the observed outcomes were retained after a 4-week period without further LS training. Some stress measures were also related to the individual changes in skill performance, thereby indicating a more direct link between the stress circuitry and skill development.

Compliance with ethical standards

Disclosures Drs. Crewther, Shetty, Jarchi, Cook and Leff, along with Mr. Selvadurai and Professors Darzi and Yang, have no conflicts of interest or financial ties to disclose.

References

1. Arora S, Sevdalis N, Nestel D, Woloshynowych M, Darzi A, Kneebone R (2010) The impact of stress on surgical performance: a systematic review of the literature. Surgery 147:318–330
2. Leblanc VR (2009) The effects of acute stress on performance: implications for health professions education. Acad Med 84:S25–S33
3. Cerqueira JJ, Almeida OF, Sousa N (2008) The stressed prefrontal cortex. Left? Right? Brain Behav Immun 22:630–638
4. Leff DR, Orihuela-Espina F, Atallah L, Athanasiou T, Leong JJ, Darzi AW, Yang GZ (2008) Functional prefrontal reorganization accompanies learning-associated refinements in surgery: a manifold embedding approach. Comput Aided Surg 13:325–339

5. Kelly AMC, Garavan H (2005) Human functional neuroimaging of brain changes associated with practice. Cereb Cortex 15:1089–1102

6. Arora S, Sevdalis N, Aggarwal R, Sirimanna P, Darzi A, Kneebone R (2010) Stress impairs psychomotor performance in novice laparoscopic surgeons. Surg Endosc 24:2588–2593

7. Schuetz M, Gockel I, Beardi J, Hakman P, Dunschede F, Moenk S, Heinrichs W, Junginger T (2008) Three different types of surgeon-specific stress reactions identified by laparoscopic simulation in a virtual scenario. Surg Endosc 22:1263–1267

8. Arora S, Aggarwal R, Moran A, Sirimanna P, Crochet P, Darzi A, Kneebone R, Sevdalis N (2011) Mental practice: effective stress management training for novice surgeons. J Am Coll Surg 212:225–233

9. Böhm B, Rötting N, Schwenk W, Grebe S, Mansmann U (2001) A prospective randomized trial on heart rate variability of the surgical team during laparoscopic and conventional sigmoid resection. Arch Surg 136:305–310

10. Klein MI, Warm JS, Riley MA, Matthews G, Doarn C, Donovan JF, Gaitonde K (2012) Mental workload and stress perceived by novice operators in the laparoscopic and robotic minimally invasive surgical interfaces. J Endourol 26:1089–1094

11. Leff DR, Orihuela-Espina F, Elwell CE, Athanasiou T, Delpy DT, Darzi AW, Yang GZ (2011) Assessment of the cerebral cortex during motor task behaviours in adults: a systematic review of functional near infrared spectroscopy (fNIRS) studies. Neuroimage 54:2922–2936

12. Ohuchida K, Kenmotsu H, Yamamoto A, Sawada K, Hayami T, Morooka K, Takasugi S, Konishi K, Ieiri S, Tanoue K, Iwamoto Y, Tanaka M, Hashizume M (2009) The frontal cortex is activated during learning of endoscopic procedures. Surg Endosc 23:2296–2301

13. Orihuela-Espina F, Leff DR, James DRC, Darzi AW, Yang GZ (2010) Quality control and assurance in functional near infrared spectroscopy (fNIRS) experimentation. Phys Med Biol 55:1–24

14. Cope M, Delpy DT (1988) System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. Med Biol Eng Comput 26:289–294

15. Yamaguchi S, Yoshida D, Kenmotsu H, Yasunaga T, Konishi K, Ieiri S, Nakashima H, Tanoue K, Hashizume M (2011) Objective assessment of laparoscopic suturing skills using a motion-tracking system. Surg Endosc 25:771–775

16. Malhotra N, Poolton JM, Wilson MR, Ngo K, Masters RSW (2012) Conscious monitoring and control (reinvestment) in surgical performance under pressure. Surg Endosc 26:2423–2429

17. Yamamoto A, Hara T, Kikuchi K, Hara T, Fujiwara T (1999) Intraoperative stress experienced by surgeons and assistants. Ophthalmic Surg Lasers 30:27–30

18. Arora S, Tierney T, Sevdalis N, Aggarwal R, Nestel D, Woloshynowych M, Darzi A, Kneebone R (2010) The Imperial Stress Assessment Tool (ISAT): a feasible, reliable and valid approach to measuring stress in the operating room. World J Surg 34:1756–1763

19. Alobid I, de Pablo J, Mullol J, Centellas S, Parramon G, Carrasco J, Armario A, Bernal-Sprekelsen M (2011) Increased cardiovascular and anxiety outcomes but not endocrine biomarkers of stress during performance of endoscopic sinus surgery. Arch Otolaryngol Head Neck Surg 137:487–492

20. Korndorffer JR Jr, Dunne JB, Sierra R, Stefanidis D, Touchard CL, Scott DJ (2005) Simulator training for laparoscopic suturing using performance goals translates to the operating room. J Am Coll Surg 201:23–29

21. Marteau T, Bekker H (1992) The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). Br J Clin Psychol 31:301–306

22. Hart SG (2006) NASA-Task Load Index (NASA-TLX); 20 years later. In: Human factors and ergonomics society annual meeting proceedings, San Francisco, pp 904–908

23. Orihuela-Espina F, Leff DR, James DR, Darzi AW, Yang GZ (2009) ICNA: a software tool for manifold embedded based analysis of functional near infrared spectroscopy data. In: 15th annual meeting of the organization for human brain mapping, San Francisco, California, USA

24. Gröschl M (2008) Current status of salivary hormone analysis. Clin Chem 54:1759–1769

25. Ishikawa W, Sato M, Fukuda Y, Matsumoto T, Takemura N, Sakatani K (2014) Correlation between asymmetry of spontaneous oscillation of hemodynamic changes in the prefrontal cortex and anxiety levels: a near-infrared spectroscopy study. J Biomed Opt 19:027005. doi:10.1117/1111.JBO.1119.1112

26. Liang K-Y, Zeger SL (1986) Longitudinal data analysis using generalized linear models. Biometrika 73:13–22

27. Chaudhry A, Sutton C, Wood J, Stone R, McClow R (1999) Learning rate for laparoscopic surgical skills on MIST VR, a virtual reality simulator: quality of human–computer interface. Ann R Coll Surg Engl 81:281–286

28. Gallagher AG, Jordan-Black JA, O'Sullivan GC (2012) Prospective, randomized assessment of the acquisition, maintenance, and loss of laparoscopic skills. Ann Surg 256:387–393

29. Prabhu A, Smith W, Yurko Y, Acker C, Stefanidis D (2010) Increased stress levels may explain the incomplete transfer of simulator-acquired skill to the operating room. Surgery 147:640–645

30. Vine SJ, Freeman P, Moore LJ, Chandra-Ramanan R, Wilson MR (2013) Evaluating stress as a challenge is associated with superior attentional control and motor skill performance: testing the predictions of the biopsychosocial model of challenge and threat. J Exp Psychol Appl 19:185–194

31. Bouarfa L, Bembnowicz P, Crewther B, Jarchi D, Yang G-Z (2013) Profiling visual and verbal stress responses using electrodermal, heart rate and hormonal measures. In: IEEE body sensor network conference, MIT, Cambridge, Massachusetts pp 1–7

32. James DR, Leff DR, Orihuela-Espina F, Kwok KW, Mylonas GP, Athanasiou T, Darzi AW, Yang GZ (2013) Enhanced frontoparietal network architectures following "gaze-contingent" versus "free-hand" motor learning. Neuroimage 64:267–276

33. Rémy F, Wenderoth N, Lipkens K, Swinnen SP (2008) Acquisition of a new bimanual coordination pattern modulates the cerebral activations elicited by an intrinsic pattern: an fMRI study. Cortex 44:482–493

34. Debaere F, Wenderoth N, Sunaert S, Van Hecke P, Swinnen SP (2004) Changes in brain activation during the acquisition of a new bimanual coordination task. Neuropsychologia 42:855–867

35. Petersen SE, van Mier H, Fiez JA, Raichle ME (1998) The effects of practice on the functional anatomy of task performance. Proc Natl Acad Sci USA 95:853–860

36. Maan ZN, Maan IN, Darzi AW, Aggarwal R (2012) Systematic review of predictors of surgical performance. Br J Surg 99:1610–1621

37. Wetzel CM, George A, Hanna GB, Athanasiou T, Black SA, Kneebone RL, Nestel D, Woloshynowych M (2011) Stress management training for surgeons—a randomized, controlled, intervention study. Ann Surg 253:488–494

38. Moore LJ, Vine SJ, Wilson MR, Freeman P (2014) Examining the antecedents of challenge and threat states: the influence of perceived required effort and support availability. Int J Psychophysiol 93:267–273