



OPEN ACCESS

EDITED BY

Elena De Momi,
Politecnico di Milano, Italy

REVIEWED BY

Yuanyuan Gao,
Boston University, United States
Arun Nemani,
Draycon Labs, United States

*CORRESPONDENCE

Christopher D'Ambrosia
✉ cdambros@eng.ucsd.edu

SPECIALTY SECTION

This article was submitted to
Clinical Neuroergonomics,
a section of the journal
Frontiers in Neuroergonomics

RECEIVED 23 September 2022

ACCEPTED 19 December 2022

PUBLISHED 09 January 2023

CITATION

D'Ambrosia C, Aronoff-Spencer E,
Huang EY, Goldhaber NH,
Christensen HI, Broderick RC and
Appelbaum LG (2023) The
neurophysiology of intraoperative
error: An EEG study of trainee
surgeons during robotic-assisted
surgery simulations.
Front. Neuroergon. 3:1052411.
doi: 10.3389/fnrgo.2022.1052411

COPYRIGHT

© 2023 D'Ambrosia, Aronoff-Spencer,
Huang, Goldhaber, Christensen,
Broderick and Appelbaum. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

The neurophysiology of intraoperative error: An EEG study of trainee surgeons during robotic-assisted surgery simulations

Christopher D'Ambrosia^{1,2*}, Eliah Aronoff-Spencer³,
Estella Y. Huang⁴, Nicole H. Goldhaber⁴, Henrik I. Christensen²,
Ryan C. Broderick⁴ and Lawrence G. Appelbaum⁵

¹College of Physicians and Surgeons, Columbia University, New York, NY, United States, ²Cognitive Robotics Laboratory, Department of Computer Science and Engineering, Contextual Robotics Institute, University of California, San Diego, La Jolla, CA, United States, ³Department of Medicine, University of California, San Diego, La Jolla, CA, United States, ⁴Division of Minimally Invasive Surgery, Department of Surgery, University of California, San Diego, La Jolla, CA, United States, ⁵Department of Psychiatry, University of California, San Diego, La Jolla, CA, United States

Surgeons operate in mentally and physically demanding workspaces where the impact of error is highly consequential. Accurately characterizing the neurophysiology of surgeons during intraoperative error will help guide more accurate performance assessment and precision training for surgeons and other teleoperators. To better understand the neurophysiology of intraoperative error, we build and deploy a system for intraoperative error detection and electroencephalography (EEG) signal synchronization during robot-assisted surgery (RAS). We then examine the association between EEG data and detected errors. Our results suggest that there are significant EEG changes during intraoperative error that are detectable irrespective of surgical experience level.

KEYWORDS

robot-assisted surgery, medical robotics, human performance, surgical training, clinical assessment

1. Introduction

Surgeons operate in mentally and physically demanding workspaces where the impact of error is highly consequential (Christian et al., 2006; Modi et al., 2018). The cognitive and affective states of surgeons in the operating room (OR) are primary determinants of surgical performance (Schuetz et al., 2008; Yurko et al., 2010; Haji et al., 2015). Measuring changes in these states contributes to insightful performance analysis, training development and skills improvement (Eversbusch and Grantcharov, 2004; Carswell et al., 2005). However, these cognitive and affective state measurements, which can be derived from neurophysiological metrics, must be coupled with measures of surgical task success or failure to enable any conclusions about the correspondences

between psychological indicators and surgical performance (Modi et al., 2017). Unfortunately, the majority of existing tools to measure surgical task performance are retrospective, reliant on subjective, manual expert annotation and often focused on entire procedures rather than on intraoperative performance (Maruthappu et al., 2014).

Examining the associations between neurophysiology and intraoperative performance requires detecting markers of intraoperative success or failure at millisecond to second resolutions and synchronizing those markers with neurophysiological data. In this study, we build and deploy an automated, objective error detection system for robotic-assisted surgery (RAS) simulations that is synchronized with an electroencephalogram (EEG) headset worn by trainee surgeons at an RAS operating console. We use this system to detect intraoperative errors and analyze the accompanying EEG changes during those errors.

Prior work examining EEG changes during RAS has primarily focused on correlating EEG measurements with retrospective, subjective surveys of perceived mental workload (Guru et al., 2015b; Yu et al., 2017; Zander et al., 2017; Zhou et al., 2018, 2020). While mental workload is a significant contributor to surgical performance (Sarker and Vincent, 2005; Healey et al., 2006; Balch et al., 2009), it is not a proxy for actual performance. Understanding the relationship between operator neurophysiology and actual surgical performance requires objective, high temporal-resolution measurements of intraoperative performance that can be synchronized with changes in neurophysiological indicators. Automated skill and error assessment algorithms for RAS platforms do exist (Chen et al., 2019) but are rarely used in conjunction with biometric data capture. Integrating intraoperative performance data with intraoperative physiological data enables investigation of the potential cognitive and affective features that influence surgical success or failure.

For neurophysiological assessment, we choose to focus on the EEG indicators of psychological features that have been shown to impact task performance in mentally and physically demanding workspaces as well as human-machine teams. "Focused attention" or "vigilance" is the ability of an individual to maintain sustained attention on a task (Oken et al., 2006). Decreases in focused attention increase error rates in human-machine teamwork (Gill, 2012). Fluctuations in levels of attention can be measured through changes in EEG spectral power analyses (Borghini et al., 2014). For this study, we examine global β : ($\alpha + \theta$) power ratio (Pope et al., 1995), occipital β power (Gola et al., 2013), occipital δ power (Harmony, 2013), occipital θ power (Putman et al., 2014), frontal θ power (Strijkstra et al., 2003), and global γ power (Kaiser and Lutzenberger, 2005) as quantitative indicators of attention.

Cognitive load and fatigue are similarly relevant to human-robot teaming. High levels of cognitive load lead to task errors

(Sweller, 2011). Cognitive fatigue, a decrease in cognitive performance after sustained cognitive load (Ackerman, 2011), is of particular importance in safety-critical aviation, transportation, aerospace, military, medicine, and industrial settings where fatigued individuals routinely operate complex, automated systems (Zhang et al., 2009). EEG spectral power ratios (Ryu and Myung, 2005) can also be used as measures of cognitive load or fatigue. We focus on frontal θ power (Shou and Ding, 2013; So et al., 2017; Raufi and Longo, 2022), frontal α power (Palva and Palva, 2007) and parietal α power (Frey et al., 2015) as indicators of cognitive load. For cognitive fatigue, we examine global α power (Santamaria and Chiappa, 1987), occipital α power (Thut et al., 2006; Zumer et al., 2014), frontal θ power (Strijkstra et al., 2003), and θ : α power ratio at Fz (Berka et al., 2007).

Affective characteristics such as valence, arousal, and dominance (VAD) may also influence human-robot system performance. Valence, the positivity or negativity of emotion, modulates attention and has shown an inverted u-shaped relationship with human performance in various tasks (Cai and Lin, 2011). Arousal, often referred to as "stress," has a similar inverted u-shaped relationship with cognitive load (Teigen, 1994; Staal, 2004). Dominance, feelings of being "in control" of a situation, has a positive relationship with performance during sustained tasks (Cohen et al., 2016). EEG correlates of valence include α power asymmetry (Ohme et al., 2010; Gordon et al., 2018), global β power (Liu and Sourina, 2013), global γ power (Oathes et al., 2008) and global α power (Ahern and Schwartz, 1985). Indicators of arousal include global θ power, global β power and global γ power (Liu and Sourina, 2013). EEG correlates of dominance include global α : β power ratio and parietal β power (Verma and Tiwary, 2017).

Finally, we examine the association of proposed error recognition-related EEG metrics with intraoperative error. We include global θ power (Cavanagh and Frank, 2014), and global β power (Ray and Cole, 1985) as quantitative EEG indicators of error recognition.

To correlate these EEG metrics with markers of intraoperative error, we require high temporal-resolution analysis of surgical performance data. Prior work in automated surgical performance analysis has primarily focused on kinematic, system event, and haptic data obtained from RAS platforms. This data is then used to find associations between features such as instrument traveling distance (kinematics), camera clutch engagement (system event), and grip force (haptics) and self-reported level of surgical experience (Chen et al., 2019). Other research has used surgical video analysis for competency assessment (Reiley and Hager, 2009; Tao et al., 2012; Wang and Majewicz Fey, 2018). While kinematic, system, haptic, and image data can be correlated with surgical experience and used for skills assessments, in order to examine the EEG correlates of actual task errors, automated surgical error detection is required. Prior work in

intraoperative error detection largely relies on retrospective, manual annotations by expert surgeons (Chen et al., 2019). These manual annotations provide sub-task level resolution for error detection but sub-tasks within surgical procedures can take seconds to minutes. A sub-task that is labeled with an error then implies an error marker of seconds to minutes in length. The EEG signal associated with a long-duration error marker would include more than the neurophysiological response to error.

To more precisely examine the neurophysiology of intraoperative error, we use a video analysis pipeline to automatically label each video frame from all surgical simulation videos for the presence or absence of an error. Because RAS operating console video is captured at a frame rate of 30 frames-per-second (fps) and each frame has an error label, we achieve error detection resolution of 30 Hertz (Hz). Each of these error labels is associated with a timestamp that is then synchronized with raw EEG data. By creating non-overlapping EEG data windows that align with each video frame, we create a more accurate measure of the neurophysiology of intraoperative error.

Parameterizing the EEG-based neurophysiology of intraoperative error creates the possibility of designing training interventions to reduce or prevent these errors. Surgeons who exhibit high correlations between EEG indices that may be related to attention and error, cognitive load and error, or emotional valence and error, for example, may benefit from EEG neurofeedback training to improve attention (Egner and Gruzelier, 2001, 2004), cognitive load management (Gruzelier et al., 2006), memory (Vernon et al., 2003), or emotional regulation (Raymond et al., 2005). Other surgeons who have high correlations between EEG indicators of arousal or stress and error, or sensorimotor desynchronization and error, for example, may benefit from focused neurostimulation for stress reduction (Modi et al., 2019) and fine motor skill augmentation (Cox et al., 2020).

In this paper, we aim to characterize the neurophysiological response to intraoperative error during RAS. To achieve this, we build and deploy a system for intraoperative error detection and EEG signal synchronization. We then use this system in a series of experiments with trainee surgeons and analyze the results. Our motivation is a better understanding of the neurophysiology of intraoperative error to help guide more accurate human performance assessments and training for surgeons and other teleoperators.

2. Materials and methods

2.1. Materials

We use a daVinci Xi operating console (Intuitive Surgical, Inc.) for image-based surgical data. The Xi operating console

provides video of the surgeon's field-of-view (FOV) at 30 fps. Video analysis of the operating console video (surgeon's FOV) is used for intraoperative error detection. For neurophysiological data capture, we use a Cognionics Quick-20 EEG headset (Cognionics, Inc.) with a sampling rate of 500 Hz.

2.2. Methods

2.2.1. Experimental protocol

Twenty participants were recruited for this study under institutional IRB approval. Of these participants, 5 were non-medical graduate students and the remainder were general surgery residents at various levels of training. All participants were naïve to the hypotheses of the study prior to participation.

Each participant completed three simulation tasks during a single sitting without breaks. These tasks are standard exercises used in robotic surgical training to teach surgeons operational procedures for robot-assisted surgery. They are part of a larger standard learning curriculum that is widely used in surgical education curricula that are designed, tested, and deployed to improve robot assisted surgery skills. These tasks were completed in the same order for all participants. The tasks utilized were "Ring Rollercoaster 1," "Ring Rollercoaster 3," and "Wrist Articulation 1" in this order for all participants (Figure 1).

An EEG headset was fitted on each participant prior to beginning the first task and checked for appropriate positioning and data transfer. We use the international 10–20 EEG system to ensure proper electrode placement. An example of a deployed headset is shown in Figure 2. Following appropriate electrode placement, we confirm adequate signal through the Cognionics Data Acquisition software (Cognionics, Inc.) which provides visual cues to monitor signal quality at individual EEG electrodes. Electrodes with low signal quality are re-adjusted until signal quality meets the data acquisition software criteria. This signal quality is monitored throughout the experiment, and the EEG headset is re-adjusted between simulation tasks if necessary.

Four minutes of baseline EEG readings were then recorded as each participant sat still at the operating console. Following baseline recording, each participant was directed to start the first surgical simulation task and given no other instruction. After completion of the first task, each participant was directed to start the second simulation, also without additional instruction or feedback. After completion of the second simulation, each participant was directed to start the third and final simulation, also without additional instruction or feedback. Following completion of the final simulation, each participant was instructed to complete a demographics survey after which the EEG headset was removed and the system was reset for future participants.



FIGURE 1
Surgical simulation tasks.



FIGURE 2
EEG headset deployed in RAS-equipped OR.

2.2.2. Video analysis

To generate error markers from operating console (i.e., surgeon FOV) video, we use each video frame as input for a classical computer vision feature extraction algorithm. Our algorithm uses pixel-based frequency filters, intensity filters and Hough transforms to generate a feature representation for each video frame. This feature representation is used to label each video frame for the presence or absence of intraoperative error. The types of error we consider are: contact between the manipulated object and an obstacle, excessive grip force, dropping the manipulated object, contact between end effectors, end effectors going out of view during camera clutch movement, contact between end effectors and obstacles.

To validate this video analysis pipeline, a single expert human adjudicator manually annotated a validation set of 18,887 randomly chosen operating console video frames for the presence of intraoperative error. We compared these manual annotations to the annotations generated by our computer vision algorithm on the same validation set of operating console video frames. Our algorithm's overall error detection accuracy on the validation set of 18,887 operating console video frames is 98.0%, precision is 90.4%, and recall is 93.7%. We then use this pipeline for frame-by-frame inference on the video data. Given the operating console video frame rate of 30 fps, frame-by-frame error detection provides intraoperative error markers at 30 Hz.

2.3. EEG analysis

To pre-process the raw EEG data, we remove the DC offset and bandpass filter the resulting signal with a passband of 0.01–50 Hz. We eliminate EEG noise using artifact subspace reconstruction (Chang et al., 2018). This pre-processed data is then filtered in parallel into five spectral power bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (30–50 Hz). This filtering is done on the entire EEG data stream i.e., from the start of the baseline or start of the simulation period to the end of the baseline or end of the simulation period. After this filtering is complete, the EEG data for each of the spectral power bands is then windowed.

To accurately associate each operating console video frame with a non-overlapping window of EEG data, we align each video frame timestamp with the closest EEG signal timestamp. This timestamp is the “start window” timestamp. Each EEG window includes all samples received between a “start window” timestamp and the “start window” timestamp corresponding to the next video frame. The operating console video has a frame rate of 30 fps which implies that each video frame has a duration of ~33 ms. Each EEG window therefore has a size of ~33 ms.

Using these non-overlapping ~33 ms EEG windows, we compute the root-mean-square amplitude of each band for every EEG channel. Channel-specific spectral band powers for each window (i.e., video frame) are then aggregated into features

that correspond to each operating console video frame. These features are compared to the average feature value for the baseline period.

2.4. Statistical analysis

For analysis of all EEG channels and all spectral power bands, we use separate linear mixed effects models for each channel and each band. Given 19 EEG channels and 5 spectral power bands, this results in 95 mixed effects models. In each of these 95 models, the dependent variable (channel-band power) is standardized, intraoperative error is modeled as a fixed effect and individual participant variance is modeled as a random effect.

For EEG feature analysis, we examine a total of 15 distinct EEG features. We select these 15 features (see Table 1 for

all features) based on previously published work from other contexts that associates these EEG features with cognitive and affective attributes that could be important during RAS. We do not attempt to reproduce earlier work from other contexts by testing whether these EEG features are appropriate indicators of cognition or affect in an RAS simulation. Instead, we test whether these 15 EEG features are associated with intraoperative error. We use a separate linear mixed effects model for each EEG feature. The dependent variable (EEG feature) is standardized, intraoperative error is modeled as a fixed effect and individual participant variance is modeled as a random effect.

Given the independent testing of 95 channel-band combinations as well as 15 EEG features in this experiment, we correct our baseline significance level of $P = 0.01$ by a factor of $\frac{1}{110}$. Our multiple comparisons corrected significance level is $P = 0.00009$.

TABLE 1 EEG correlates of selected cognitive and affective features from previously published research.

Cognitive/affective feature	EEG correlate	Spectral bands: EEG electrodes
Attention	Global β : ($\alpha + \theta$) power ratio (Pope et al., 1995)	β, α, θ : All
	Occipital β power (Gola et al., 2013)	β : O1, O2
	Occipital δ power (Harmony, 2013)	δ : O1, O2
	Occipital θ power (Putman et al., 2014)	θ : O1, O2
	Frontal θ power (Strijkstra et al., 2003)	θ : F4, Fz, F3
	Global γ power (Kaiser and Lutzenberger, 2005)	γ : all
Cognitive load	Frontal θ power (Shou and Ding, 2013; So et al., 2017; Raufi and Longo, 2022)	θ : F4, Fz, F3
	Frontal α power (Palva and Palva, 2007)	α : F4, Fz, F3
	Parietal α power (Frey et al., 2015)	α : P3, P4, Pz, P7, P8
Cognitive fatigue	Global α power (Santamaria and Chiappa, 1987)	α : All
	Occipital α power (Thut et al., 2006; Zumer et al., 2014)	α : O1, O2
	Frontal θ power (Strijkstra et al., 2003)	θ : F4, Fz, F3
	θ : α power ratio (Berka et al., 2007)	θ, α : Fz
Valence	α power asymmetry (Ohme et al., 2010; Gordon et al., 2018)	α : (F7, Fp1, F3, C3, P7, T3, P3, O1) / (Fp2, F8, F4, P8, P4, O2, C4, T4)
	Global β power (Liu and Sourina, 2013)	β : all
	Global γ power (Oathes et al., 2008)	γ : all
	Global α power (Ahern and Schwartz, 1985)	α : all
Arousal	Global θ power	θ : all
	Global β power	β : all
	Global γ power (Liu and Sourina, 2013)	γ : all
Dominance	Global α : β power ratio	α, β : all
	Parietal β power (Verma and Tiwary, 2017)	β : P3, P4, Pz, P7, P8
Error recognition	Global θ power (Cavanagh and Frank, 2014)	θ : all
	Global β power (Ray and Cole, 1985)	β : all

All electrodes: Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, P7, P3, Pz, P4, P8, O1, and O2.

We conducted a simulation-based power analysis (Green and MacLeod, 2016) of the EEG feature correlates using simulated effect sizes approximately equal to those reported in our results (Brybaert and Stevens, 2018) and significance levels of $P = 0.00009$. With a simulated effect size of 0.05, our statistical power to detect error-related changes in occipital delta band power has a 95% confidence interval (C.I.) of 99.6–100.0%. With a simulated effect size of -0.03 , the 95% statistical power C.I. for parietal alpha band power is 99.6 to 100.0%. With a simulated effect size of 0.03, the 95% statistical power C.I. for frontal theta band power is 99.6–100.0%. With a simulated effect size of -0.05 , the 95% statistical power C.I. for left-to-right alpha band power asymmetry is 99.6–100.0%. With a simulated effect size of -0.02 , the 95% statistical power C.I. for global beta band power is 99.3–99.9%. With a simulated effect size of -0.04 , the 95% statistical power C.I. for parietal beta band power is 99.6–100.0%.

3. Results

3.1. Video analysis

The sample size of detected errors using was smaller than that of the non-errors. There were a total of 62,809 detected error frames (17.1%) as compared to 305,234 non-error frames (82.9%). See Figure 3 for example error detection frames from each simulation task.

3.2. All channels all bands analysis

Seventy-nine of 95 channel-band power combinations differ significantly between error and non-error. The five largest magnitude standardized effect sizes correspond to δ band power relative to baseline at channels Fp2 (estimate: 0.106, standard error: 0.003, $P < 2E-16$), F4 (estimate: 0.143, standard error: 0.003, $P < 2E-16$), Pz (estimate: 0.175, standard error: 0.004, $P < 2E-16$), and P4 (estimate: 0.149, standard error: 0.004, $P < 2E-16$), and β band power relative to baseline at channel P4 (estimate: -0.119 , standard error: 0.003, $P < 2E-16$). Channel-band combinations that show no significant difference between error and non-error are: δ at Fp1, F8, P3, P7; θ at Cz, O1, C4; α at F4, Cz, P8, P4, O1; β at P7, T4; and γ at P7, T3 (Figure 4).

3.3. All features analysis

Twelve of 15 EEG features differ significantly between error and non-error. For attention, δ power at the occipital electrodes shows the largest magnitude difference between error and non-error (estimate: 0.05, standard error: 0.004, $P < 2E-16$). For cognitive load, α power at the parietal electrodes has the largest magnitude difference (estimate: -0.037 , standard error: 0.003, P

$< 2E-16$). The most informative feature for cognitive fatigue is θ power at the frontal electrodes (estimate: 0.034, standard error: 0.003, $P < 2E-16$). For valence, the most informative feature is α power asymmetry between the left and right electrodes (estimate: -0.059 , standard error: 0.004, $P < 2E-16$). For arousal, β power globally has the largest magnitude difference (estimate: -0.027 , standard error: 0.003, $P < 2E-16$). For dominance, global $\alpha:\beta$ power ratio (estimate: -0.040 , standard error: 0.004, $P < 2E-16$) and β power at the parietal electrodes (estimate: -0.040 , standard error: 0.003, $P < 2E-16$) have similar effect sizes. The most informative feature for error recognition is global β power (estimate: -0.027 , standard error: 0.003, $P < 2E-16$) (Figure 5).

To validate our EEG feature results against potential signals due primarily to EEG pre-processing methods, we compare our full model (a fixed effect of intraoperative error and a random effect of individual variance) with a model that only incorporates the random effect of individual variance in post-processed EEG data. Using a simulation based approach, the likelihood ratio lower bounds for the 95% statistical power C.I.s to detect intraoperative error effects in the full EEG feature models as compared to the simple random effect models ranged from 98.84 to 99.63%. To validate our EEG feature results against potential signals in EEG noise, we compare our full model using post-processed EEG data with our full model using raw EEG data. The likelihood ratio lower bounds for the 95% statistical power C.I.s to detect intraoperative error effects using post-processed EEG data as compared to raw EEG data ranged from 99.3 to 99.6%.

4. Discussion

Potential cognitive and affective determinants of surgical performance, and their EEG correlates, have been explored in prior research (Gruzelier, 2014; Guru et al., 2015a; Ndaro and Wang, 2018). This work, however, does not couple EEG analysis with automated surgical error detection to capture high temporal-resolution indicators of intraoperative error. By integrating intraoperative error detection with synchronized EEG analysis, we aimed to characterize the neurophysiology of surgeons during robot-assisted surgery (RAS), particularly during intraoperative error.

To achieve this, we built and deployed a novel RAS data capture and analysis platform to examine the relationships among EEG spectral band powers at various channels, previously proposed EEG correlates of various cognitive and affective features in other contexts, and intraoperative error. Our primary statistical tools, linear mixed effects models (LMMs), are widely used in psychophysiological research particularly because of the non-independence of the data in hierarchical structures (Bagiella et al., 2000; Koerner and Zhang, 2017; Frömer et al., 2018). While sequential 30ms EEG samples may be correlated, we are not examining a temporal effect in our model, so the impact of potential residual error autocorrelation



FIGURE 3 Error detection example for simulation tasks.

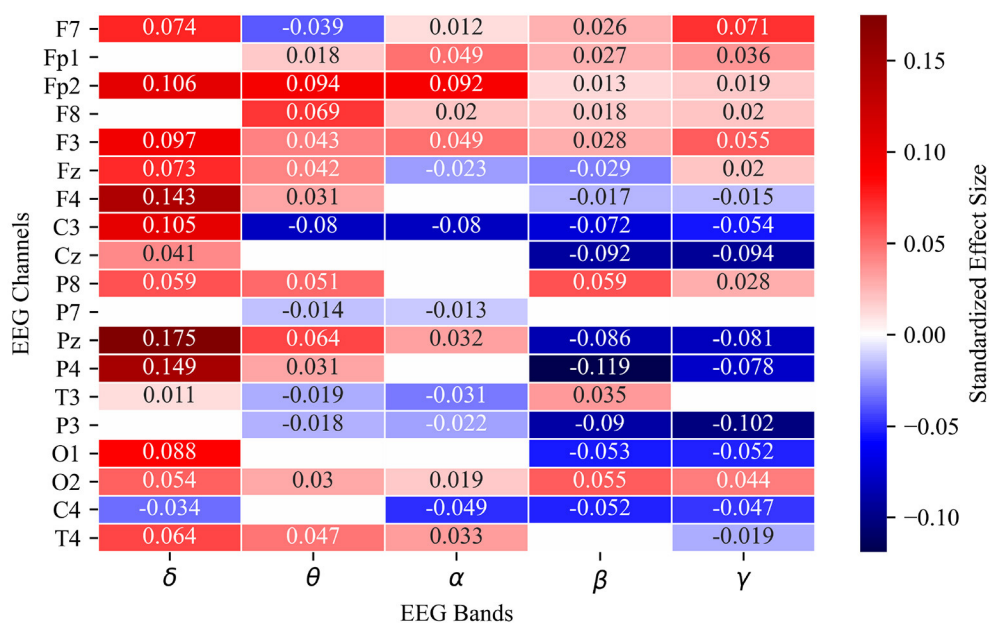


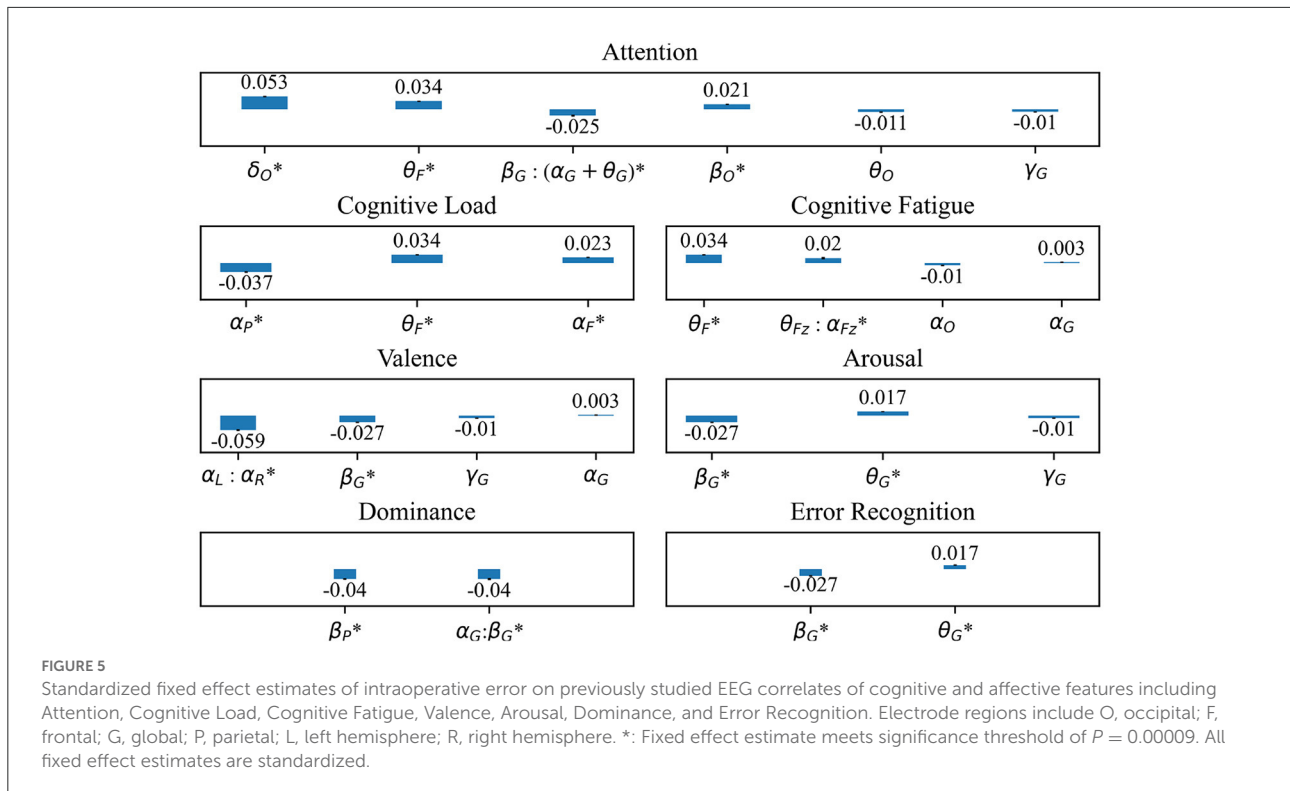
FIGURE 4 Heat map of estimated fixed effects for 95 linear mixed effects models representing all channel-band combinations of 19 EEG channels and 5 spectral power bands. Fixed effect estimate labeled numerically and by color intensity in each cell. Only significant channel-band combinations are included in heat map. Blank cells did not meet corrected significance threshold of $P = 0.00009$.

on the error effect is not meaningful for our hypotheses. For this reason, and to our knowledge, the standard LMMs in neurophysiological research do not use autoregressive residual covariance matrices (West et al., 2006). Additionally, when we randomly shuffle our datasets to remove any potential residual autocorrelation that may impact our results, our effect sizes and significance tests remain unchanged.

Our results suggest that relative band powers at most channels demonstrate significant changes during errors as compared to when errors are not being committed. We observe that EEG features which are proposed correlates of attention, cognitive load, cognitive fatigue, valence, arousal, dominance, and error recognition in other contexts (Klimesch, 1999; Strijkstra et al., 2003; Harmony, 2013; Liu and Sourina, 2013;

Verma and Tiwary, 2017; Tempel et al., 2020) demonstrate significant changes during intraoperative error and may underlie the error-related signals detected in this study.

The direction of change in EEG features largely aligns with proposed associations between EEG statistics and cognitive or affective features in other contexts. Increases in occipital δ power have been associated with increases in attention (Harmony, 2013), and in our study, these increases are also observed during intraoperative error. Decreases in parietal α power have been associated with decreases in cognitive load (Klimesch, 1999). In our study, decreases in parietal α power were also observed during intraoperative error. Increases in frontal θ power have been associated with increases in cognitive fatigue (Strijkstra et al., 2003), and we observed an increase



in frontal θ power during error. Left and right hemisphere α power asymmetry has been proposed as an indicator of emotional valence (Strijkstra et al., 2003) with higher asymmetry corresponding to more positive valence. We observed a decrease in α power asymmetry during error. Decreases in global β power have been associated with increased arousal (Liu and Sourina, 2013), and we observed a decrease in global β power during error. Increases in parietal β power have been associated with increases in dominance (Verma and Tiwary, 2017), and we observed decreases in parietal β power during error. For error recognition, increases in global β power have been associated with increased response inhibition during voluntary motor tasks (Tempel et al., 2020). In our study, global β power decreased during intraoperative error.

While this study attempts to characterize the physiological correlates of intraoperative error, this study is not designed to test the hypothesis that the cognitive and affective states correlate with physiological metrics that are causative of error. The EEG activations described in these results are exploratory and intended to generate hypotheses for future studies that can formally test hypotheses generated from the current findings. Moreover, correlations between physiological signals are not indicative of causation and therefore future studies may attempt to implement causal interventions to test the directional relationship between errors and physiological states.

There are several limitations of this study. First, the number of participants was neither large nor evenly distributed across

skill level, potentially skewing cohort analysis. Second, using bandpass filters and RMS power calculations for frame-by-frame channel-band and EEG feature power measurements are accompanied by a tradeoff in the settling time of the filter. Third, the error detection video pipeline we use is based on color and text-based error indicators that appear in the operating console during surgical simulations. These error indicators are part of the built-in simulation software and cannot be turned off or altered during simulation tasks. This presents the possibility that the neurophysiological changes that accompany error may be confounded by a response to the specific error indicator in the stimulator. That said, we observed significant changes in parietal channels that would seem to be less likely if the neurophysiological changes were due to vision-based stimuli alone. Finally, these results may be affected by other variables such as movement at the console and time at the console. Based on our observations, low performers tended to move and shift body positions more frequently than high performers, potentially indicative of discomfort at the console. Low performers also required more time to complete the simulation tasks.

Despite these limitations, this study shows that a novel RAS data capture and analysis platform can enable the detection of distinct neurophysiological changes associated with intraoperative error. These changes are largely consistent with known relationships between EEG metrics and cognitive or affective factors that impact performance. This data highlights

the possibility of using objective biometric data capture and error detection to better understand the neurophysiology of intraoperative error and help guide more accurate human performance assessments and precision training for surgeons and other teleoperators.

Data availability statement

The datasets presented in this article are not readily available because HIPAA. Requests to access the datasets should be directed to CD'A.

Ethics statement

The studies involving human participants were reviewed and approved by University of California, San Diego School of Medicine Institutional Review Board IRB#: 800026. The patients/participants provided their written informed consent to participate in this study.

Author contributions

CD'A: conceptualization, software, formal analysis, and writing—original draft. EA-S: supervision and funding acquisition. EH: investigation and project administration.

References

- Ackerman, P. L. (2011). *Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications*. Washington, DC: American Psychological Association.
- Ahern, G. L., and Schwartz, G. E. (1985). Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis. *Neuropsychologia* 23, 745–755. doi: 10.1016/0028-3932(85)90081-8
- Bagiella, E., Sloan, R. P., and Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology* 37, 13–20. doi: 10.1111/1469-8986.3710013
- Balch, C. M., Freischlag, J. A., and Shanafelt, T. D. (2009). Stress and burnout among surgeons: understanding and managing the syndrome and avoiding the adverse consequences. *Arch. Surg.* 144, 371–376. doi: 10.1001/archsurg.2008.575
- Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., et al. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviat. Space Environ. Med.* 78, B231–B244.
- Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neurosci. Biobehav. Rev.* 44, 58–75. doi: 10.1016/j.neubiorev.2012.10.003
- Brybaert, M., and Stevens, M. (2018). Power analysis and effect size in mixed effects models: a tutorial. *J. Cogn.* 1, 9. doi: 10.5334/joc.10
- Cai, H., and Lin, Y. (2011). Modeling of operators' emotion and task performance in a virtual driving environment. *Int. J. Hum. Comput. Stud.* 69, 571–586. doi: 10.1016/j.ijhcs.2011.05.003
- Carswell, C. M., Clarke, D., and Seales, W. B. (2005). Assessing mental workload during laparoscopic surgery. *Surg. Innov.* 12, 80–90. doi: 10.1177/155335060501200112

NG: investigation. HC and LA: writing—review and editing, supervision, and funding acquisition. RB: writing—review and editing, project administration, supervision, and funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

LA was supported by the Army Research Office, Grant#: W911NF-15-1-0390.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Cavanagh, J. F., and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* 18, 414–421. doi: 10.1016/j.tics.2014.04.012
- Chang, C.-Y., Hsu, S.-H., Pion-Tonachini, L., and Jung, T.-P. (2018). "Evaluation of artifact subspace reconstruction for automatic eeg artifact removal," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Honolulu, HI: IEEE), 1242–1245.
- Chen, J., Cheng, N., Cacciamani, G., Oh, P., Lin-Brandt, M., Remulla, D., et al. (2019). Objective assessment of robotic surgical technical skill: a systematic review. *J. Urol.* 201, 461–469. doi: 10.1016/j.juro.2018.06.078
- Christian, C. K., Gustafson, M. L., Roth, E. M., Sheridan, T. B., Gandhi, T. K., Dwyer, K., et al. (2006). A prospective study of patient safety in the operating room. *Surgery* 139, 159–173. doi: 10.1016/j.surg.2005.07.037
- Cohen, I., den Braber, N., Smets, N. J., van Diggelen, J., Brinkman, W.-P., and Neerinx, M. A. (2016). Work content influences on cognitive task load, emotional state and performance during a simulated 520-days' mars mission. *Comput. Hum. Behav.* 55, 642–652. doi: 10.1016/j.chb.2015.10.011
- Cox, M. L., Deng, Z.-D., Palmer, H., Watts, A., Beynel, L., Young, J. R., et al. (2020). Utilizing transcranial direct current stimulation to enhance laparoscopic technical skills training: a randomized controlled trial. *Brain Stimul.* 13, 863–872. doi: 10.1016/j.brs.2020.03.009
- Egner, T., and Gruzelić, J. H. (2001). Learned self-regulation of EEG frequency components affects attention and event-related brain potentials in humans. *Neuroreport* 12, 4155–4159. doi: 10.1097/00001756-200112210-00058
- Egner, T., and Gruzelić, J. H. (2004). EEG biofeedback of low beta band components: frequency-specific effects on variables of attention and event-related brain potentials. *Clin. Neurophysiol.* 115, 131–139. doi: 10.1016/S1388-2457(03)00353-5

- Eversbusch, A., and Grantcharov, T. (2004). Learning curves and impact of psychomotor training on performance in simulated colonoscopy: a randomized trial using a virtual reality endoscopy trainer. *Surg. Endoscopy Other Intervent. Techn.* 18, 1514–1518. doi: 10.1007/s00464-003-9264-9
- Frey, J. N., Ruhnau, P., and Weisz, N. (2015). Not so different after all: the same oscillatory processes support different types of attention. *Brain Res.* 1626, 183–197. doi: 10.1016/j.brainres.2015.02.017
- Frömer, R., Maier, M., and Abdel Rahman, R. (2018). Group-level eeg-processing pipeline for flexible single trial-based analyses including linear mixed models. *Front. Neurosci.* 12, 48. doi: 10.3389/fnins.2018.00048
- Gill, K. S. (2012). *Human Machine Symbiosis: The Foundations of Human-Centred Systems Design*. Springer Science & Business Media.
- Gola, M., Magnuski, M., Szumska, I., and Wróbel, A. (2013). EEG beta band activity is related to attention and attentional deficits in the visual performance of elderly subjects. *Int. J. Psychophysiol.* 89, 334–341. doi: 10.1016/j.ijpsycho.2013.05.007
- Gordon, R., Ciorciari, J., and van Laer, T. (2018). Using eeg to examine the role of attention, working memory, emotion, and imagination in narrative transportation. *Eur. J. Mark.* 52, 92–117. doi: 10.1108/EJM-12-2016-0881
- Green, P., and MacLeod, C. J. (2016). Simr: an r package for power analysis of generalized linear mixed models by simulation. *Methods Ecol. Evolut.* 7, 493–498. doi: 10.1111/2041-210X.12504
- Gruzelier, J., Egner, T., and Vernon, D. (2006). Validating the efficacy of neurofeedback for optimising performance. *Prog. Brain Res.* 159, 421–431. doi: 10.1016/S0079-6123(06)59027-2
- Gruzelier, J. H. (2014). Eeg-neurofeedback for optimising performance. i: a review of cognitive and affective outcome in healthy participants. *Neurosci. Biobehav. Rev.* 44, 124–141. doi: 10.1016/j.neubiorev.2013.09.015
- Guru, K. A., Esfahani, E. T., Raza, S. J., Bhat, R., Wang, K., Hammond, Y., et al. (2015a). Cognitive skills assessment during robot-assisted surgery: separating the wheat from the chaff. *BJU Int.* 115, 166–174. doi: 10.1111/bju.12657
- Guru, K. A., Shafiei, S. B., Khan, A., Hussein, A. A., Sharif, M., and Esfahani, E. T. (2015b). Understanding cognitive performance during robot-assisted surgery. *Urology* 86, 751–757. doi: 10.1016/j.urology.2015.07.028
- Haji, F. A., Rojas, D., Childs, R., de Ribaupierre, S., and Dubrowski, A. (2015). Measuring cognitive load: performance, mental effort and simulation task complexity. *Med. Educ.* 49, 815–827. doi: 10.1111/medu.12773
- Harmony, T. (2013). The functional significance of delta oscillations in cognitive processing. *Front. Integr. Neurosci.* 7, 83. doi: 10.3389/fnint.2013.00083
- Healey, A., Sevdalis, N., and Vincent, C. (2006). Measuring intra-operative interference from distraction and interruption observed in the operating theatre. *Ergonomics* 49, 589–604. doi: 10.1080/00140130600568899
- Kaiser, J., and Lutzenberger, W. (2005). Human gamma-band activity: a window to cognitive processing. *Neuroreport* 16, 207–211. doi: 10.1097/00001756-200502280-00001
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3
- Koerner, T. K., and Zhang, Y. (2017). Application of linear mixed-effects models in human neuroscience research: a comparison with pearson correlation in two auditory electrophysiology studies. *Brain Sci.* 7, 26. doi: 10.3390/brainsci7030026
- Liu, Y., and Sourina, O. (2013). “EEG databases for emotion recognition,” in *2013 International Conference on Cyberworlds* (Yokohama: IEEE), 302–309.
- Maruthappu, M., El-Harasis, M., Nagendran, M., Orgill, D., McCulloch, P., Duclos, A., et al. (2014). Systematic review of methodological quality of individual performance measurement in surgery. *J. Br. Surg.* 101, 1491–1498. doi: 10.1002/bjs.9642
- Modi, H. N., Singh, H., Fiorentino, F., Orihuela-Espina, F., Athanasiou, T., Yang, G.-Z., et al. (2019). Association of residents’ neural signatures with stress resilience during surgery. *JAMA Surg.* 154, e192552-e192552. doi: 10.1001/jamasurg.2019.2552
- Modi, H. N., Singh, H., Orihuela-Espina, F., Athanasiou, T., Fiorentino, F., Yang, G.-Z., et al. (2018). Temporal stress in the operating room: brain engagement promotes “coping” and disengagement prompts “choking”. *Ann. Surg.* 267, 683–691. doi: 10.1097/SLA.0000000000002289
- Modi, H. N., Singh, H., Yang, G.-Z., Darzi, A., and Leff, D. R. (2017). A decade of imaging surgeons’ brain function (part i): terminology, techniques, and clinical translation. *Surgery* 162, 1121–1130. doi: 10.1016/j.surg.2017.05.021
- Ndaro, N. Z., and Wang, S.-Y. (2018). Effects of fatigue based on electroencephalography signal during laparoscopic surgical simulation. *Minim. Invasive Surg.* 2018, 2389158. doi: 10.1155/2018/2389158
- Oathes, D. J., Ray, W. J., Yamasaki, A. S., Borkovec, T. D., Castonguay, L. G., Newman, M. G., et al. (2008). Worry, generalized anxiety disorder, and emotion: evidence from the eeg gamma band. *Biol. Psychol.* 79, 165–170. doi: 10.1016/j.biopsycho.2008.04.005
- Ohme, R., Reykowska, D., Wiener, D., and Choromanska, A. (2010). Application of frontal eeg asymmetry to advertising research. *J. Econ. Psychol.* 31, 785–793. doi: 10.1016/j.joep.2010.03.008
- Oken, B. S., Salinsky, M. C., and Elsas, S. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clin. Neurophys.* 117, 1885–1901. doi: 10.1016/j.clinph.2006.01.017
- Palva, S., and Palva, J. M. (2007). New vistas for α -frequency band oscillations. *Trends Neurosci.* 30, 150–158. doi: 10.1016/j.tins.2007.02.001
- Pope, A. T., Bogart, E. H., and Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biol. Psychol.* 40, 187–195. doi: 10.1016/0301-0511(95)05116-3
- Putman, P., Verkuil, B., Arias-Garcia, E., Pantazi, I., and van Schie, C. (2014). EEG theta/beta ratio as a potential biomarker for attentional control and resilience against deleterious effects of stress on attention. *Cogn. Affect. Behav. Neurosci.* 14, 782–791. doi: 10.3758/s13415-013-0238-7
- Raufi, B., and Longo, L. (2022). An evaluation of the eeg alpha-to-theta and theta-to-alpha band ratios as indexes of mental workload. *Front. Neuroinform.* 16, 861967. doi: 10.3389/fninf.2022.861967
- Ray, W. J., and Cole, H. W. (1985). Eeg alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science* 228, 750–752. doi: 10.1126/science.3992243
- Raymond, J., Varney, C., Parkinson, L. A., and Gruzelier, J. H. (2005). The effects of alpha/theta neurofeedback on personality and mood. *Cogn. Brain Res.* 23, 287–292. doi: 10.1016/j.cogbrainres.2004.10.023
- Reiley, C. E., and Hager, G. D. (2009). “Task versus subtask surgical skill evaluation of robotic minimally invasive surgery,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer), 435–442. doi: 10.1007/978-3-642-04268-3_54
- Ryu, K., and Myung, R. (2005). Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *Int. J. Ind. Ergon.* 35, 991–1009. doi: 10.1016/j.ergon.2005.04.005
- Santamaria, J., and Chiappa, K. H. (1987). The eeg of drowsiness in normal adults. *J. Clin. Neurophysiol.* 4, 327–382. doi: 10.1097/00004691-198710000-00002
- Sarker, S. K., and Vincent, C. (2005). Errors in surgery. *Int. J. Surg.* 3, 75–81. doi: 10.1016/j.ijsu.2005.04.003
- Schuetz, M., Gockel, I., Beardi, J., Hakman, P., Dunschede, F., Moenk, S., et al. (2008). Three different types of surgeon-specific stress reactions identified by laparoscopic simulation in a virtual scenario. *Surg. Endosc.* 22, 1263–1267. doi: 10.1007/s00464-007-9605-1
- Shou, G., and Ding, L. (2013). “Frontal theta eeg dynamics in a real-world air traffic control task,” in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Osaka: IEEE), 5594–5597.
- So, W. K., Wong, S. W., Mak, J. N., and Chan, R. H. (2017). An evaluation of mental workload with frontal EEG. *PLoS ONE* 12, e0174949. doi: 10.1371/journal.pone.0174949
- Staal, M. A. (2004). *Stress, cognition, and human performance: A literature review and conceptual framework*. Technical Memorandum (TM) 20060017835.
- Strijkstra, A. M., Beersma, D. G., Drayer, B., Halbesma, N., and Daan, S. (2003). Subjective sleepiness correlates negatively with global alpha (8–12 Hz) and positively with central frontal theta (4–8 Hz) frequencies in the human resting awake electroencephalogram. *Neurosci. Lett.* 340, 17–20. doi: 10.1016/S0304-3940(03)00033-8
- Sweller, J. (2011). Cognitive load theory. *Psychol. Learn. Motiv.* 55, 37–76. doi: 10.1016/B978-0-12-387691-1.00002-8
- Tao, L., Elhamifar, E., Khudanpur, S., Hager, G. D., and Vidal, R. (2012). “Sparse hidden markov models for surgical gesture classification and skill evaluation,” in *International Conference on Information Processing in Computer-Assisted Interventions* (Springer), 167–177. doi: 10.1007/978-3-642-30618-1_17
- Teigen, K. H. (1994). Yerkes-dodson: a law for all seasons. *Theory Psychol.* 4, 525–547. doi: 10.1177/09593543940444004
- Tempel, T., Frings, C., and Pastötter, B. (2020). EEG beta power increase indicates inhibition in motor memory. *Int. J. Psychophysiol.* 150, 92–99. doi: 10.1016/j.ijpsycho.2020.02.008
- Thut, G., Nietzel, A., Brandt, S. A., and Pascual-Leone, A. (2006). α -band electroencephalographic activity over occipital cortex indexes visuospatial

- attention bias and predicts visual target detection. *J. Neurosci.* 26, 9494–9502. doi: 10.1523/JNEUROSCI.0875-06.2006
- Verma, G. K., and Tiwary, U. S. (2017). Affect representation and recognition in 3D continuous valence-arousal-dominance space. *Multimed Tools Appl.* 76, 2159–2183. doi: 10.1007/s11042-015-3119-y
- Vernon, D., Egner, T., Cooper, N., Compton, T., Neilands, C., Sheri, A., et al. (2003). The effect of training distinct neurofeedback protocols on aspects of cognitive performance. *Int. J. Psychophysiol.* 47, 75–85. doi: 10.1016/S0167-8760(02)00091-0
- Wang, Z., and Majewicz Fey, A. (2018). Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. *Int. J. Comput. Assist. Radiol. Surg.* 13, 1959–1970. doi: 10.1007/s11548-018-1860-1
- West, B. T., Welch, K. B., and Galecki, A. T. (2006). *Linear Mixed Models: A Practical Guide Using Statistical Software*. Chapman and Hall; CRC. doi: 10.1201/b17198
- Yu, D., Dural, C., Morrow, M., Yang, L., Collins, J. W., Hallbeck, S., et al. (2017). Intraoperative workload in robotic surgery assessed by wearable motion tracking sensors and questionnaires. *Surg. Endosc.* 31, 877–886. doi: 10.1007/s00464-016-5047-y
- Yurko, Y. Y., Scerbo, M. W., Prabhu, A. S., Acker, C. E., and Stefanidis, D. (2010). Higher mental workload is associated with poorer laparoscopic performance as measured by the nasa-tlx tool. *Simulat. Healthcare* 5, 267–271. doi: 10.1097/SIH.0b013e3181e3f329
- Zander, T. O., Shetty, K., Lorenz, R., Leff, D. R., Krol, L. R., Darzi, A. W., et al. (2017). Automated task load detection with electroencephalography: towards passive brain-computer interfacing in robotic surgery. *J. Med. Rob. Res.* 2, 1750003. doi: 10.1142/S2424905X17500039
- Zhang, C., Zheng, C.-X., and Yu, X.-L. (2009). Automatic recognition of cognitive fatigue from physiological indices by using wavelet packet transform and kernel learning algorithms. *Exp. Sys. App.* 36, 4664–4671. doi: 10.1016/j.eswa.2008.06.022
- Zhou, T., Cha, J. S., Gonzalez, G., Wachs, J. P., Sundaram, C. P., and Yu, D. (2020). Multimodal physiological signals for workload prediction in robot-assisted surgery. *ACM Trans. Hum. Rob. Interact.* 9, 1–26. doi: 10.1145/3368589
- Zhou, T., Cha, J. S., Gonzalez, G. T., Wachs, J. P., Sundaram, C., and Yu, D. (2018). “Joint surgeon attributes estimation in robot-assisted surgery,” in *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (IEEE)*, 285–286. doi: 10.1145/3173386.3176981
- Zumer, J. M., Scheeringa, R., Schoffelen, J.-M., Norris, D. G., and Jensen, O. (2014). Occipital alpha activity during stimulus processing gates the information flow to object-selective cortex. *PLoS Biol.* 12, e1001965. doi: 10.1371/journal.pbio.1001965