



OPEN ACCESS

EDITED BY

Frederic Dehais,
Institut Supérieur de l'Aéronautique et de
l'Espace (ISAE-SUPAERO), France

REVIEWED BY

Ahmet Omurtag,
Nottingham Trent University, United Kingdom
Daniel Callan,
Advanced Telecommunications Research
Institute International (ATR), Japan

*CORRESPONDENCE

Katharina Lingelbach
✉ katharina.lingelbach@iao.fraunhofer.de

RECEIVED 27 November 2024

ACCEPTED 30 January 2025

PUBLISHED 19 February 2025

CITATION

Lingelbach K, Rips J, Karstensen L,
Mathis-Ullrich F and Vukelić M (2025)
Evaluating robotic actions: spatiotemporal
brain dynamics of performance assessment in
robot-assisted laparoscopic training.
Front. Neuroergonomics 6:1535799.
doi: 10.3389/fnrgo.2025.1535799

COPYRIGHT

© 2025 Lingelbach, Rips, Karstensen,
Mathis-Ullrich and Vukelić. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Evaluating robotic actions: spatiotemporal brain dynamics of performance assessment in robot-assisted laparoscopic training

Katharina Lingelbach^{1,2*}, Jennifer Rips¹, Lennart Karstensen³,
Franziska Mathis-Ullrich³ and Mathias Vukelić¹

¹Applied Neurocognitive Systems, Fraunhofer Institute for Industrial Engineering IAO, Stuttgart, Germany, ²Applied Neurocognitive Psychology, Department of Psychology, Carl von Ossietzky University, Oldenburg, Germany, ³Department Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-University, Erlangen, Germany

Introduction: Enhancing medical robot training traditionally relies on explicit feedback from physicians to identify optimal and suboptimal robotic actions during surgery. Passive brain-computer interfaces (BCIs) offer an emerging alternative by enabling implicit brain-based performance evaluations. However, effectively decoding these evaluations of robot performance requires a comprehensive understanding of the spatiotemporal brain dynamics identifying optimal and suboptimal robot actions within realistic settings.

Methods: We conducted an electroencephalographic study with 16 participants who mentally assessed the quality of robotic actions while observing simulated robot-assisted laparoscopic surgery scenarios designed to approximate real-world conditions. We aimed to identify key spatiotemporal dynamics using the surface Laplacian technique and two complementary data-driven methods: a mass-univariate permutation-based clustering and multivariate pattern analysis (MVPA)-based temporal decoding. A second goal was to identify the optimal time interval of evoked brain signatures for single-trial classification.

Results: Our analyses revealed three distinct spatiotemporal brain dynamics differentiating the quality assessment of optimal vs. suboptimal robotic actions during video-based laparoscopic training observations. Specifically, an enhanced left fronto-temporal current source, consistent with P300, LPP, and P600 components, indicated heightened attentional allocation and sustained evaluation processes during suboptimal robot actions. Additionally, amplified current sinks in right frontal and mid-occipito-parietal regions suggested prediction-based processing and conflict detection, consistent with the oERN and interaction-based ERN/N400. Both mass-univariate clustering and MVPA provided convergent evidence supporting these neural distinctions.

Discussion: The identified neural signatures propose that suboptimal robotic actions elicit enhanced, sustained brain dynamics linked to continuous attention allocation, action monitoring, conflict detection, and ongoing evaluative processing. The findings highlight the importance of prioritizing late evaluative

brain signatures in BCIs to classify robotic actions reliably. These insights have significant implications for advancing machine-learning-based training paradigms.

KEYWORDS

robot training, performance monitoring, spatio-temporal clustering, temporal decoding, machine learning, electroencephalography (EEG), passive brain-computer interfaces (BCIs), current source density (CSD)

1 Introduction

Current research is advancing the development and optimization of robotic systems capable of autonomously performing specialized tasks and providing adaptive assistance to support surgeons during various stages of procedures (Moustris et al., 2011; Richter et al., 2019; Thananjeyan et al., 2017). These tasks include camera guidance (Pandya et al., 2014), tissue clamping (Nguyen et al., 2019), tissue manipulation (Scheikl et al., 2024), and surgical knot tying (Osa et al., 2014; Van Den Berg et al., 2010).

Machine learning, particularly reinforcement learning, is well-suited for training robots efficiently, allowing them to learn tasks autonomously (Iturrate et al., 2010; Vukelić et al., 2023). A key challenge, however, lies in providing effective feedback to the reinforcement learning agent. The agent requires frequent and continuous evaluation of its actions via a reward function to distinguish between successful and unsuccessful outcomes. Reinforcement learning is typically trained in simulated environments using this reward function before being adapted to real-world settings for fine-tuning or deployment. The design of the reward function and the real-world fine-tuning both rely on the expertise of physicians. However, obtaining explicit feedback in the form of labels for robot actions from physicians is challenging, as it further burdens their already demanding workload.

Passive brain-computer interfaces (BCIs) offer a promising approach by enabling direct, implicit and continuous feedback loops in human-robot interactions (e.g., Aricò et al., 2018; Protzak et al., 2013), thereby alleviating the burden on physicians (Zander et al., 2017). Brain signals elicited during the observation and mental assessment of robot actions can serve as an evaluation function for reinforcement learning models (Kim et al., 2017; Vukelić et al., 2023).

Previous studies on performance monitoring tasks, including those in BCI applications (Chavarriaga et al., 2010; Iturrate et al., 2015; Ehrlich and Cheng, 2019; Ferrez and Millán, 2005; Ferrez and Millán, 2008; Kreilinger et al., 2012; Spüler and Niethammer, 2015), have shown that observing errors is associated with pronounced event-related potential (ERP) deflections, particularly in the following components (see Somon et al., 2017 for review): Across various tasks, an observation-based error-related negativity (oERN; Somon et al., 2017) has been consistently identified, resembling the ERN observed in self-generated errors (Gehring et al., 1993; also referred to as error negativity (Ne) in early studies; Falkenstein et al., 1991). However, the oERN peaks slightly later, between 250 and 270 ms, in frontocentral regions and is enhanced

in response to erroneous actions (Chavarriaga et al., 2010; Ferrez and Millán, 2005; Ferrez and Millán, 2008; Somon et al., 2017; Pavone et al., 2016).

The oERN is sometimes followed by a frontocentral positivity known as error positivity (oPe), which responds to errors depending on contextual factors such as task engagement and error relevance. This component tends to be absent when another observed agent produces the error without relational impact or direct consequence for the observer (Chavarriaga et al., 2010; van Schie et al., 2004; Koban et al., 2010; Padrao et al., 2016). The oPe peaks between 350 and 450 ms and is thought to reflect conscious recognition and high-level evaluation of errors (Ferrez and Millán, 2005; Ferrez and Millán, 2008; Somon et al., 2017; Pavone et al., 2016).

Many of the studies on error monitoring in observed agents and systems (Ferrez and Millán, 2005; Chavarriaga et al., 2010; Padrao et al., 2016; Pavone et al., 2016) have identified a further negative ERP deflection, likely linked to prediction violations and unexpected events. This monitoring-related ERP termed the interaction ERN by Ferrez and Millán (2005), peaks at frontocentral sites between 400 and 550 ms and is proposed to be related to the N400. Initially linked to semantic inconsistencies, the N400 typically peaks around 450 ms post-stimulus at centroparietal sites (Kutas and Hillyard, 1980). However, it has also been observed in non-semantic contexts, such as unexpected outcomes in movement sequences, with a more frontocentral and temporoparietal distribution (Balconi and Vitaloni, 2014).

Building on this foundation, promising results have emerged in training non-medical robots using these error-related ERPs (Iturrate et al., 2010, 2015; Kim et al., 2017, 2020; Luo et al., 2018; Penalzoza et al., 2015; Salazar-Gomez et al., 2017; Vukelić et al., 2023). Despite these advances, the application of BCI-based training for medical robots in realistic scenarios remains scarce.

This study investigated evoked spatiotemporal dynamics associated with evaluating optimal and suboptimal robot actions during a robot-assisted laparoscopic simulation using electroencephalography (EEG). Our objectives were twofold: (a) to determine whether the spatiotemporal dynamics evoked by observing optimal and suboptimal robotic actions in near-naturalistic laparoscopic robot training videos resemble commonly reported error-related potentials, using two complementary analytical approaches; and (b) to identify the optimal time interval of these evoked brain signatures for single-trial classification, with potential application for feedback loops in BCI-driven reinforcement learning systems.

2 Materials and methods

2.1 Participants

Sixteen volunteers ($M_{age} = 24.88$ years, $SD = 4.88$, range: 19–38 years, 14 females, two males) with no prior experience in surgical procedures participated in the study. Eligibility criteria included age between 18 and 40 years, right-handedness, absence of diagnosed neurological, physiological, or psychological disorders, no regular use of centrally acting substances, and no head implants or history of brain surgeries. Participants provided written informed consent before participation and received monetary compensation. The study complied with the Declaration of Helsinki and was approved by the University of Tübingen Ethics Committee (ID: 827/2020BO1).

2.2 Procedure

At the beginning of the experiment, EEG signals were recorded during a 2-min resting period while participants focused on a fixation cross with their eyes open. Participants subsequently undertook an evaluation task, requiring them to observe laparoscopic video sequences and mentally assess the quality of the robotic action depicted in each sequence.

The video sequences illustrated simulated tissue-cutting procedures performed by a robotic arm using a rod instead of a scalpel. These procedures were conducted on a phantom torso model with replicated organs, offering realistic representations of robotic actions in laparoscopic surgery. Light-emitting diodes (LEDs) were used to mark the target organ and tissue for the surgical cut. The target organs included the right kidney, stomach and spleen, each equipped with a single LED point sensor, and the left kidney, which was fitted with a line sensor consisting of a row of seven LEDs (Figures 1A–D). The optimal action required the robot to press the rod with sufficient pressure onto the target organ for the point sensor and to move the rod along the organ's surface for the line sensor. If the robot applied adequate pressure to the marked tissue, the LEDs turned off (Figure 1A). Conversely, if the tissue was missed or the pressure was insufficient, the LEDs remained fully or partially lit. Detailed information about the stimulus material and an illustrative overview video are provided in the [Supplementary material](#). The stimulus database is accessible upon request through the OSF repository at <https://osf.io/6ndsv/>.

Participants rated each robotic action as good or bad, giving verbal responses during the practice phase to confirm task comprehension, and conducting mental evaluations during the actual experiment. High classification accuracy of robot action assessments was confirmed in a preliminary behavioral study ($N = 9$; see [Supplementary material](#) for details).

Following a brief practice session consisting of 15 video sequences to familiarize them with the task and video material at the beginning of the experiment, participants were presented with 1,000 video sequences across 10 blocks. Each block included an overview video, a countdown, and randomized combination of 65 sequences showing optimal robotic actions and 35 showing suboptimal actions (100 sequences per block; Figures 1, 2). A 1-min break followed each block. The overview video at the beginning

of each block depicted a sequence of optimal and suboptimal robotic actions from two viewpoints (zoomed-out in the top left corner of the screen and zoomed-in in the bottom right corner of the screen), providing context for the medical scenario and upcoming 100 zoomed-in video sequences. Each zoomed-in video sequence contained a single robotic action. It lasted 1.5 s and was followed by a jittered interstimulus interval ranging from 0.75 to 1 s, during which a fixation cross appeared at the center of the screen (Figure 2).

2.3 Data acquisition and preprocessing

EEG potentials were recorded according to the international 10–20 system with 64 electrodes and at a sampling rate of 1,000 Hz (actiCAP and BrainAmp, BrainProducts GmbH, Germany). The locations of the electrodes were Fp1, Fp2, Fz, AF3, AF4, AF7, AF8, F1, F2, F3, F4, F5, F6, F7, F8, FC1, FC2, FC3, FC4, FC5, FC6, FT7, FT8, FT9, FT10, Cz, C1, C2, C3, C4, C5, C6, T7, T8, CPz, CP1, CP2, CP3, CP4, CP5, CP6, TP7, TP8, TP9, TP10, Pz, P1, P2, P3, P4, P5, P6, P7, P8, POz, PO3, PO4, PO7, PO8, Oz, O1, O2, and Iz. The ground electrode was positioned on FPz and the reference electrode on FCz. Impedance of electrodes was kept below 25 k Ω at the beginning of the experiment.

All analyses were performed in Python and MNE Python (Gramfort et al., 2014). The EEG signals were de-trended and bandpass filtered using a fourth-order infinite impulse response (IIR) Butterworth filter with cut-off frequencies of 0.2 and 10 Hz (see also Iturrate et al., 2010, 2015; Kim et al., 2017, 2020; Vukelić et al., 2023). The signals were then segmented into 2.2-second epochs, each beginning 200 ms before the onset of each zoomed-in video sequence. Epoched data was subsequently down-sampled to 250 Hz. To remove cardiac, muscle, and ocular artifacts, epochs were cleaned using an independent component analysis (ICA; Chaumon et al., 2015; Hipp and Siegel, 2013; Lee et al., 1999) within an automated pipeline called FASTER (Nolan et al., 2010) as implemented in mne-python version 1.6.1 (Gramfort et al., 2014). To generate an electro-oculography (EOG) surrogate for the ICA, a virtual EOG channel was constructed using the frontal Fp1 and Fp2 electrode signals. After cleaning the signals, the epochs were baseline corrected by subtracting the mean amplitude of the time interval before the video onset (200 ms) and bad channels were interpolated per epoch using a spline interpolation (Gramfort et al., 2014; Nolan et al., 2010). Finally, the reference-free current source density (CSD) transformation was applied to the data to enhance spatial resolution by minimizing volume conduction effects and estimating local electrical activity (current sources and sinks) at the scalp surface (Perrin et al., 1989; Kayser and Tenke, 2015).

CSD is a mathematical transformation of EEG signals that estimates local current sources and sinks across the cortical surface at the sensor level. By computing the second spatial derivative of the electric potential field, it determines the spatial distribution and direction of current flow. Notably, the number of output channels matches the input channels, as the transformation is applied directly to the data from each electrode without changing the input dimensionality. CSD distinguishes between current sources (positive polarity) and sinks (negative polarity). In a CSD map, a source indicates outward current flow from a cortical region,

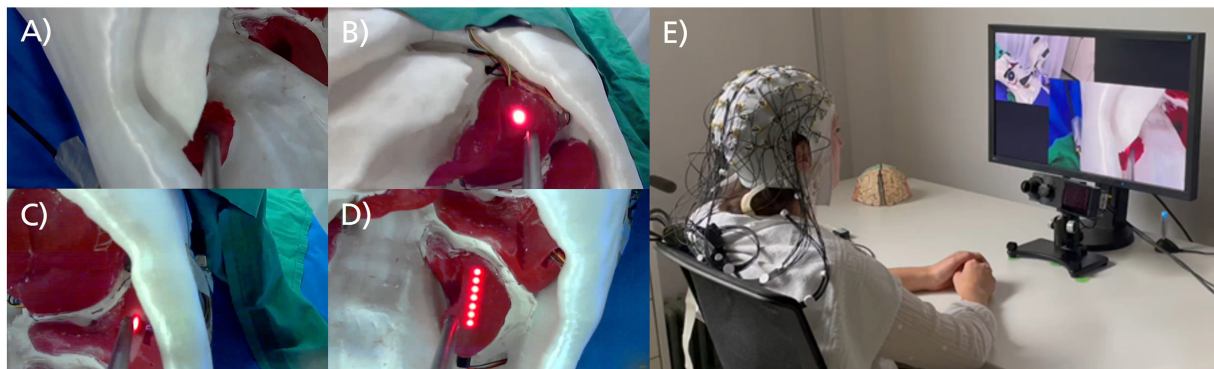


FIGURE 1

Excerpt from zoomed-in video sequences and laboratory setup. (A) Video sequence of the right kidney with a point sensor. The LED light turned off during the video, indicating the robot's action was successful. (B) Video sequence of the stomach with a point sensor. The sensor was not touched during the video, and the robot's action failed. (C) Video sequence of the spleen with a point sensor. The sensor was touched during the video, but the pressure was insufficient, resulting in a failed action. (D) Video sequence of the left kidney with a line sensor. The line sensor was not touched during the video, leading to a failed action. (E) Laboratory setup with a participant seated in front of the monitor and eye-tracking system, wearing a 64-channel EEG.

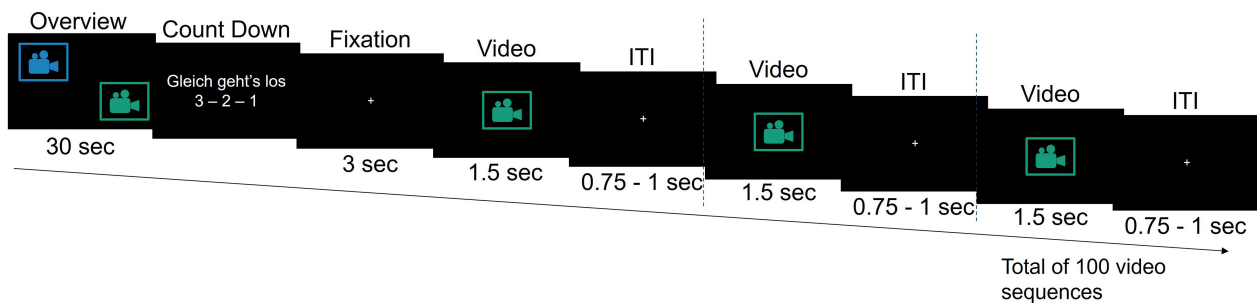


FIGURE 2

Overview of an experimental block in the robot action evaluation task. The overview phase at the beginning of each block includes two camera views. The following countdown is displayed in German (English: "It's about to start 3 - 2 - 1"). ITI, interstimulus interval.

reflecting reduced excitatory postsynaptic potentials (EPSPs). In contrast, a sink represents inward current flow linked to increased EPSPs (Perrin et al., 1989; Kayser and Tenke, 2015). This approach offers a more localized and directly interpretable representation of neural activity than standard reference-dependent EEG potentials (Perrin et al., 1989; Kayser and Tenke, 2015).

For subsequent analyses, the number of epochs was equalized across conditions by minimizing timing discrepancies across trial lists, ensuring an identical epoch count per condition.

2.4 Mass-univariate permutation-based clustering

To examine differences in brain signatures evoked by the robot actions, we used mass-univariate permutation-based spatiotemporal clustering (Maris and Oostenveld, 2007) with a paired *t*-test. The clustering was performed on contrast data, calculated by subject-wise subtracting suboptimal from optimal evoked responses.

Compared to traditional univariate approaches, such as performing an ANOVA or *t*-test on the mean or peak amplitude within a predefined time interval, mass-univariate statistics allow

statistical testing at every location and time point (e.g., Maris and Oostenveld, 2007; Pernet et al., 2015; Groppe et al., 2011).

This approach is particularly advantageous when addressing variability in ERP latencies arising from experimental parameters, such as complex stimulus material (e.g., in the case of the P300; Bentin et al., 1999). However, the multiple comparisons problem—occurring when testing across many locations and time points—must be accounted for. Mass-univariate permutation-based spatiotemporal clustering addresses this issue by identifying clusters of contiguous samples (i.e., time points and sensors) that exhibit similar effects, thereby reducing the number of comparisons to the cluster level (Maris and Oostenveld, 2007). Neighboring effects (test statistics of time points and sensors) that exceed a predefined univariate cluster-forming threshold (here $\alpha < 0.05$) are grouped into clusters. Statistical values (e.g., *t*- or *F*-values) within these clusters are aggregated, for instance by summing them, into cluster-mass scores (Maris and Oostenveld, 2007). Statistical significance is then determined by comparing the observed cluster-mass scores to a reference null distribution, generated via random resampling of condition labels (e.g., using Monte Carlo permutations or bootstrapping). A *p*-value is calculated for each cluster as the proportion of permutations in which the cluster-level statistic from the null distribution equals or exceeds the observed

cluster-mass score obtained from the original dataset. To control the overall Type I error rate (false positives) across all clusters, only clusters with a p -value below a predefined group-level threshold (here $\alpha < 0.05$) are considered statistically significant.

2.5 Temporal decoding with a linear machine learning model

Temporal decoding with subject-wise multivariate pattern analysis (MVPA) provides an alternative to mass-univariate analyses, offering enhanced sensitivity and statistical power (Holdgraf et al., 2017; Kriegeskorte and Douglas, 2019). MVPA leverages the multidimensional characteristics of neurophysiological data from each subject, thereby accounting for anatomical and functional inter-individual neural variability (Marsicano et al., 2024).

Figure 3 illustrates the input data structure and pipeline steps applied in temporal decoding. For the machine-learning (ML) based analyses, epoched data were downsampled to 100 Hz to reduce computational costs. Linear discriminant analysis (LDA), using a least-squares solution and automatic shrinkage via the Ledoit-Wolf lemma (as implemented in scikit-learn version 1.4.1), was applied as a sliding supervised ML algorithm (i.e., the Base Estimator) on a time-point-by-time-point basis (implemented in mne-python version 1.6.1; Gramfort et al., 2014). The data of each participant (shape: n epochs, n channels, n timepoints; Figure 3) was split into training and testing sets using a repeated stratified five-fold cross-validation with 20 iterations, resulting in 100 folds per time point. In total, 220 (timepoints) \times 100 (cross-validation folds) \times 16 (participants) models were trained and fitted in the time decoding. The Area Under the Receiver Operating Characteristic Curve (ROC-AUC, henceforth referred to as AUC) was used as performance metric. Classification performance was statistically evaluated by bootstrapping the AUC scores across participants and folds in a Monte Carlo simulation (MCS; 5,000 iterations), yielding the bootstrapped mean and its 95% confidence interval (CI; Cumming, 2014). Time intervals were considered significant if the lower CI boundary of the average LDA performance exceeded the upper CI boundary of an average dummy performance (i.e., an empirical baseline estimated by chance-level stratified classification in scikit-learn version 1.4.1).

After fitting the linear models, model decoding weights were transformed into activation patterns representing their contribution to classification through inverse computations (Haufe et al., 2014). These activation patterns were averaged across participants and visualized using topographic maps. A spatiotemporal mask was applied to identify statistically significant activation patterns using univariate bootstrapped means and CIs (MCS with 5,000 iterations). Only patterns at electrode positions where the CI for the average evoked response contrast (suboptimal–optimal robot actions) excluded zero were considered significant and visualized. Positive values in the activation patterns indicate that the region contributes to the classification of evaluated suboptimal robot actions, whereas negative values indicate a contribution to the classification of evaluated optimal robot actions. Pattern values closer to zero indicate lower confidence

in their contribution. To assess the relationship between evoked response amplitudes and significant patterns, the time course of contributing regions was visualized, along with bootstrapped means and CIs for each condition at the time point of maximal classification performance.

2.6 Single-trial decoding

In the final analysis, we decoded the observer's evaluation of robot actions from brain signatures on a trial-by-trial basis using three distinct time intervals for feature extraction identified through MVPA time decoding (see Figure 3 for an illustrative overview). These intervals were defined as (1) 0–750 ms, (2) 750–1,350 ms, and (3) 1,350–2,000 ms after video onset. As in the time decoding, data were downsampled to 100 Hz to reduce computational costs.

An LDA classifier with automatically extracted features based on Riemannian geometry has been proven effective for state decoding in passive BCIs (Lotte et al., 2018; Vukelić et al., 2023) and was, thus, applied to each time interval in a within-subject single-trial decoding (implemented in pyRiemann; version 0.5). The Riemannian-based method operates directly on the epoched EEG time series (data shape: n epochs, n channels, n timepoints; Figure 3), obviating the need for manual feature extraction. It converts the EEG time series into symmetric positive definite (SPD) covariance matrices and applies Riemannian geometry to analyse these matrices (Congedo et al., 2017; Appriou et al., 2020; Vukelić et al., 2023). In the Riemannian manifold, covariance matrices were spatially filtered with the xDAWN algorithm (Rivet et al., 2009) before being projected into tangent space for transformation into Euclidean vectors (Barachant et al., 2011). This tangent space projection preserves the manifold structure while enabling effective classification (Appriou et al., 2020).

Classification was performed using an LDA classifier (with default settings as implemented in scikit-learn version 1.4.1). Performance was quantified using a repeated stratified k -fold cross-validation (five splits, 20 iterations) with AUC as metric. As with temporal decoding, a dummy classifier estimated chance-level performance. Non-parametric bootstrapping of classification scores across folds and subjects yielded the average performance and corresponding CI for each classifier, enabling statistical evaluation (Cumming, 2014).

3 Results

3.1 Mass-univariate permutation-based clustering

The non-parametric permutation-based clustering identified significant spatiotemporal differences in evoked responses when observing suboptimal compared to optimal robot actions across five clusters.

The first two clusters emerged \sim 440 ms after video onset, revealing lateralised frontal responses. Observing suboptimal robot actions resulted in a reduced left-hemispheric frontal current sink (Figure 4A; 13 electrodes; $p < 0.001$) and an enhanced

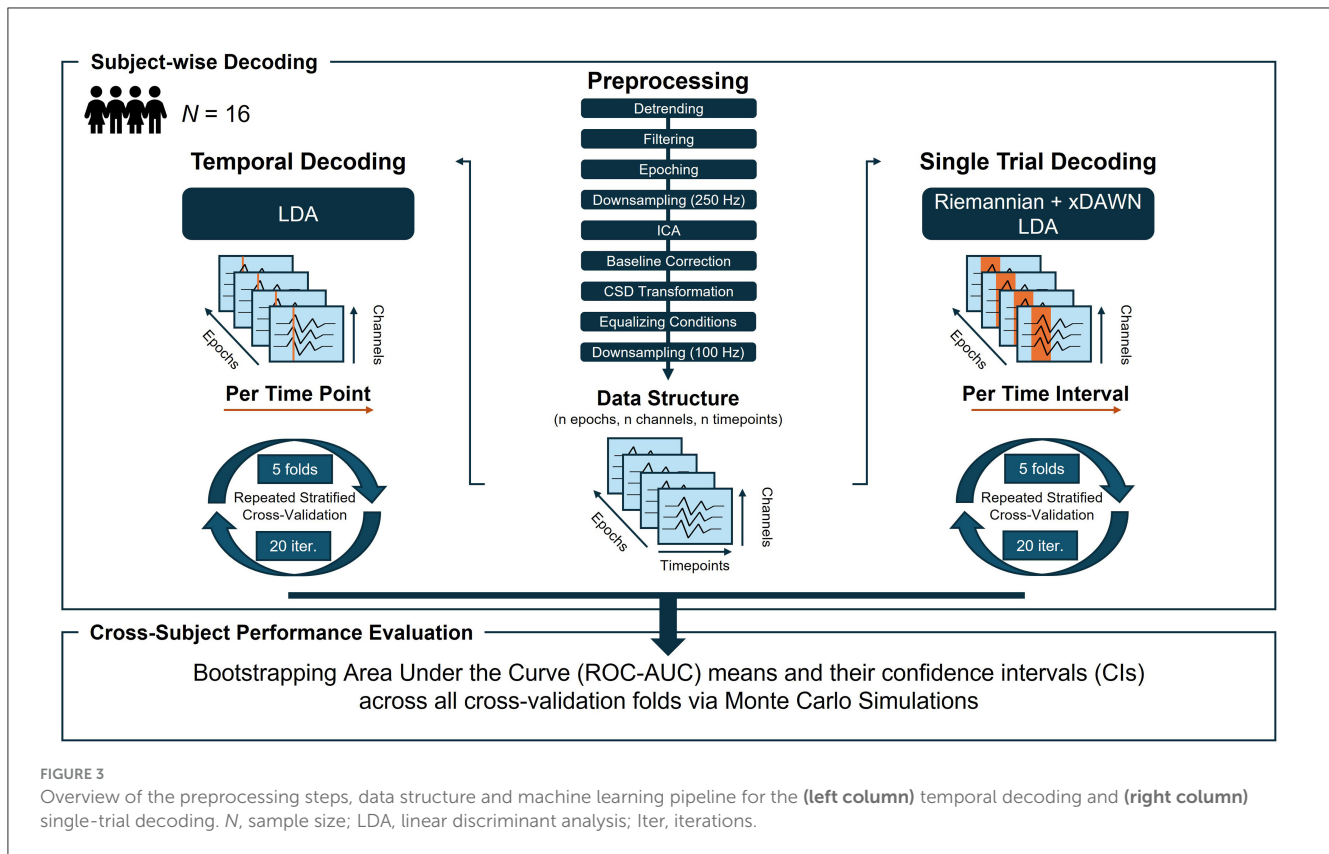


FIGURE 3

Overview of the preprocessing steps, data structure and machine learning pipeline for the (left column) temporal decoding and (right column) single-trial decoding. *N*, sample size; LDA, linear discriminant analysis; iter, iterations.

right-hemispheric frontal current sink (Figure 4B; six electrodes; $p < 0.001$). In electrodes overlying right-hemispheric frontal regions, observing optimal robot actions was even associated with current sources (i.e., a positive deflection) from around 500 ms until the analysis window's end (Figure 4B). The third cluster, including 25 electrodes over occipital, parietal, and left temporal regions, emerged at 448 ms. It differentiated robot actions by showing a reduced current source peak around 550 ms, followed by an increased current sink from 800 to 1,760 ms for suboptimal compared to optimal actions (Figure 4C; $p < 0.001$). The fourth cluster, with five electrodes over right parieto-temporal regions, appeared at 460 ms, showing increased current sources for suboptimal actions (Figure 4D; $p < 0.017$). Finally, the fifth cluster over fronto-central regions, emerging after 576 ms, showed a decreased current sink for suboptimal actions (Figure 4E; eight electrodes; $p < 0.013$). All clusters persisted almost until the end of the 2-s analysis interval (1,760–1,996 ms).

3.2 Temporal decoding with a linear machine learning model

Temporal decoding using MVPA and LDA successfully distinguished the brain signatures evoked by observing optimal vs. suboptimal robot actions.

The empirical chance level of the dummy classifier was estimated at an AUC score of 48.4 95% CI [48.06, 48.79]. In later intervals, beginning 750 ms post-stimulus onset and continuing until the end of the 2-s analysis period, classification

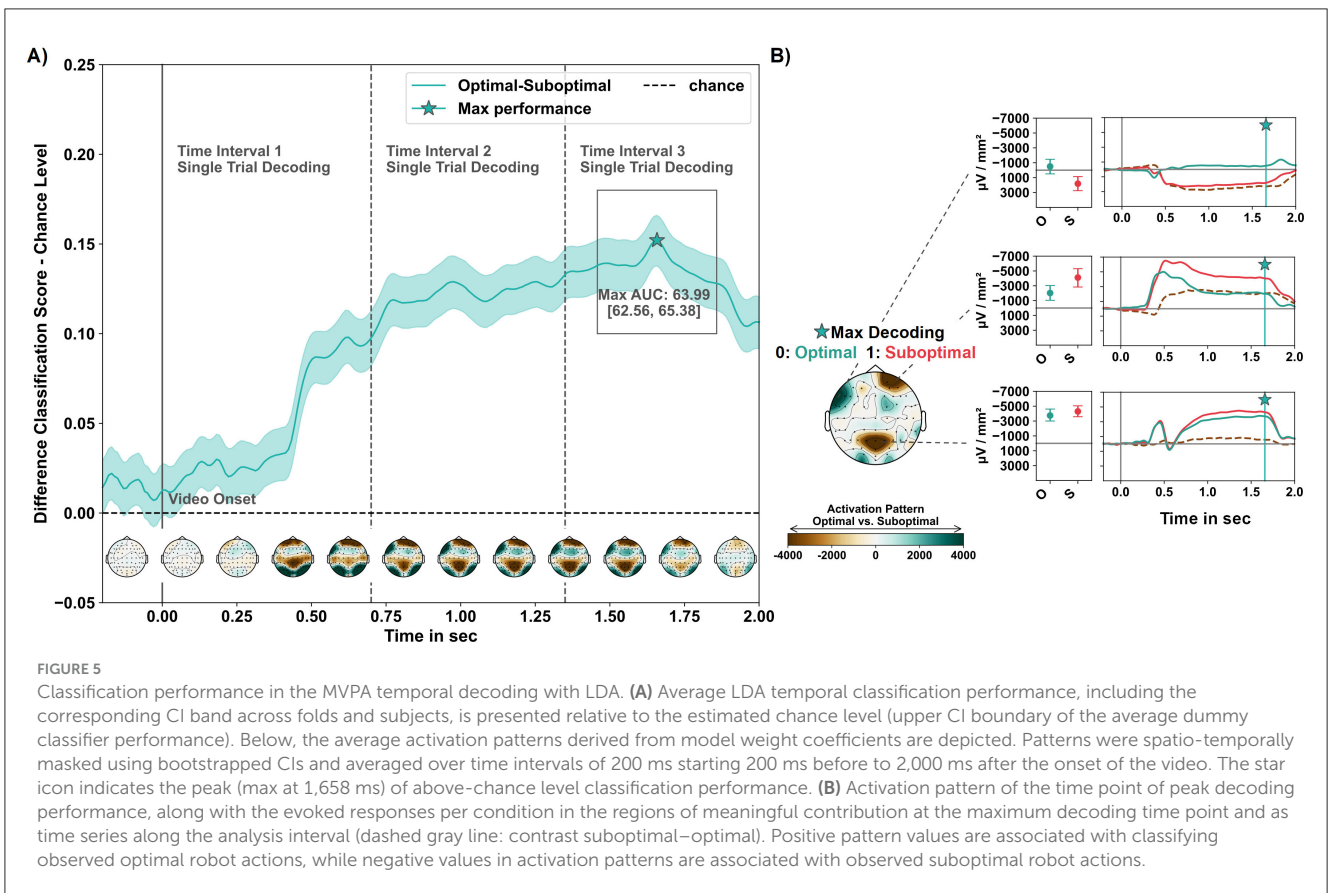
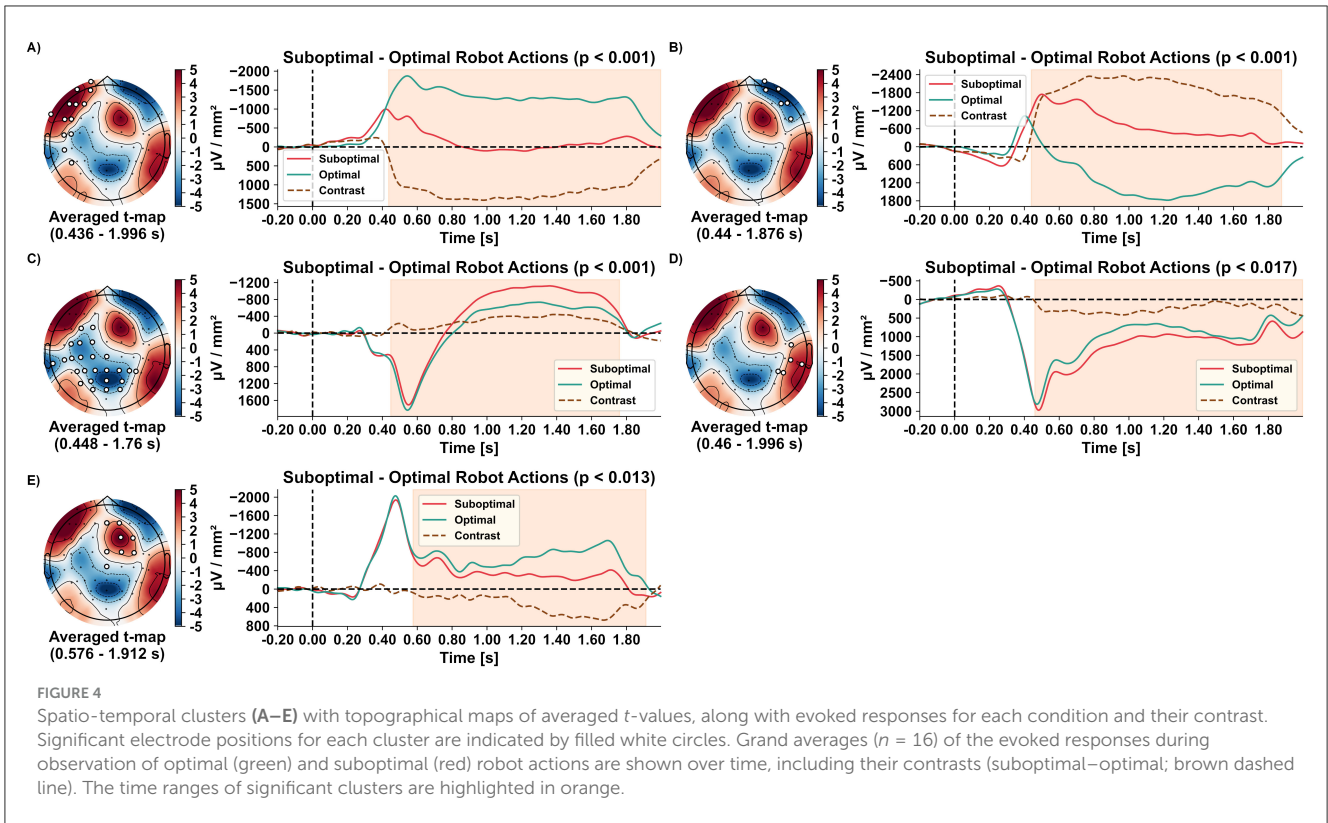
performance consistently exceeded a 60% AUC score. The classification performance varied over the analysis interval, with a standard deviation of 4.13 (4.11, 4.13). The highest classification performance was observed after 1,658 ms with an AUC score of 63.99 95% CI [62.56, 65.38], representing a difference of 15.21 (95% CI [13.78, 16.6]) to the upper CI boundary of the mean chance performance (see Figure 5A).

At the peak decoding time, significant activation patterns highlighted three regions of interest that differentiated between optimal and suboptimal robot actions. These regions included electrodes over the right frontal, left fronto-temporal, and mid-parietal areas, corresponding to three clusters identified in the mass-univariate permutation-based analysis.

Figure 5B shows the relationship between classification-contributing regions and the brain signatures evoked in these regions by the conditions. The pattern that classified suboptimal robot actions comprised electrodes positioned over a left fronto-temporal region (F7 and FT9) and revealed a current source for suboptimal actions, while optimal robot actions elicited a current sink. Two other regions contributed to classifying optimal robot actions: Current sinks in a right frontal electrode (Fp2) and electrodes overlying the mid-parietal region (P1, Pz, P2) were reduced for evaluating optimal compared to suboptimal actions.

3.3 Single-trial decoding

In the single-trial decoding of robot performance evaluations, the Riemannian LDA combined with xDawn spatial filtering yielded classification results above chance level for all selected



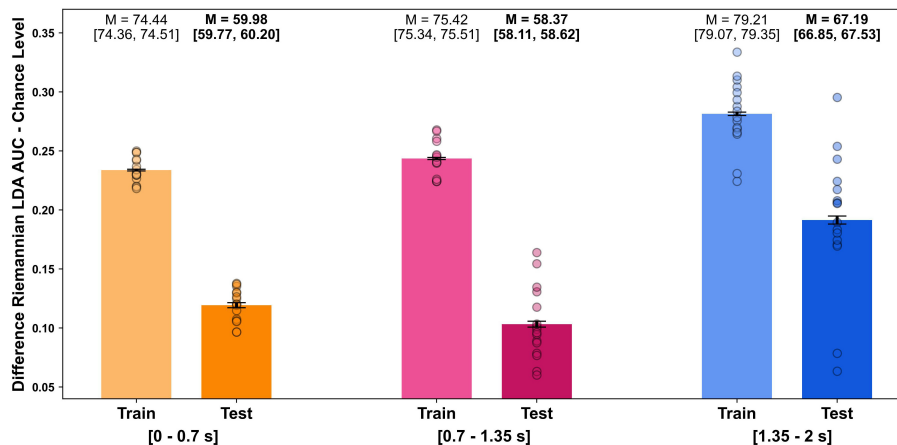


FIGURE 6

Riemannian LDA single-trial decoding performance by time interval. Bar plots display the average AUC classification scores of the Riemannian LDA relative to the upper CI boundary of the estimated chance level (dummy performance: train AUC = 48.02, 95% CI [47.99, 51.06]; test AUC = 51.04, 95% CI [51.02, 51.06]). The mean AUC score (M) and its 95% CI (I), estimated via bootstrapping and represented by error bars, are displayed above each bar (for training and test datasets across time intervals). Individual subject decoding performances are depicted as scattered dots. Time intervals of the decoding were (1) 0 to 700 ms, (2) 700 to 1,350 ms, and (3) 1,350 to 2,000 ms after video onset.

time intervals (dummy performance: train AUC = 48.21, 95% CI [48.15, 48.28]; test AUC = 51.37, 95% CI [51.36, 51.38]). The highest classification performance was observed using the latest interval including evoked responses from 1,350 to 2,000 ms after video onset, with a test AUC of 67.19 (95% CI [66.85, 67.53]). This interval also included the time point of peak decoding performance in the MVPA-based temporal decoding. In contrast, earlier intervals cropped before 750 ms post-stimulus showed a significant decrease in performance, with test AUCs of 59.98 (95% CI [59.77, 60.20]) for an interval from 0 to 700 ms and 58.37 (95% CI [58.11, 58.62]) for an interval from 700 to 1,350 ms (Figure 6).

4 Discussion

Our study identified distinct spatiotemporal brain dynamics that reliably differentiate the mental performance evaluation of optimal and suboptimal robotic actions observed in video excerpts of laparoscopic training procedures.

We assessed the robustness of neural signatures by employing surface Laplacian transformations to enhance the spatial resolution of evoked responses (see Somon et al., 2019) and two complementary data-driven methods - a mass-univariate permutation-based clustering and multivariate pattern analysis (MVPA) temporal decoding. The identified discriminative spatiotemporal brain signatures suggest that differentiation between optimal and suboptimal actions does not occur during early perceptual stages but rather at later evaluative stages (Somon et al., 2017; Ferrez and Millán, 2005; Chavarriaga et al., 2010; Oliveira et al., 2007). This finding was observed despite a perceptual component introduced by LED feedback in the evaluation task.

In addition to analyzing stimulus-locked evoked responses, we examined single-trial decoding performance of robot action evaluations across different time intervals of averaged evoked brain responses. The Riemannian LDA with xDawn filtering reliably

classified observers' electrophysiological responses to optimal and suboptimal robot actions on a trial-by-trial level. Temporal dynamics of classification performance revealed that late intervals (from 1,350 to 2,000 ms post-stimulus) significantly outperformed earlier intervals aligning with findings from MVPA temporal decoding. This indicates that the most informative brain patterns are linked to attentional and evaluative processes related to prediction violations and unexpected events (Somon et al., 2017; Chavarriaga et al., 2010; Ferrez and Millán, 2005; Oliveira et al., 2007).

4.1 Convergent findings for evoked responses differentiating robot performance

Our clustering analyses revealed five spatiotemporal brain signatures associated with robot performance evaluation, of which three were replicated in the temporal decoding. The other two clusters including electrodes localized over right mid-fronto-central and temporal areas were exclusively identified in the mass-univariate analysis. Thus, they exhibited limited reliability as distinctive patterns for evaluating robot performance in near-naturalistic scenarios. Therefore, the next section focuses on the remaining three signatures located over the left fronto-temporal, right frontal, and mid-occipito-parietal regions.

4.1.1 Left fronto-temporal spatiotemporal signature

Evoked responses in electrodes overlying left fronto-temporal regions differentiated the evaluation of optimal and suboptimal robot performance in both the clustering (Figure 4A) and temporal decoding (Figure 5B, upper row). This brain dynamic is characterized by differences in current direction-switching around 420 ms post-stimulus onset between suboptimal and

optimal robot actions. Observing suboptimal performance evoked a persistent current source (see [Figure 5B](#), upper row), while optimal performance elicited a sustained current sink during late time intervals (see [Figures 4A, 5B](#), upper row).

The late shift to a current source during suboptimal actions may indicate the allocation of additional cognitive resources for conflict processing and deviation detection ([Botvinick et al., 2001](#); [Ullsperger et al., 2014](#); [Bartholow et al., 2005](#); [Pailing and Segalowitz, 2004](#)). Although typical oERN or oPE responses were not observed in this study, the sustained fronto-temporal responses align with components such as the P300, late positive potential (LPP) and P600 ([Somon et al., 2017](#); [Sassenhagen et al., 2014](#); [Oliveira et al., 2007](#)). The P300 and LPP are positive deflections that typically emerge around 300 ms after significant and emotionally salient stimuli, respectively, at centroparietal electrode sites ([Polich, 2007](#); see [Hajcak and Foti, 2020](#) for review). The P300 appears as a broad peak, while the LPP can be sustained for up to 1,000 ms or more. Their amplitudes increase in response to motivationally significant but also deviant and uncertain stimuli ([Scheffers and Coles, 2000](#); [Sutton et al., 1965](#)), indicating sustained attentional allocation toward these stimuli ([Ridderinkhof et al., 2009](#); [Hajcak and Foti, 2020](#); [Falkenstein et al., 2000](#)). The P600, initially linked to processing linguistic anomalies ([Sassenhagen et al., 2014](#)), has also been observed during error processing in choice-reaction time tasks with enhanced amplitudes following errors [Falkenstein et al. \(1991\)](#).

In summary, the sustained current source observed during suboptimal robot actions likely reflects increased cognitive and attentional engagement in a persistent evaluative stance. This state likely facilitates conflict detection by assessing action accuracy and adequacy, monitors deviations, and supports cognitive flexibility.

4.1.2 Right frontal and mid-occipito-parietal spatiotemporal signatures

In addition to the left fronto-temporal signature, we observed two spatiotemporal signatures characterized by enhanced current sinks for evaluated suboptimal robot performance in both, the cluster analysis ([Figures 4B, C](#)) and temporal decoding activation patterns ([Figure 5B](#), middle and lower row).

After ~300 ms, a right frontal current sink emerged, peaking between 400 and 600 ms, with a delayed but pronounced deflection in response to suboptimal actions ([Figures 4B, 5B](#), middle). Another spatiotemporal brain signature, indicative of suboptimal actions and located over mid-occipito-parietal areas, appeared between 350 and ~500 ms ([Figures 4C, 5B](#), lower row). This mid-occipito-parietal signature is characterized by a current sink deflection in response to both optimal and suboptimal robot actions, followed by a short time interval of current source with a peak at 550 ms. Afterwards, another directional switch from source to sink is observed, occurring around 600 ms second in the decoding and 700 ms in the clustering analysis. In both analyses, this sustained current sink in late time intervals after stimulus onset was more pronounced when observing suboptimal compared to optimal robot performance.

These time windows and sustained current sinks for suboptimal robot actions likely reflect a combination of a delayed oERN and

an interaction ERN/N400 ([Chavarriga et al., 2010](#); [Ferrez and Millán, 2005](#); [Ferrez and Millán, 2008](#); [Somon et al., 2017](#)). The delay in evoked response intervals is potentially attributable to the erroneous robot action occurring shortly after the video onset. Notably, the N400 has previously been observed in non-linguistic contexts over parietal areas in response to unexpected motor sequences ([Balconi and Vitaloni, 2014](#)). Both ERP components are amplified when observing erroneous, suboptimal actions. In their sustained form, they may reflect ongoing quality evaluation, signaling deviations from predicted trajectories and expected movements, thereby indicating suboptimal performance.

To summarize, through temporal decoding and clustering analyses, we identified three consistent spatiotemporal signatures that distinguish the evaluation of optimal and suboptimal robot performance. A left fronto-temporal signature, characterized by an enhanced current source resembling ERP components such as the P300, LPP, and P600, suggests increased attentional allocation and sustained evaluation of suboptimal robot actions. Furthermore, right frontal and mid-occipito-parietal signatures displayed amplified current sinks in response to suboptimal actions, suggesting prediction-based processing of deviations and errors, consistent with the oERN and interaction-based ERN/N400.

4.2 Effects of task load and video stimulus material

The identified discriminative evoked signatures reflect a sustained, step-by-step evaluation of robot actions from continuous video excerpts. They persisted even after deviations from expected (optimal) performance were detected. Consequently, optimal robot actions were characterized by the absence of deviations throughout the entire video. In our specific task, participants were required to monitor and mentally assess multiple aspects of the action, including the position, length, and pressure of the intended cut. Thus, even if the robot correctly reached the target position, participants needed to verify that all criteria were met. Accordingly, it is noteworthy that the continuous video stimulus, coupled with the ongoing monitoring and evaluation of robot actions in an applied scenario, likely imposed a substantial perceptual and cognitive load on participants.

This task-induced load may have reduced differences in the amplitude of evoked responses between observed suboptimal and optimal actions ([Somon et al., 2017, 2019](#); see [Endrass et al., 2012a,b](#) for load effects during self-monitoring). It could explain the lack of modulated amplitudes in early components during the observation of suboptimal actions. In addition, although the task instructions aimed to emphasize the importance of errors and the potentially serious consequences of mistakes in laparoscopic surgery, the absence of a modulated Pe component in response to suboptimal robot actions may be due to the low (self-related) relevance of negative outcomes for participants in a passive observation role ([Chavarriga et al., 2010](#); [Somon et al., 2017](#)).

To conclude, given that task-induced cognitive load on the observer may be inherently present and unavoidable in real-world applications, further investigation is warranted to ensure ecologically valid and robust correlates of performance assessment.

4.3 Limitations and future directions

The study offers valuable insights into the neural mechanisms underlying robot performance evaluation and error monitoring in a near-naturalistic laparoscopic surgical training context. However, several limitations must be considered.

To aid non-medical participants in judging whether the robot's actions were optimal or suboptimal - particularly for subtle criteria such as applied pressure - LEDs were placed along the tissue to be cut. While this LED feedback during suboptimal actions (i.e., LEDs remained lit) vs. optimal actions (LEDs turned off) was essential for participants' understanding and engagement, it may have influenced evoked responses, introducing a perceptual component to the task and complicating comparisons with previous studies. Future research on passive BCIs for robotic training should explore alternative guidance methods and additional information sources, such as haptic feedback representing applied pressure, to reduce reliance on perceptual feedback and better isolate evaluative processes.

Our findings revealed a lateralised error-related brain signature, marked by enhanced left fronto-temporal current sources and right frontal current sinks. This lateralisation has not been previously reported and may represent a novel correlate of performance assessment during robot action monitoring in applied scenarios. Given the rather small sample size in this study, further research should confirm the robustness and replicability of the identified spatiotemporal brain signatures linked to robot action evaluation in realistic scenarios.

Moreover, including medical students or even physicians would enable an assessment of the impact of expert knowledge. Future research should also explore the potential effects of participant fatigue or fluctuating task engagement throughout prolonged video sequences, as this may introduce variability in evoked responses. Addressing these factors could further clarify the robustness of our identified brain signatures.

Precise onset detection of suboptimal actions is challenging in near-naturalistic experiments and is often shaped by subjective observer criteria. Therefore, eye-related measures, such as fixations (Simola et al., 2015; Ladouce et al., 2022) or blinks (Alyan et al., 2023), may provide an ecologically valid approach to further investigate attentional shifts toward significant deviations and the associated evaluative processing.

Combining deviation onset detection through eye-based approaches with findings from temporal and single-trial decoding establishes a foundation for developing passive BCIs to reliably label robot actions for reinforcement-learning-based training paradigms. In our study, the most informative signals were extracted from late evoked responses linked to attentional or evaluative processes. Consequently, BCI algorithms should focus on these late evaluative intervals (beyond 1,000 ms after eye-based deviation detection) to enhance decoding accuracy. However, it is important to note that while late evoked responses are suitable for training robots, the delay of a few hundred milliseconds following error detection may restrict their effectiveness for real-time interventions. Such real-time interventions could provide a safeguard in robot-assisted surgeries. To overcome this limitation, future studies could investigate a multisensory decoding approach that integrates electrophysiological, peripheral-physiological, and

eye-based data, combined with a conservative stop criterion (high sensitivity/true positive rate), to develop a system capable of intervening and eliminating suboptimal robotic actions in real-life surgical scenarios.

The next steps toward BCI-assisted robot training in real-world settings include replicating these findings in (a) dual-task paradigms that simulate collaborative scenarios with individual and shared tasks, and (b) using mobile, dry EEG systems suited for unobtrusive, everyday measurements (e.g., Vukelić et al., 2023).

5 Conclusion

Our study reveals three robust spatiotemporal brain signatures that distinguish between evaluated optimal and suboptimal robotic actions during laparoscopic training. The findings emphasize the critical role of late-stage evaluative brain processes in detecting deviations in robotic performance. Specifically, the left fronto-temporal signature, associated with ERP components such as the P300, LPP, and P600, indicates sustained attentional and evaluative engagement in response to suboptimal actions. Additionally, amplified current sinks in right frontal and mid-occipito-parietal regions, consistent with error-related responses like the oERN and ERN/N400, suggest prediction-based processing of errors and deviations.

By delineating distinct electrophysiological patterns, our results deepen the understanding of the neural mechanisms underpinning mental assessments of robotic performance in near-naturalistic scenarios. These insights hold promise for advancing passive BCIs capable of facilitating real-time, automated evaluations in robotic training and collaborative surgical contexts. The research highlights the role of late-stage electrophysiological responses, linked to attentional and evaluative processes, in detecting significant deviations from optimal robotic actions. Integrating these findings into reinforcement-learning-based training frameworks could reduce reliance on explicit feedback from human instructors, enabling more efficient and intuitive human-centered robotic training systems.

Data availability statement

The datasets presented in this study can be found in an online OSF repository and are accessible upon request at <https://osf.io/6ndsv/>.

Ethics statement

The studies involving humans were approved by the local ethics committee of the University of Tübingen (ID: 827/2020BO1). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

KL: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project

administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. JR: Data curation, Investigation, Validation, Writing – original draft, Writing – review & editing. LK: Conceptualization, Funding acquisition, Methodology, Resources, Software, Validation, Visualization, Writing – review & editing. FM-U: Funding acquisition, Resources, Supervision, Writing – review & editing. MV: Conceptualization, Funding acquisition, Methodology, Resources, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The research was supported by the Vector Stiftung Baden-Württemberg (MINT-Innovationen P-2020-0116) and Ministry of Economic Affairs, Labor, and Tourism Baden-Württemberg (KI-Fortschrittszentrum Lernende Systeme und Kognitive Robotik).

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.

References

- Alyan, E., Wascher, E., Arnau, S., Kaesemann, R., and Reiser, J. E. (2023). Operator state in a workplace simulation modulates eye-blink related EEG activity. *IEEE Trans. Neural Syst. Rehabil. Eng.* 31, 1167–1179. doi: 10.1109/TNSRE.2023.3241962
- Appriou, A., Cichocki, A., and Lotte, F. (2020). Modern machine-learning algorithms: for classifying cognitive and affective states from electroencephalography signals. *IEEE Syst. Man Cybern. Mag.* 6, 29–38. doi: 10.1109/MSMC.2020.2968638
- Aricò, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., and Babiloni, F. (2018). Passive BCI beyond the lab: current trends and future directions. *Physiol. Meas.* 39:08TR02. doi: 10.1088/1361-6579/aad57e
- Balconi, M., and Vitaloni, S. (2014). N400 effect when a semantic anomaly is detected in action representation. a source localization analysis. *J. Clin. Neurophysiol.* 31, 58–64. doi: 10.1097/WNP.0000000000000017
- Barachant, A., Bonnet, S., Congedo, M., and Jutten, C. (2011). Multiclass brain-computer interface classification by Riemannian geometry. *IEEE Trans. Biomed. Eng.* 59, 920–928. doi: 10.1109/TBME.2011.2172210
- Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., Gratton, G., et al. (2005). Strategic control and medial frontal negativity: beyond errors and response conflict. *Psychophysiology* 42, 33–42. doi: 10.1111/j.1469-8986.2005.00258.x
- Bentin, S., Mouchetant-Rostaing, Y., Giard, M.-H., Echallier, J.-F., and Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *J. Cogn. Neurosci.* 11, 235–260. doi: 10.1162/089892999563373
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108:624. doi: 10.1037/0033-295X.108.3.624
- Chaumon, M., Bishop, D. V., and Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *J. Neurosci. Methods* 250, 47–63. doi: 10.1016/j.jneumeth.2015.02.025
- Chavarriaga, R., Biasucci, A., Förster, K., Roggen, D., Tröster, G., and Millán, J. R. (2010). “Adaptation of hybrid human-computer interaction systems using EEG error-related potentials,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (Buenos Aires: IEEE), 4226–4229. doi: 10.1109/IEMBS.2010.5627376
- Congedo, M., Barachant, A., and Bhatia, R. (2017). Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. *Brain-Comput. Interfaces* 4, 155–174. doi: 10.1080/2326263X.2017.1297192
- Cumming, G. (2014). The new statistics: why and how. *Psychol. Sci.*, 25, 7–29. doi: 10.1177/0956797613504966
- Ehrlich, S. K., and Cheng, G. (2019). A feasibility study for validating robot actions using EEG-based error-related potentials. *Int. J. Soc. Robot.* 11, 271–283. doi: 10.1007/s12369-018-0501-8
- Endrass, T., Klawohn, J., Gruetzmann, R., Ischebeck, M., and Kathmann, N. (2012a). Response-related negativities following correct and incorrect responses: evidence from a temporospatial principal component analysis. *Psychophysiology* 49, 733–743. doi: 10.1111/j.1469-8986.2012.01365.x
- Endrass, T., Klawohn, J., Preuss, J., and Kathmann, N. (2012b). Temporospatial dissociation of pe subcomponents for perceived and unperceived errors. *Front. Hum. Neurosci.* 6:178. doi: 10.3389/fnhum.2012.00178
- Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. ii. error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* 78, 447–455. doi: 10.1016/0013-4694(91)90062-9
- Falkenstein, M., Hoormann, J., Christ, S., and Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biol. Psychol.* 51, 87–107. doi: 10.1016/S0301-0511(99)00031-9
- Ferrez, P. W., and Millán, J. D. R. (2005). “You are wrong! automatic detection of interaction errors from brain waves,” in *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI’05* (San Francisco, CA: Morgan Kaufmann Publishers Inc), 1413–1418.
- Ferrez, P. W., and Millán, J. R. (2008). Error-related EEG potentials generated during simulated brain-computer interaction. *IEEE Trans. Biomed. Eng.* 55, 923–929. doi: 10.1109/TBME.2007.908083
- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390. doi: 10.1111/j.1467-9280.1993.tb00586.x

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

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.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnrgo.2025.1535799/full#supplementary-material>

- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2014). MNE software for processing meg and EEG data. *Neuroimage* 86, 446–460. doi: 10.1016/j.neuroimage.2013.10.027
- Groppe, D. M., Urbach, T. P., and Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology* 48, 1711–1725. doi: 10.1111/j.1469-8986.2011.01273.x
- Hajcak, G., and Foti, D. (2020). Significance?... Significance! empirical, methodological, and theoretical connections between the late positive potential and p300 as neural responses to stimulus significance: an integrative review. *Psychophysiology* 57:e13570. doi: 10.1111/psyp.13570
- Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., et al. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* 87, 96–110. doi: 10.1016/j.neuroimage.2013.10.067
- Hipp, J. F., and Siegel, M. (2013). Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. *Front. Hum. Neurosci.* 7:338. doi: 10.3389/fnhum.2013.00338
- Holdgraf, C. R., Rieger, J. W., Micheli, C., Martin, S., Knight, R. T., Theunissen, F. E., et al. (2017). Encoding and decoding models in cognitive electrophysiology. *Front. Neurosci.* 11:61. doi: 10.3389/fnpsy.2017.00061
- Iturrate, I., Chavarriaga, R., Montesano, L., Minguez, J., and Millán, J. R. (2015). Teaching brain-machine interfaces as an alternative paradigm to neuroprosthetics control. *Sci. Rep.* 5:13893. doi: 10.1038/srep13893
- Iturrate, I., Montesano, L., and Minguez, J. (2010). “Robot reinforcement learning using EEG-based reward signals,” in *2010 IEEE International Conference on Robotics and Automation* (Anchorage, AK: IEEE), 4822–4829. doi: 10.1109/ROBOT.2010.5509734
- Kayser, J., and Tenke, C. E. (2015). On the benefits of using surface Laplacian (current source density) methodology in electrophysiology. *Int. J. Psychophysiol.* 97, 171. doi: 10.1016/j.ijpsycho.2015.06.001
- Kim, S. K., Kirchner, E. A., and Kirchner, F. (2020). “Flexible online adaptation of learning strategy using EEG-based reinforcement signals in real-world robotic applications,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)* (Paris: IEEE), 4885–4891. doi: 10.1109/ICRA40945.2020.9197538
- Kim, S. K., Kirchner, E. A., Stefes, A., and Kirchner, F. (2017). Intrinsic interactive reinforcement learning-using error-related potentials for real world human-robot interaction. *Sci. Rep.* 7, 1–16. doi: 10.1038/s41598-017-17682-7
- Koban, L., Pourtois, G., Vocat, R., and Vuilleumier, P. (2010). When your errors make me lose or win: event-related potentials to observed errors of cooperators and competitors. *Soc. Neurosci.* 5, 360–374. doi: 10.1080/17470911003651547
- Kreilinger, A., Neuper, C., and Müller-Putz, G. R. (2012). Error potential detection during continuous movement of an artificial arm controlled by brain-computer interface. *Med. Biol. Eng. Comput.* 50, 223–230. doi: 10.1007/s11517-011-0858-4
- Kriegeskorte, N., and Douglas, P. K. (2019). Interpreting encoding and decoding models. *Curr. Opin. Neurobiol.* 55, 167–179. doi: 10.1016/j.conb.2019.04.002
- Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205. doi: 10.1126/science.7350657
- Ladouce, S., Mustle, M., Ietswaart, M., and Dehais, F. (2022). Capturing cognitive events embedded in the real world using mobile electroencephalography and eye-tracking. *J. Cogn. Neurosci.* 34, 2237–2255. doi: 10.1162/jocn_a_01903
- Lee, T.-W., Girolami, M., and Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput.* 11, 417–441. doi: 10.1162/089976699300016719
- Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., et al. (2018). A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *J. Neural Eng.* 15:031005. doi: 10.1088/1741-2552/aab2f2
- Luo, T.-j., Fan, Y.-c., Lv, J.-t., and Zhou, C. (2018). “Deep reinforcement learning from error-related potentials via an EEG-based brain-computer interface,” in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (Madrid: IEEE), 697–701. doi: 10.1109/BIBM.2018.8621183
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and meg-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024
- Marsicano, G., Bertini, C., and Ronconi, L. (2024). Decoding cognition in neurodevelopmental, psychiatric and neurological conditions with multivariate pattern analysis of EEG data. *Neurosci. Biobehav. Rev.* 164:105795. doi: 10.1016/j.neubiorev.2024.105795
- Moustris, G. P., Hiridis, S. C., Deliparaschos, K. M., and Konstantinidis, K. M. (2011). Evolution of autonomous and semi-autonomous robotic surgical systems: a review of the literature. *Int. J. Med. Robot. Comput. Assist. Surg.* 7, 375–392. doi: 10.1002/rcs.408
- Nguyen, N. D., Nguyen, T., Nahavandi, S., Bhatti, A., and Guest, G. (2019). “Manipulating soft tissues by deep reinforcement learning for autonomous robotic surgery,” in *2019 IEEE International Systems Conference (SysCon)* (Orlando, FL: IEEE), 1–7. doi: 10.1109/SYSCON.2019.8836924
- Nolan, H., Whelan, R., and Reilly, R. B. (2010). Faster: fully automated statistical thresholding for EEG artifact rejection. *J. Neurosci. Methods* 192, 152–162. doi: 10.1016/j.jneumeth.2010.07.015
- Oliveira, F. T., McDonald, J. J., and Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J. Cogn. Neurosci.* 19, 1994–2004. doi: 10.1162/jocn.2007.19.12.1994
- Osa, T., Sugita, N., and Mitsuishi, M. (2014). “Online trajectory planning in dynamic environments for surgical task automation,” in *Robotics: Science and Systems X*, eds. D. Fox, L. E. Kavvaki, and H. Kurniawati (Berkeley, CA: University of California), 1–9. doi: 10.15607/RSS.2014.X.011
- Padrao, G., Gonzalez-Franco, M., Sanchez-Vives, M. V., Slater, M., and Rodriguez-Fornells, A. (2016). Violating body movement semantics: neural signatures of self-generated and external-generated errors. *Neuroimage* 124, 147–156. doi: 10.1016/j.neuroimage.2015.08.022
- Pailing, P. E., and Segalowitz, S. J. (2004). The effects of uncertainty in error monitoring on associated ERPs. *Brain Cogn.* 56, 215–233. doi: 10.1016/j.bandc.2004.06.005
- Pandya, A., Reisner, L. A., King, B., Lucas, N., Composto, A., Klein, M., et al. (2014). A review of camera viewpoint automation in robotic and laparoscopic surgery. *Robotics* 3, 310–329. doi: 10.3390/robotics3030310
- Pavone, E. F., Tieri, G., Rizza, G., Tidoni, E., Grisoni, L., Aglioti, S. M., et al. (2016). Embodying others in immersive virtual reality: electro-cortical signatures of monitoring the errors in the actions of an avatar seen from a first-person perspective. *J. Neurosci.* 36, 268–279. doi: 10.1523/JNEUROSCI.0494-15.2016
- Penaloza, C. I., Mae, Y., Kojima, M., and Arai, T. (2015). Brain signal-based safety measure activation for robotic systems. *Adv. Robot.* 29, 1234–1242. doi: 10.1080/01691864.2015.1057615
- Pernet, C., Latinus, M., Nichols, T., and Rousselle, G. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: a simulation study. *J. Neurosci. Methods* 250, 85–93. doi: 10.1016/j.jneumeth.2014.08.003
- Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72, 184–187. doi: 10.1016/0013-4694(89)90180-6
- Polich, J. (2007). Updating p300: an integrative theory of p3a and p3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Protzak, J., Ihme, K., and Zander, T. O. (2013). “A passive brain-computer interface for supporting gaze-based human-machine interaction,” in *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion: 7th International Conference, UAHCI 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21–26, 2013, Proceedings, Part 1 7* (Cham: Springer), 662–671. doi: 10.1007/978-3-642-39188-0_71
- Richter, F., Orosco, R. K., and Yip, M. C. (2019). Open-sourced reinforcement learning environments for surgical robotics. *arXiv [Preprint]*. arXiv:1903.02090. doi: 10.48550/arXiv.1903.02090
- Ridderinkhof, K. R., Ramautar, J. R., and Wijnen, J. G. (2009). To pe or not to pe: a p3-like ERP component reflecting the processing of response errors. *Psychophysiology* 46, 531–538. doi: 10.1111/j.1469-8986.2009.00790.x
- Rivet, B., Souloumiac, A., Attina, V., and Gibert, G. (2009). Xdawn algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Trans. Biomed. Eng.* 56, 2035–2043. doi: 10.1109/TBME.2009.2012869
- Salazar-Gomez, A. F., DelPreto, J., Gil, S., Guenther, F. H., and Rus, D. (2017). “Correcting robot mistakes in real time using EEG signals,” in *2017 IEEE international conference on robotics and automation (ICRA)* (Singapore: IEEE), 6570–6577. doi: 10.1109/ICRA.2017.7989777
- Sassenhagen, J., Schlesewsky, M., and Bornkessel-Schlesewsky, I. (2014). The p600-as-p3 hypothesis revisited: Single-trial analyses reveal that the late EEG positivity following linguistically deviant material is reaction time aligned. *Brain Lang.* 137, 29–39. doi: 10.1016/j.bandl.2014.07.010
- Scheffers, M. K., and Coles, M. G. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *J. Exp. Psychol.* 26:141. doi: 10.1037//0096-1523.26.1.141
- Scheikl, P. M., Schreiber, N., Haas, C., Freymuth, N., Neumann, G., Lioutikov, R., et al. (2024). Movement primitive diffusion: learning gentle robotic manipulation of deformable objects. *IEEE Robot. Autom. Lett.* 9, 5338–5345. doi: 10.1109/LRA.2024.3382529
- Simola, J., Le Fevre, K., Tormainen, J., and Baccino, T. (2015). Affective processing in natural scene viewing: valence and arousal interactions in eye-fixation-related potentials. *Neuroimage* 106, 21–33. doi: 10.1016/j.neuroimage.2014.11.030
- Somon, B., Campagne, A., Delorme, A., and Berberian, B. (2017). Performance monitoring applied to system supervision. *Front. Hum. Neurosci.* 11:360. doi: 10.3389/fnhum.2017.00360
- Somon, B., Campagne, A., Delorme, A., and Berberian, B. (2019). Evaluation of performance monitoring ERPs through difficulty manipulation in a response-feedback paradigm. *Brain Res.* 1704, 196–206. doi: 10.1016/j.brainres.2018.10.007

- Spüler, M., and Niethammer, C. (2015). Error-related potentials during continuous feedback: using EEG to detect errors of different type and severity. *Front. Hum. Neurosci.* 9:155. doi: 10.3389/fnhum.2015.00155
- Sutton, S., Braren, M., Zubin, J., and John, E. (1965). Evoked-potential correlates of stimulus uncertainty. *Science* 150, 1187–1188. doi: 10.1126/science.150.3700.1187
- Thananjeyan, B., Garg, A., Krishnan, S., Chen, C., Miller, L., Goldberg, K., et al. (2017). "Multilateral surgical pattern cutting in 2d orthotropic gauze with deep reinforcement learning policies for tensioning," *2017 IEEE International Conference on Robotics and Automation (ICRA)* (Singapore: IEEE), 2371–2378. doi: 10.1109/ICRA.2017.7989275
- Ullsperger, M., Fischer, A. G., Nigbur, R., and Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn. Sci.* 18, 259–267. doi: 10.1016/j.tics.2014.02.009
- Van Den Berg, J., Miller, S., Duckworth, D., Hu, H., Wan, A., Fu, X.-Y., et al. (2010). "Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations," in *2010 IEEE International Conference on Robotics and Automation* (Anchorage, AK: IEEE), 2074–2081. doi: 10.1109/ROBOT.2010.5509621
- van Schie, H. T., Mars, R. B., Coles, M. G., and Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nat. Neurosci.* 7, 549–554. doi: 10.1038/nn1239
- Vukelić, M., Bui, M., Vorreuther, A., and Lingelbach, K. (2023). Combining brain-computer interfaces with deep reinforcement learning for robot training: a feasibility study in a simulation environment. *Front. Neuroergonom.* 4:1274730. doi: 10.3389/fnrgo.2023.1274730
- 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. Robot. Res.* 2:1750003. doi: 10.1142/S2424905X17500039