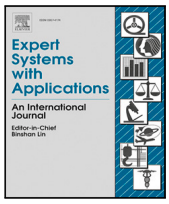
Expert Systems With Applications 275 (2025) 127076

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa



Machine learning classification of motivational states: Insights from EEG analysis of perception and imagery

Tommaso Colafiglio [a,b], Angela Lombardi [b,*], Tommaso Di Noia [b], Maria Luigia Natalia De Bonis [b], Fedelucio Narducci [b], Alice Mado Proverbio [c]

[a] Department of Computer, Control, and Management Engineering, Sapienza University, Rome, Italy
[b] Department of Electrical and Information Engineering, Polytechnic University of Bari, Bari, Italy
[c] Milan Center for Neuroscience, University of Milano-Bicocca, Milan, Italy

ARTICLE INFO

Keywords:
Brain-computer interfaces (BCIs)
Imagery
Perception
Event-related potentials (ERPs)
Motivational states
Machine learning

ABSTRACT

The investigation of neural correlates of mental imagery and perception has been a pivotal area of research in cognitive neuroscience, offering significant insights into how the brain represents both imagined and perceived experiences. While previous studies have successfully identified electrophysiological markers associated with motor and perceptual imagery, the neural signatures of motivational imagery — encompassing desires, needs, and cravings — remain underexplored. This study employs different machine learning classifiers applied to EEG data to classify and compare neural representations of twelve distinct motivational states under perception and imagery conditions. We conducted experiments using 14-channel and 18-channel EEG configurations to capture and analyze the neural responses of participants exposed to various stimuli. Our primary aims were to evaluate classification performance, measured by accuracy, and assess the impact of electrode density on performance. The results indicate that perception conditions generally yield higher accuracy in distinguishing motivational states than imagery conditions. Specifically, primary needs and somatosensory states exhibited strong and clear neural patterns in perception, with a peak accuracy of 88% in the 18-channel setup, while the accuracy for imagined states was more variable. Comparisons between 14-channel and 18-channel configurations revealed that higher electrode density slightly improved performance but was not significantly superior.

1. Introduction

The exploration of the neural correlates of mental imagery and perception has been a pivotal area of research in cognitive neuroscience, yielding significant insights into how the brain represents both imagined and perceived experiences (Alho, Gotsopoulos, & Silvanto, 2023; Dijkstra, 2024; Dijkstra & Fleming, 2023; Motoyama & Hishitani, 2024; Spagna, Hajhajate, Liu, & Bartolomeo, 2021).

Previous studies have provided a foundational understanding of the electrophysiological markers associated with motor imagery (Echtioui, Zouch, Ghorbel, Mhiri, & Hamam, 2024; Morozova et al., 2024), perceptual and cognitive imagery (Syrov, Yakovlev, Nikolaeva, Kaplan, & Lebedev, 2022), and, to a lesser extent, motivational imagery, highlighting their relevance for brain-computer interface (BCI) systems and potential applications in clinical settings for patients with disorders of consciousness (Cai et al., 2022; Leoni, Strada, Tanelli, Brusa, & Proverbio, 2022; Leoni et al., 2021; Mattioli, Porcaro, & Baldassarre, 2022; Proverbio, Tacchini, & Jiang, 2022). Despite these advancements, the comparison between perceptual signals and those generated by imaginative processes, especially within the domain of motivational imagery — encompassing desires, needs, and craves — remains underexplored. This gap in research is particularly poignant given the implications such studies have for understanding consciousness and enhancing communication methods for non-verbal patients, such as those with locked-in syndrome (Kavanagh, Andrade, & May, 2005; Vanhaudenhuyse et al., 2007). Ji et al.'s meta-analysis showed that mental images can evoke cognitive and emotional responses, enabling individuals not only to represent possible scenarios to themselves but also to anticipate the emotional consequences of such choices. In healthy subjects, mental images would also be able to modify physiological parameters, such as heart rate or skin conductance (Ji, Heyes, MacLeod, & Holmes, 2016). Electrophysiological markers of craving, whether related to food or substance use, have been investigated in a few studies. For example, in cases of food craving, studies using EEG have identified specific event-related potentials components that are heightened in response

* Correspondence to: Politecnico di Bari, Dipartimento di Ingegneria Elettrica e dell'Informazione, via E. Orabona 4, 70125 Bari, Italy.
E-mail addresses: tommaso.colafiglio@poliba.it (T. Colafiglio), angela.lombardi@poliba.it (A. Lombardi), tommaso.dinoia@poliba.it (T.D. Noia), marialuigia.debonis@poliba.it (M.L.N. De Bonis), fedelucio.narducci@poliba.it (F. Narducci), mado.proverbio@unimib.it (A.M. Proverbio).

https://doi.org/10.1016/j.eswa.2025.127076
Received 5 September 2024; Received in revised form 24 January 2025; Accepted 25 February 2025
Available online 7 March 2025

to food stimuli, such as the P1, P2, P300, LPP, and N200. These components show increased amplitude for palatable or calorie-dense foods, especially when participants are hungry (Carbine et al., 2018). In addiction-related cravings, research has found that the P300 and LPP are particularly relevant, with increased amplitudes in individuals with pathological eating habits (Carbine et al., 2018). Similarly, in substance cravings, studies have highlighted ERP components such as a heightened N170 in occipito-temporal areas, a reduced N400 in centro-parietal sites, and a delayed P300, all of which are linked to difficulty inhibiting cravings (Habelt, Arvaneh, Bernhardt, & Minev, 2020). These neural markers overlap with those observed during food cravings, suggesting shared brain mechanisms between different types of cravings, as seen in both food and substance use disorders (Rogers, 2017). While there are not BCI studies on motivational or physiological states, several shreds of evidence were provided of effective BCI systems for detecting affective states such as loneliness (Walsh & Schlauch, 2024), fear and satisfaction (Pan, Yang, Qiu, & Huang, 2022; Rodriguez, Del-Valle-Soto, & Gonzalez-Sanchez, 2022) or frustration (Jochumsen, Hougaard, Kristensen, & Knoche, 2022).

In this context, the application of machine learning (ML) methodologies, specifically Support Vector Machines (SVM), to analyze electroencephalogram (EEG) data emerges as a promising approach (Aggarwal & Chugh, 2022; Colafiglio et al., 2023; Kim et al., 2024). ML algorithms offer the potential to dissect and classify the intricate patterns present in EEG signals associated with a wide range of mental states, from physiological needs to motivational desires (Chattopadhyay, Zary, Quek, & Prasad, 2021; Colafiglio et al., 2024; Sairamya, Subathra, & George, 2022; Tang et al., 2024). The integration of ML with EEG analysis represents a novel path forward, enabling the detailed comparison of imaginative and perceptual signals across a spectrum of stimuli (Sharma, Kim, & Gupta, 2022).

While previous studies have provided a foundational understanding of the electrophysiological markers associated with motor and perceptual imagery, the neural signatures of motivational imagery — encompassing desires, needs, and cravings — remain underexplored. This study addresses this gap by employing ML algorithms to classify and compare neural representations of twelve distinct motivational states under both perception and imagery conditions. Unlike prior work focused primarily on motor or perceptual processes, our study uniquely investigates how ML techniques can be used to analyze motivational states, which has implications for enhancing BCI systems. Furthermore, we provide a systematic comparison of neural patterns in perception versus imagination, offering insights into the differential representation of motivational states in the brain, an area that has been largely overlooked in previous research. Important questions central to this investigation include (i) determining the differential neural signatures of imagination and perception, (ii) evaluating the efficacy of ML algorithms in classifying these complex states, and (iii) understanding how these distinctions contribute to our knowledge of brain function and potential BCI applications.

To achieve some of these objectives, our study employed a structured pipeline, beginning with the meticulous collection of EEG data from participants as they were exposed to various stimuli designed to evoke specific motivational and physiological responses. The experiment was structured to capture the neural representations of both imagined and perceived states, categorizing these into one of 12 predefined groups based on the nature of the stimulus. Subsequent analysis utilized state-of-the-art ML techniques to process the EEG data, identifying and comparing the electrophysiological markers associated with each state. This involved extracting features relevant to known ERP components and applying different classification models to discern between imagined and perceived experiences.

Through the lens of ML-enhanced EEG analysis, this study endeavors to advance our comprehension of the neural dynamics underlying motivational imagery and refine the capabilities of BCI technologies. By elucidating the distinct neural signatures of imaginative and perceptual processes, our research contributes to the broader fields of neuroscience and cognitive science, paving the way for innovative therapeutic approaches and enhanced communication strategies for individuals with severe communication impairments.

2. Related work

A BCI aims to acquire brain signals that can be analyzed to control devices or provide feedback to a smart application. Wolpaw, Birbaumer, McFarland, Pfurtscheller, and Vaughan (2002) first proposed the idea that BCI could potentially offer individuals who are fully paralyzed or "locked in" the ability to communicate basic needs to caregivers, operate word processing programs, or control neuroprosthetic devices. BCIs typically work by detecting, interpreting, and translating brain signals into actionable commands for applications that exploit them (Gembler, Stawicki, & Volosyak, 2017). Once the brain signals are captured, they are processed using signal processing algorithms and ML techniques to extract relevant information (Aggarwal & Chugh, 2022). BCIs have a wide range of potential applications, including assistive technologies for individuals with disabilities, rehabilitation for stroke patients, neurofeedback training for cognitive enhancement, gaming and entertainment, and even controlling smart home devices using just brain activity.

The most common types of EEG-based BCI systems deal with motor potentials, slow cortical potentials (SCPs), sensorimotor rhythms, and P300 detection, but also use transient and steady states evoked potentials (Kuś, Duszyk, Milanowski, Łabęcki, Bierzyńska, Radzikowska, Michalska, Żygierewicz, Suffczyński, & Durka, 2013) and EEG frequency analysis (Chang, Wang, Yan, & Liu, 2020). The P300 speller (Hu et al., 2024; Khan et al., 2023; Maslova et al., 2023) involves a grid of symbols, including letters and digits, displayed on a computer screen. The rows and columns are randomly flashed to the user. The user focuses on the character they want to communicate, silently counting the number of times it flashes while eliciting cognitive P300 responses, reflecting their decision-making processes. One issue with this paradigm is that it necessitates the patient's active participation. The patient must be aware of what they want to communicate and make a decision, such as selecting a letter or command. This method is not suitable for patients who are comatose or unable to interact voluntarily with the equipment.

To address this issue, in a recent study by Proverbio et al. (2022), ERPs elicited by 10 stimulus categories across both visual and auditory domains have been analyzed to discern consistent markers for the perception process. By examining the spatio-temporal patterns of brain activity, the authors identified significant voltage peaks and statistically explored their association with specific stimulus categories. Despite the interpretive advantages of an expert BCI system, establishing strict rules can be challenging due to the intricate nature of EEG data. EEG recordings typically encompass multiple channels corresponding to distinct brain regions. Additionally, the morphology and amplitude of brain waves vary among individuals, necessitating a time-intensive fine-tuning process for each new user.

For that reason, ML and deep learning approaches are exploited to identify ERPs automatically (Du, Tang, Jiang, & Tian, 2022). Based on ERP data collected in healthy controls during perceptual and imagery conditions and referred to mental processing of 14 stimulus categories, Proverbio, Tacchini, and Jiang (2023) identified unprecedented electrophysiological markers of imagination without sensory stimulation at BCI purposes. The following peaks were identified at specific scalp sites and latencies during imagination of infants (centroparietal positivity, CPP, and late CPP), human faces (anterior negativity, AN), animals (anterior positivity, AP), music (P300-like), speech (N400-like), affective vocalizations (P2-like) and sensory (visual vs. auditory) modality (PN300). Leoni, Strada, Tanelli, and Proverbio (2023) investigated the development of a Mind Reading Classification Engine (MIRACLE) based on these signals. They showcase the potential of

this engine in enhancing the accuracy of stimuli classification from EEG data. Again, Leoni et al. (2022) presents a deep-learning approach for single-trial stimuli classification using detected P300 waves. This novel method aims to improve the performance of Augmented BCI systems for individuals with limited speech or motor abilities. In another study, Leoni et al. (2021) focuses on automatic stimuli classification from ERP data. The researchers investigate the potential of utilizing ML algorithms to enhance communication via BCIs. Overall, these studies collectively contribute to the advancement of BCI technology by exploring innovative methodologies for stimuli classification and communication enhancement, ultimately aiming to improve the quality of life for individuals with neurological disorders. Previous studies (Llorella, Patow, & Azorín, 2020; Shen, Dwivedi, Majima, Horikawa, & Kamitani, 2019) utilized deep-learning techniques to differentiate between ERPs generated by the imagery of various stimuli. Specifically, Llorella et al. (2020) focused on discerning between ERPs linked to the visualization of homes and human faces with a 68% accuracy rate. In contrast, Shen et al. (2019) employed convolution neural networks and genetic algorithms to distinguish ERPs associated with imagery of dogs, aeroplanes, and houses, achieving a 60% accuracy rate. However, these results may not be sufficient for developing a reliable BCI system due to the limited range of stimulus categories.

A form of explanation is implemented by ranking the features according to their contribution to the prediction. The interesting result is that the rules blindly learned by the algorithm are based on the same features that the neuroscientist would have considered carrying out the stimulus identification based on their expertise. The most used technique for recognizing the presence of the ERP is the Linear Discriminant Analysis (LDA). Zhang et al. (2013) provides an example wherein they utilize spatiotemporal LDA to detect the visual P300, achieving an accuracy rate of 80.0%. A noteworthy aspect of their study is the recognition of the dual nature of the ERP. They emphasize that besides its temporal dynamics, the ERP also exhibits a spatial dimension, both contributing to its recognition. Support Vector Machines are exploited in Rakotomamonjy, Guigue, Mallet, and Alvarado (2005) with an accuracy of 100% in evoked potential identification. However, we must be careful with such high accuracy values since they are clues to overfitting. Through convolutional neural networks (CNN), (Cecotti & Graeser, 2011) achieve 88% accuracy. While neural networks have demonstrated their effectiveness in classifying ERPs since early applications, their not transparent decision-making process has led them to take a backseat compared to more interpretable algorithms like LDA. A further limitation of these approaches is that they require very large datasets that are not always available in this context since, generally, a small number of subjects are involved, given the complexity of the data acquisition process. The imagination of tactile stimulation of a body part as a stimulus is used in Yao, Mrachacz-Kersting, Sheng, Zhu, Farina, and Jiang (2018). The authors implement an LDA-based classifier that can distinguish the different somatosensory cortical patterns evoked by imagining a tactile stimulation to the right hand, left hand, both hands and idle condition. In their study, Zhang, Zhou, and Jiang (2020) show that using a tactile and an auditory stimulus simultaneously increases the implemented system's accuracy. While uncommon, existing literature includes efforts aimed at multi-class stimulus identification. However, the scope remains limited, typically encompassing a pool of five or fewer distinct stimuli. In addition to prior studies focusing on somatosensory stimuli, an illustrative instance is presented in the work by Fan, Wade, Key, Warren, and Sarkar (2017). Utilizing k-nearest neighbors, they successfully recognized five emotional states, achieving an average accuracy of 87.4%. While various BCI studies have effectively demonstrated the ability to classify human intention to communicate, with P300 (Kanoh, Miyamoto, & Yoshinobu, 2011; Pan & Li, 2017; Salvaris & Sepulveda, 2009) or eye movement detection (Zhang, Gao, Zhou, Cheng, & Mao, 2023), or to classify categorical specific content of the imagination (Leoni et al., 2023), there are currently (to our knowledge) no BCI studies on the ability to detect motivational states (such as the desire for food, the feeling of cold or the desire to listen to music). Such studies would advance the possibility of communicating with patients in a coma or altered states of consciousness, such as a vegetative state, in the absence of any awareness on their part. In this regard, Sitaram et al. (2011) demonstrated that it is possible to determine a person's emotional state based on brain activity alone using SVMs and fMRI signals. They also developed a real-time classification system to recognize discrete emotional states such as happiness, disgust, and sadness that were paralleled by subjective reports of the emotions felt. Similarly, Doyle, Kucerovsky, and Ieta (2006) used an SVM-based binary classifier. Their model successfully demonstrated the efficacy of single-trial identification of affective states, but motivational states and craves remain an unexplored realm of the human mind.

3. Problem statement

This study aims to bridge the gap in understanding the electrophysiological markers of motivational states through the use of ML algorithms applied to EEG data. While previous research has successfully utilized EEG signals for motor and perceptual imagery in BCI applications, exploring EEG markers for motivational states such as cravings and desires remains largely unexplored. Our research focuses on the following key research questions (RQs) and aims to address them through a series of carefully designed experiments:

1. RQ1 Intra-condition differences in motivational states: which combinations of motivational states exhibit the most significant differences within each of the two conditions (perception and imagery)? This investigation seeks to identify the specific pairs of motivational states that are most distinguishable from one another independently within the perception and imagery conditions. By analyzing the intra-condition differences, we aim to uncover distinct neural signatures that characterize each motivational state when they are perceived and when they are imagined. This will involve training binary classifiers for each pair of motivational states and evaluating their performance within each condition.
2. RQ2 Inter-condition differences in motivational states: how do the patterns of pairs of motivational states differ the most between the two conditions (perception and imagery)? This investigation compares the neural patterns associated with pairs of motivational states across the perception and imagery conditions. The goal is to understand how the brain's representation of these states changes when transitioning from perception to imagination. By examining inter-condition differences, we aim to identify unique electrophysiological markers that differentiate perceived motivational states from imagined ones. This will involve comparing classifiers for each pair of states across the two conditions.
3. RQ3 How does performance vary between different electrode configurations? This investigation aims to determine the necessity and effectiveness of using high-density electrode configurations compared to configurations with fewer electrodes. The impact of different electrode configurations on classification performance was then assessed through a posteriori comparison of results across configurations. By evaluating the classification performance of different electrode setups, we seek to establish whether additional electrodes significantly enhance the ability to differentiate between motivational states and improve the overall explainability of the models. This will involve comparing the performance of classifiers trained on data from both high-density and reduced-electrode configurations.

To systematically address these research questions, the study was conducted as follows:

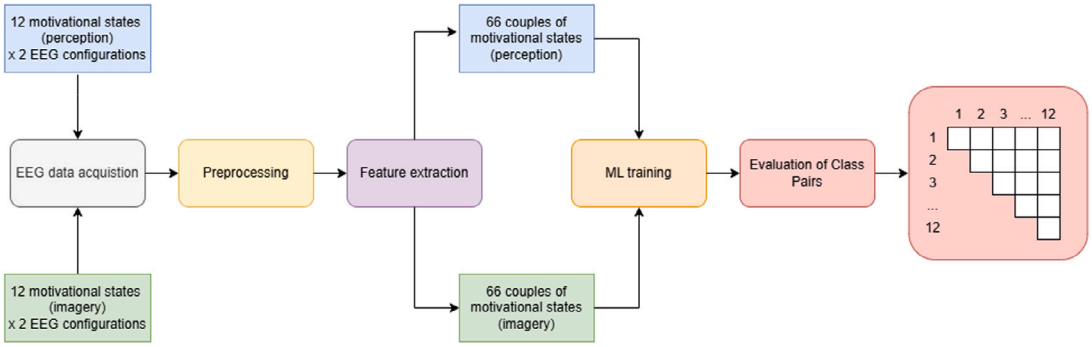


Fig. 1. Overview of the experimental framework used in this study. This framework outlines the key steps, including data collection, preprocessing, feature extraction, and model training and evaluation. Each phase is designed to systematically address the research questions by focusing on different aspects of EEG data related to motivational states.

- Data collection: EEG data were collected from participants exposed to stimuli designed to evoke specific motivational states. The data include both perception and imagery conditions, with stimuli categorized into 12 distinct motivational states. In this study, motivational states referred specifically to cravings and desires linked to four key categories of needs: primary visceral needs (*e.g.*, hunger, linked to the desire for food), somatosensory thermal and pain sensations (*e.g.*, cold, linked to the desire for warmth), affective states (*e.g.*, fear, linked to the desire for reassurance), and secondary needs (*e.g.*, desire to exercise or listen to music). These states were measured through EEG signals to explore their neural signatures.
- Preprocessing: the collected EEG data were preprocessed to remove artifacts and noise, ensuring high-quality signals for subsequent analysis.
- Feature extraction: relevant features were extracted from the preprocessed EEG data, focusing on known ERP components and other significant electrophysiological markers.
- Model training and evaluation:

RQ1 intra-condition analysis: binary classifiers were trained for each pair of motivational states within the perception and imagery conditions separately. Performance metrics were used to identify the most distinguishable pairs within each condition.

RQ2 Inter-conditions analysis: the classifiers were compared across the perception and imagery conditions for each pair of motivational states. Differences in the patterns between the conditions were analyzed.

RQ3 Electrode configuration analysis: the classifiers were trained using data from different electrode configurations to assess the impact of electrode density on classification performance and model explainability.

Fig. 1 depicts an overview of the framework. More details are included in Section 6. Fig. 2 represents a schematic diagram of the RQs.

4. Dataset collection

4.1. Subject information

The data collection process involved testing of 20 healthy students (8 males, 12 females), aged on average 23.20 years (SE = 1.70), right-handed (dominance score = 0.84, SD = 0.17) and schooling = 16.80, SE = 1.58. They had normal or corrected vision, had no current or previous neurological or psychiatric disorders, and were not taking psychotropic drugs or substances that could affect brain activity. All participants were right-handed, as assessed through the Edinburgh inventory questionnaire. They were recruited through the SONA system and received academic credits for their participation. The experiment was conducted in accordance with international ethical standards and was approved by the Ethics Committee of the local University (protocol no: RM-2020-242). Each participant was given written instructions regarding the standardized experimental protocol.

4.2. Experimental protocol

Firstly, a high-density electrode cap was carefully applied to the volunteer's scalp. After electrode application and impedance checking, participants entered a faradized and anechoic chamber and comfortably sat in front of a PC screen where visual stimulation was provided. They were instructed to fixate the center of the screen and completely refrain from moving their body and eyes during the recording sessions. The experiment consisted in the recording of visual evoked responses to pictograms describing specific motivational and emotional states (perception condition), and in the recording of mental responses during recall and imagery of subjective states prompted by a signal, *i.e.* a luminous frame surrounding the screen (imagery condition). The stimuli used in this study were sourced from a validated Motivational Pictionary with 60 colorful plates depicting male and female adults in 12 different motivational states (Proverbio & Pischedda, 2023a). Each motivational state was represented by five variants for better generalizability. The images were rated highly for communicative validity and accuracy. The needs expressed by the stimuli were categorized into primary needs for maintaining homeostasis (feeling hunger, thirst and sleep), somatosensory sensations (feeling hot, cold and pain), affective states (feeling sad, scared or happy), and secondary needs (wanting to move, listen to music and play with friends). The plates depicted characters in specific emotional and motivational states, with a cloud symbol indicating their circumstantial need or desire. The luminance values of the stimuli were consistent across micro-categories and macro-categories. A total of 360 stimuli were shown in a randomized order, with each of the 60 original stimuli being repeated 6 times. The stimuli were presented in 10 runs, each containing 36 stimuli. Each stimulus lasted 2000 ms and was followed by a brief blank screen (ISI) lasting 900 ± 100 ms to prevent retinal after-images. A visual cue, a bright yellow frame on a gray background, appeared for 2000 ms after each stimulus. The Inter Trial Interval (ITI) between stimuli was 150 ± 50 ms. The stimuli were displayed at 640×480 pixels in size, subtending $6°47'16'' \times 4°46'37''$ of visual angle. The PC screen was positioned 114 cm from the observer, and participants were instructed to maintain their gaze on a fixation point throughout the task. A schematic diagram of the procedure used for acquiring EEG data is depicted in Fig. 3. Written instructions prompted participants to imagine the emotion or feeling associated with each stimulus vividly and personally. A brief training session was provided before recording to familiarize participants with the task.

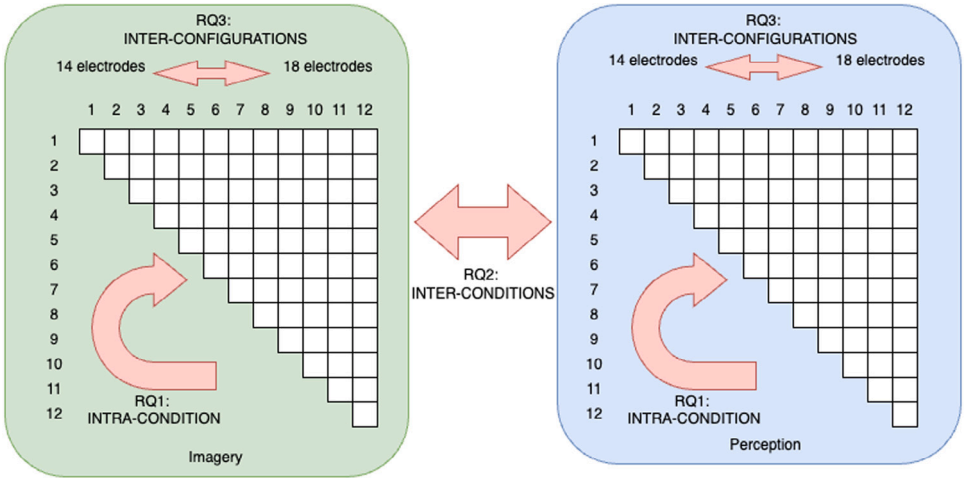


Fig. 2. Schematic representation of the research questions (RQs) addressed in the study. The diagram illustrates the intra-condition analysis (RQ1), inter-conditions analysis (RQ2), and electrode configuration analysis (RQ3), highlighting the relationship between each research question and the corresponding analytical approach used to investigate motivational state distinctions.

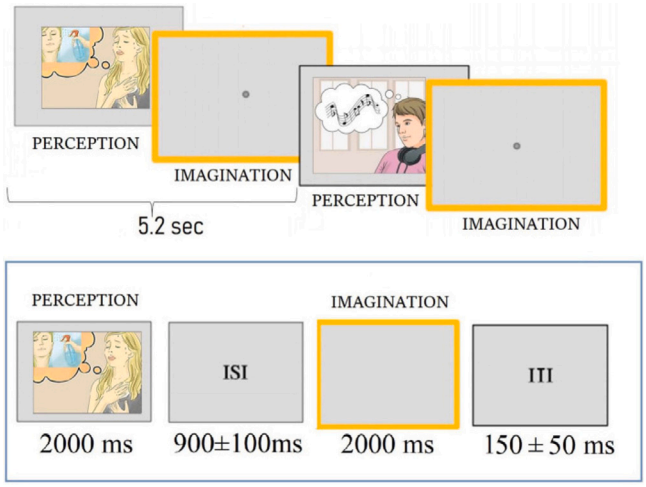


Fig. 3. Schematic diagram of the procedure used for acquiring EEG data. Taken from (Proverbio & Pischedda, 2023a) with their permission. Creative Commons Attribution License (CC BY).

4.3. Experiment process

EEG signals were recorded from 128 electrode sites mounted on elastic caps (Eci-electrocaps) and based on the 10-5 International system (Oostenveld & Praamstra, 2001), with horizontal (hEOG) and vertical (vEOG) electro-oculograms also recorded. EEG recording was carried out using the EEProbe ANT system (Advanced Neuro Technology, Enschede, The Netherlands). Filtering and AD conversion were performed using Synamps digital amplifiers. Averaged mastoids were used as the reference lead, and electrode impedance was maintained below 5 kΩ. A sampling rate of 512 Hz was utilized. The EEG and EOG signals were recorded with the Cognitrace system (ANT Software) and amplified with a bandpass filter (0.16–70 Hz).

5. Preprocessing

EEG artifacts exceeding ±50 μV were automatically removed before averaging, a threshold chosen based on standard EEG research practices to ensure high signal quality while minimizing the impact of muscle and eye movement artifacts. EEG epochs synchronized with stimulus presentation were processed with the EEProbe system (ANT Software). ERPs were averaged offline from 100 ms before to 1200 ms after stimulus onset. The initial 100 ms prior to the stimulus onset ([−100, 0] ms) serves as a baseline period, which is essential for establishing a reference point against which the subsequent activity can be measured. This approach is widely supported in the ERP literature, according to which baseline correction is essential for minimizing the influence of pre-stimulus activity on the measured ERP components (Woodman, 2010). Separate sets of ERP waveforms were generated in response to pictograms and imagery probes in both perception and imagination conditions. ERPs were aligned with the onset of stimuli (pictograms in the perception condition and empty yellow frames in the imagination condition). ERP components were identified and measured relative to the average baseline voltage recorded in the −100/0 ms pre-stimulus interval. For further details, refer to Proverbio and Pischedda (2023b). To provide a clear overview of the EEG data characteristics and acquisition setup, we summarize the most relevant features of the dataset in Table 1.

6. Proposed methodology

6.1. Problem description

This study aims to categorize twelve distinct mental states associated with specific human needs: craving relief from (i) hot, (ii) cold, (iii) pain, (iv) hunger, (v) thirst, (vi) sleep sensations, seeking comfort for (vii) fear, (viii) sadness, (ix) happiness feelings, longing to (x) play, (xi) move,(xii) listen to music. From a machine learning perspective, this is approached as a multi-class classification problem involving a relatively large number of classes. To enhance model development and ensure reliable generalization across subjects, we simplified the task into a series of binary classifications between pairs of mental states. This approach involves defining models capable of distinguishing between each possible pair of mental states. The number of possible pair combinations can be calculated using the binomial coefficient formula:

$$C(n,k) = \frac{k! \cdot (n-k)!}{n!}, \tag{1}$$

where n is the total number of elements (mental states, *i.e.*, 12), and k is the combination size (in this case, 2). Using this formula, the number of pair combinations is $C(12,2) = 66$. Consequently, we conducted 132 experiments in total: 66 for the imagery condition and 66 for the perception condition. For each pair of mental states, the classifiers were trained to differentiate between the two states. The overall process is illustrated in Figs. 1 and 2.

Table 1
Summary of the main features of the captured EEG data.

Feature	Description
Participants	20 healthy right-handed students (8 males, 12 females), average age 23.2 years
Electrode Configuration	EEG signals were recorded from 128 electrode sites mounted on elastic caps (Eci-electrocaps) and based on the 10-5 International system
Sampling Rate	512 Hz
Data Acquisition System	EEProbe ANT system, with Synamps digital amplifiers
Preprocessing	EEG artifacts exceeding ±50 μV automatically removed
Baseline Correction	Baseline from −100 ms to stimulus onset (0 ms)
Stimuli Types	360 total stimuli: 12 motivational states, each with 5 variations
Conditions	Perception and imagery
Session Duration	Each stimulus shown for 2000 ms, with ISI of 900 ±100 ms; 10 runs of 36 stimuli each
Artifact Rejection	Exclusion of artifacts with amplitude beyond ±50 μV
EEG Recording Setup	Impedance below 5 kΩ; referenced to averaged mastoids

6.2. Electrode subselection

The initial data structure includes 126 EEG channels, encompassing the Mastoids, Vertical, and Horizontal EOG channels. We aimed to reduce the number of channels required for the BCI domain. For this purpose, we considered the channel configuration of a well-known BCI, such as the Emotiv headset.[1] The initial EEG setup used the 10-5 International System, an extension of the 10-20 System, to accommodate a higher density of 128 electrode positions (Oostenveld & Praamstra, 2001). For the Emotiv-based configuration with 14 electrodes, we adopted the 10-20 international system to maintain consistency with standard low-density setups commonly used in BCI applications. Fig. 4 illustrates the electrode positions according to the 10-20 international system. Additionally, we explored the feasibility of a BCI with four extra electrodes positioned along the head's centerline at Fz, Cz, Pz, and Oz. Fig. 5 shows the placement of these additional electrodes. It should be noted that a BCI version with these four additional electrodes is not currently available on the market. Topographical electrodes distribution for ML classification of ERP is capable of addressing some of the areas of interest related to our task. We specify that the electrodes with the black circle inside represent the reference.

Hence, two potential EEG caps were evaluated with 14- and 18-channel configurations. To accommodate the BCIs considered, we created two new data structures with their respective electrode configurations using the MNE Python framework (Gramfort et al., 2013). Once the data structures with the appropriate number of channels were established, we extracted the epochs of the EEG signal for each class of target stimuli, excluding the baseline segment of −0.100 ms. The epochs considered start at 0 ms and end at 1199 ms. The final data structure consists of 12 stimuli, 20 subjects, and 2 types of BCI cap configurations.

6.3. Time sliding estimator

We adopted an approach inspired by fMRI techniques to categorize the twelve emotional states. This method trains a classifier for each sample to identify the area of maximum separability between two classes. We evaluated three machine learning models: SVM, Random Forest (RF), and a feedforward Neural Network (NN). The tuning parameters for cross-validation are detailed in the tables provided in the Supplementary Material file (Section S1). This comparison aims to ensure a more comprehensive evaluation of the most suitable model for the task.

A grid search was conducted to determine the optimal model parameters for each classifier. These parameter variations do not imply a change in the classifier type, and they only optimize the performance of each model for the dataset. Notably, grid searches were performed on every possible combination of the 12 classes, resulting in a total of

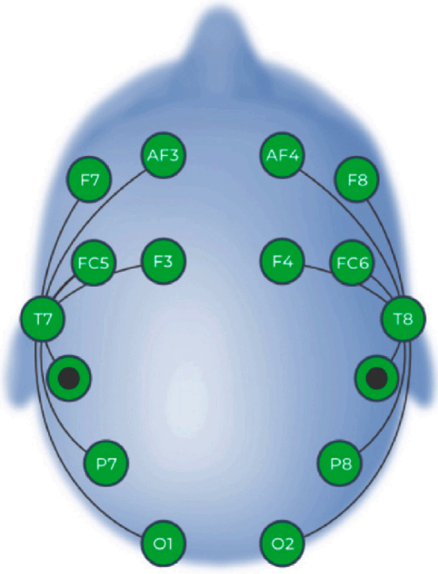


Fig. 4. Electrode placement according to the 10-20 international system used in this study. This configuration illustrates the standard electrode positions employed for the EEG-based BCI system Emotiv headset device. Electrodes with a black circle inside indicate CMS/DRL references at P3/P4 positions utilized during signal acquisition.

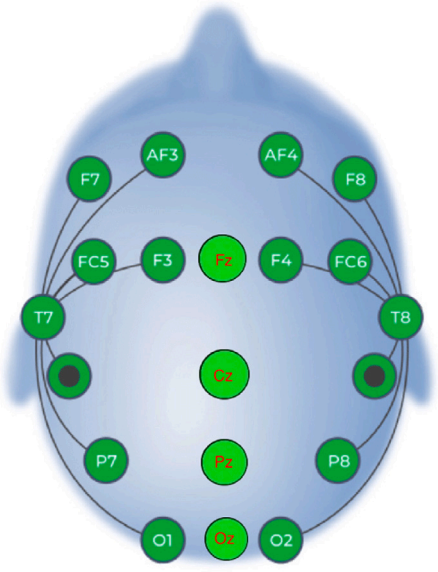


Fig. 5. Electrode placement with the inclusion of four additional electrodes at Fz, Cz, Pz, and Oz along the head's centerline. This configuration aims to enhance the BCI's capability for ERP-based ML classification tasks. The electrodes marked with a black circle indicate CMS/DRL references at P3/P4 positions.

[1] https://www.emotiv.com/products/epoc-x

132 combinations. All experiments were conducted in Python using the scikit-learn framework.[2]

To evaluate how well the models generalize to a single independent test subject, the Leave-One-Subject-Out (LOSO) cross-validation strategy was employed. This technique, widely recommended in biometric data processing, involves excluding one entire subject from the dataset for testing while using the remaining subjects for training (Kunjan et al., 2021; Lombardi et al., 2022). This process is repeated for each subject in the dataset. Specifically, the dataset is divided into 19 training subjects and one independent test subject. Each model in the grid search is tested on the test subject, yielding a prediction score. This procedure produces 20 top models, each tested on a different test subject.

The best model for each subject was selected based on the highest accuracy scores obtained during the grid search. The results for each classification pair were averaged across the 20 LOSO runs, where each result represents the best-performing model identified for that pair. This approach allows us to evaluate reliable generalization across all subjects without treating the 132×20 cases as distinct executions. The classifiers were statistically compared for each condition and electrode count using one-way ANOVA to assess differences in performance, followed by Tukey's HSD post-hoc tests (Tukey, 1949) to identify significant pairwise differences in the distributions of F1 scores obtained from all pairs of motivational states.

7. Experimental evaluation

7.1. Evaluation of final predictions

The final prediction of each model is evaluated using a frequency-based metric.

We define the following notations:

- S_n: number of EEG samples;
- F_m: number of final models.

Let M_i denote the i-th classification model, where $i \in \{1, 2, \ldots, Fm\}$.

For each model M_i, each prediction p_{ij} can be either 0 or 1, where $j \in \{1, 2, \ldots, Sn\}$.

Define n_{i0} and n_{i1} as the number of times model M_i predicts 0 and 1, respectively. These can be calculated as:

$$n_{i0} = \sum_{j=1}^{Sn} 1\{p_{ij} = 0\} \tag{2}$$

$$n_{i1} = \sum_{j=1}^{Sn} 1\{p_{ij} = 1\} \tag{3}$$

Where $1\{\cdot\}$ is the indicator function that returns 1 if the condition is true, otherwise 0.

Each model M_i predicts two classes. Let C_{i0} denote the class predicting "0" and C_{i1} denote the class predicting "1". The final prediction class for each model M_i is given by:

$$C_i = \begin{cases} C_{i0} & \text{if } n_{i0} > n_{i1} \\ C_{i1} & \text{if } n_{i0} < n_{i1} \end{cases} \tag{4}$$

7.2. Performance metrics

The performance of our models was evaluated using several key metrics: accuracy, precision, recall, specificity, and F1 score. The specific methods for calculating these metrics are detailed below.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{5}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{if} \quad (\text{TP} + \text{FP}) \neq 0 \quad \text{else} \quad 0 \tag{6}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad \text{if} \quad (\text{TP} + \text{FN}) \neq 0 \quad \text{else} \quad 0 \tag{7}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad \text{if} \quad (\text{TN} + \text{FP}) \neq 0 \quad \text{else} \quad 0 \tag{8}$$

$$\text{F1 Score} = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}} \quad \text{if} \quad (\text{Precision} + \text{Recall}) \neq 0 \quad \text{else} \quad 0 \tag{9}$$

Where:

- TP (True Positive): the number of instances correctly classified as the positive class for each binary comparison.
- TN (True Negative): the number of instances correctly classified as the negative class for each binary comparison.
- FP (False Positive): the number of instances incorrectly classified as the positive class for each binary comparison.
- FN (False Negative): the number of instances incorrectly classified as the negative class for each binary comparison.

For each binary classification, the positive class corresponds to one specific motivational state under analysis, while the negative class corresponds to the other motivational state in the pairwise comparison. For instance, if the binary classification is between the motivational states of "hunger" and "thirst", "hunger" could be designated as the positive class and "thirst" as the negative class. This designation is applied consistently across the different pairwise comparisons to maintain uniformity in calculating the performance metrics.

8. Results

All results reported in this section are based on the SVM classifier, which demonstrated superior performance compared to the other tested models. Details of the comparative evaluation, including results for RF and the NN, are provided in the Supplementary Material File (Section S2 and S3). The matrices in Figs. 6, 7, 8 and 9 report the F1 scores resulting from the classifiers trained to discriminate each pair of motivational states for both perceived and imagined states and for both EEG configurations. The reported F1 score is calculated after completing the full LOSO cross-validation process. For each classification pair, predicted labels from all 20 LOSO iterations are aggregated, and the F1 score is then computed based on these combined predictions, providing a comprehensive measure of model performance across all subjects.

For each matrix, the extracted values were used to plot the probability density distributions for the four conditions: 14 electrodes perception, 14 electrodes imagery, 18 electrodes perception, and 18 electrodes imagery. The resulting plot is shown in Fig. 10 and illustrates the performance distributions for each configuration and task.

8.1. Overall statistical comparison

To statistically analyze the differences between the conditions, we performed t-tests for the four comparisons. To ensure robustness given multiple comparisons, we applied a Bonferroni correction, setting $\alpha = 0.0125$ (0.05/4) as the significance threshold. The results remain statistically significant where noted:

- 14 Electrodes Imagery vs 14 Electrodes Perception: the t-test result ($t = -6.06$, $p < 0.0001$) indicates a significant difference, with perception data yielding higher accuracy than imagery signals.
- 18 Electrodes Imagery vs 18 Electrodes Perception: the t-test result ($t = -7.01$, $p < 0.0001$) also shows a significant difference, with perception outperforming imagery even with increased electrode density.
- 14 Electrodes Imagery vs 18 Electrodes Imagery: the t-test result ($t = 0.27$, $p = 0.78$) indicates no significant difference, suggesting that increasing electrode density does not significantly improve the accuracy for imagery tasks.

[2] https://scikit-learn.org/1.6/index.html

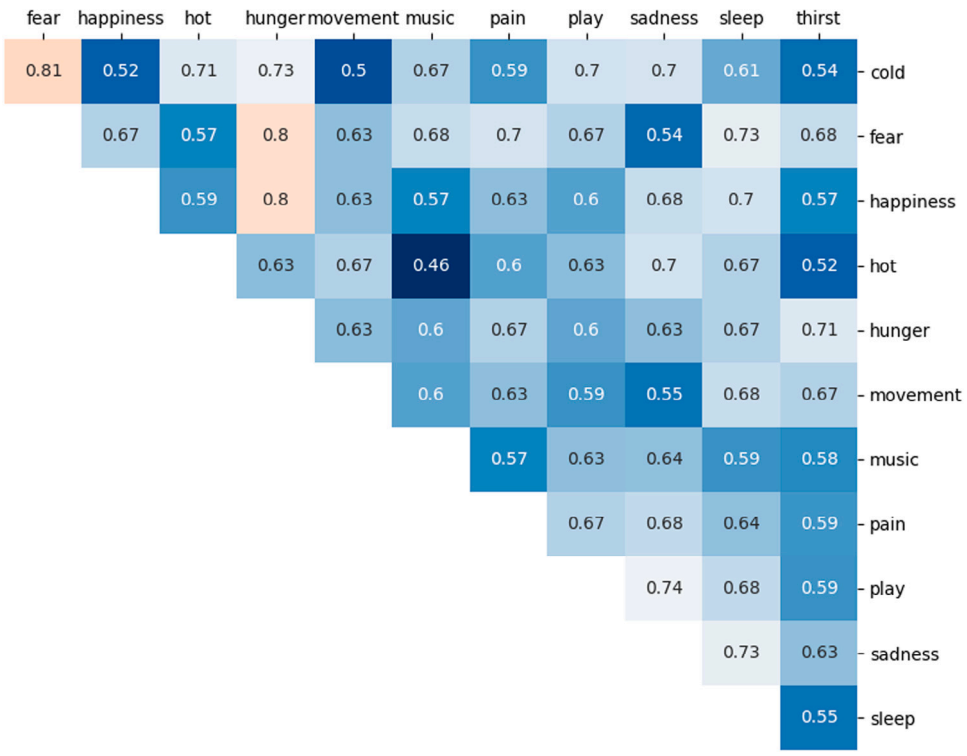


Fig. 6. F1 scores resulting from the classifiers for the imagery condition with 14 channels configuration.

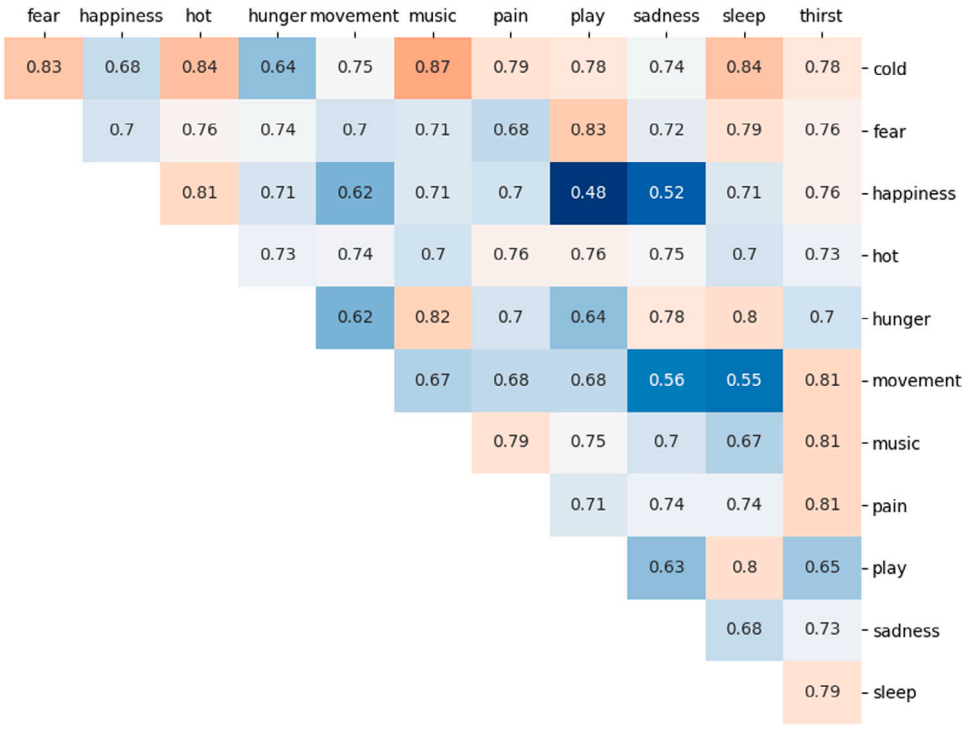


Fig. 7. F1 scores resulting from the classifiers for the perception condition with 14 channels configuration.

- 14 Electrodes Perception vs 18 Electrodes Perception: the t-test result ($t = -0.49$, $p = 0.61$) shows no significant difference, implying that the performance remains stable regardless of electrode density in perception tasks.

8.2. Best performing pairs of states

The performance for the best-performing pairs of states for each condition and configuration are reported in Figs. 11, 12, 13 and 14.

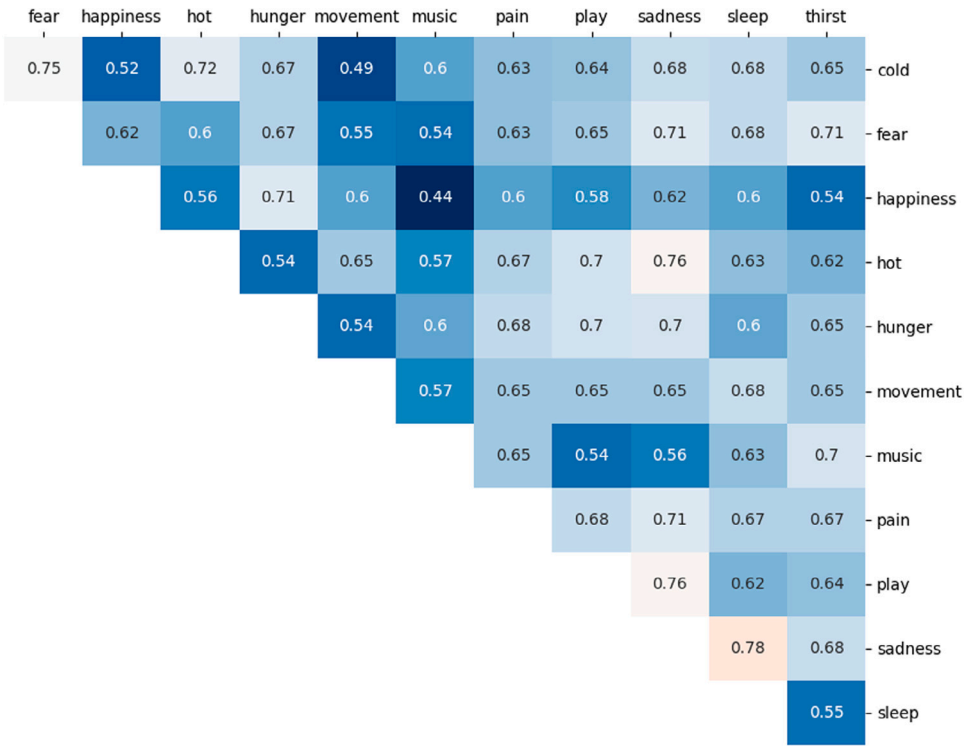


Fig. 8. F1 scores resulting from the classifiers for the imagery condition with 18 channels configuration.

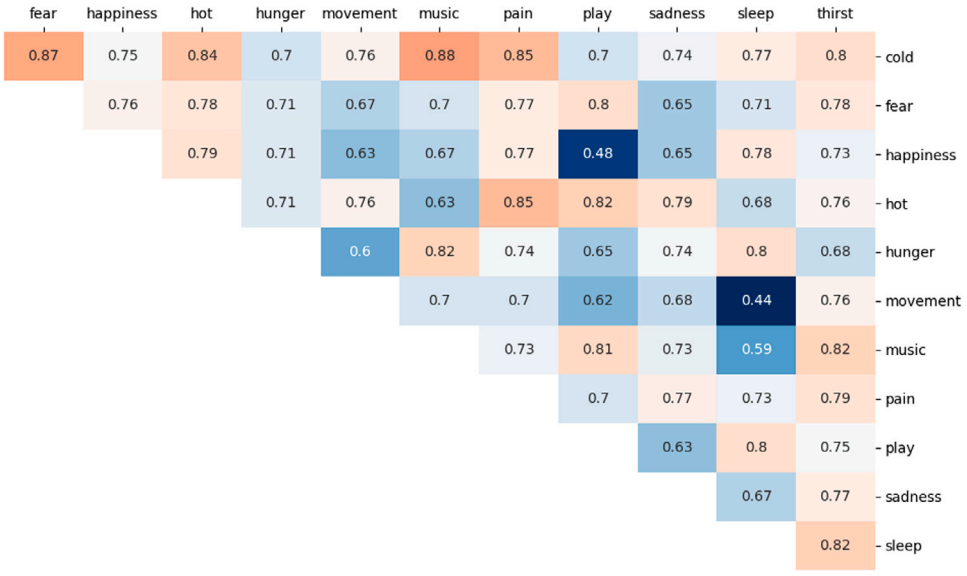


Fig. 9. F1 scores resulting from the classifiers for the perception condition with 18 channels configuration.

Table 2 summarizes the accuracy performance for each pair of states grouped into the four macro-categories.

8.2.1. RQ1: Intra-condition analysis

Imagery (14 Channels)

In the imagery condition using a 14-channel configuration, the highest accuracy is observed for the pair 'cold_fear' with an accuracy of 0.82. The pair 'happiness_hunger' also maintains a high accuracy of 0.80. However, the overall accuracy for imagined states is more variable, with some pairs such as 'thirst_pain' and 'music_hot' showing moderate accuracies of 0.65 and 0.47, respectively. This variability highlights the challenge of vividly imagining certain states compared to perceiving them.

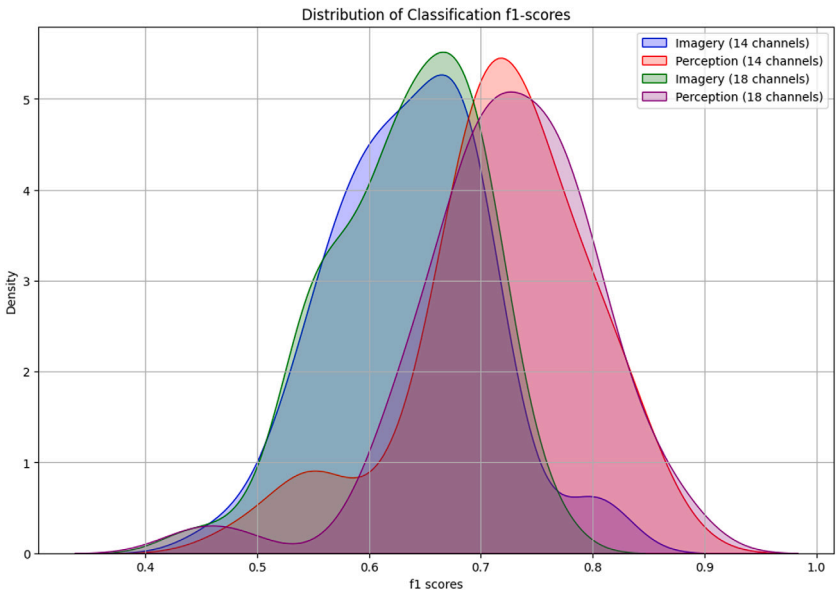


Fig. 10. Probability density distributions of the F1 scores resulting from each couple of classifiers trained to discriminate between each pair of states (both perceived and imagined and for both electrode configurations).

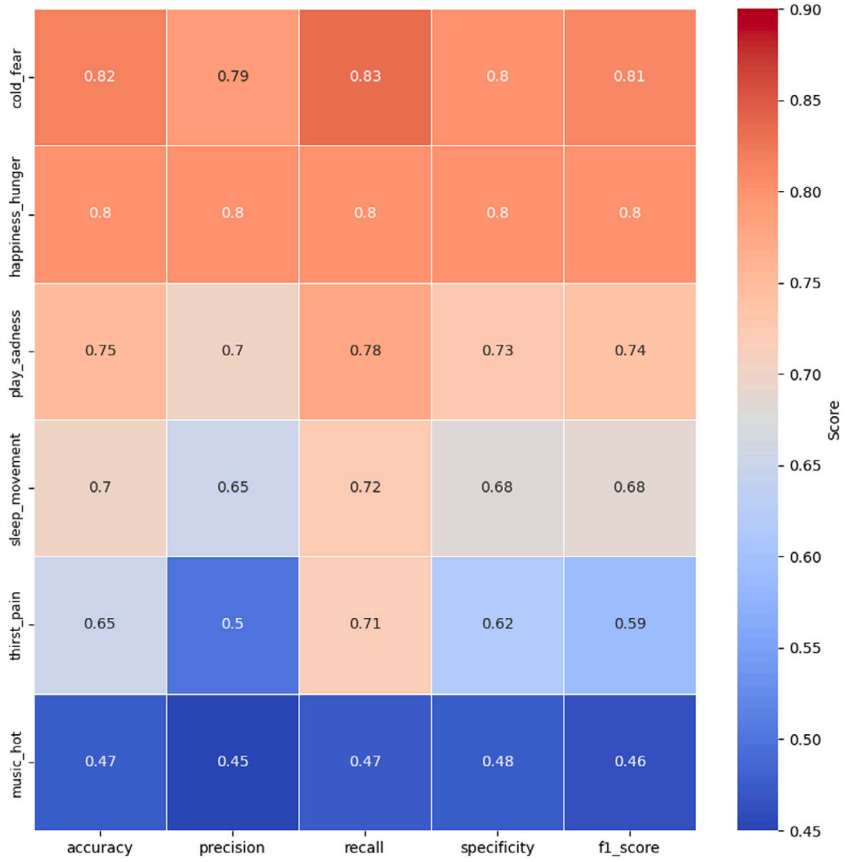


Fig. 11. Performance resulting from the best pairs of classifiers for the imagery stimuli for 14 channels.

Imagery (18 Channels)

In the imagery condition using an 18-channel configuration, the highest accuracy is observed for the pairs 'sadness_sleep' and 'cold_fear' with an accuracy of 0.75. However, similar to the 14-channel configuration, the accuracy for imagined states shows variability, with pairs such as 'hot_pain' and 'movement_play' showing accuracies below 0.7. This indicates consistent challenges in the vivid imagination of these states.

Perception (14 Channels)

The perception condition with a 14-channel EEG configuration shows robust neural pattern detection. The highest accuracy is observed for the pair 'music_cold' with an accuracy of 0.88. This indicates that the neural responses for these perceived states are distinctly

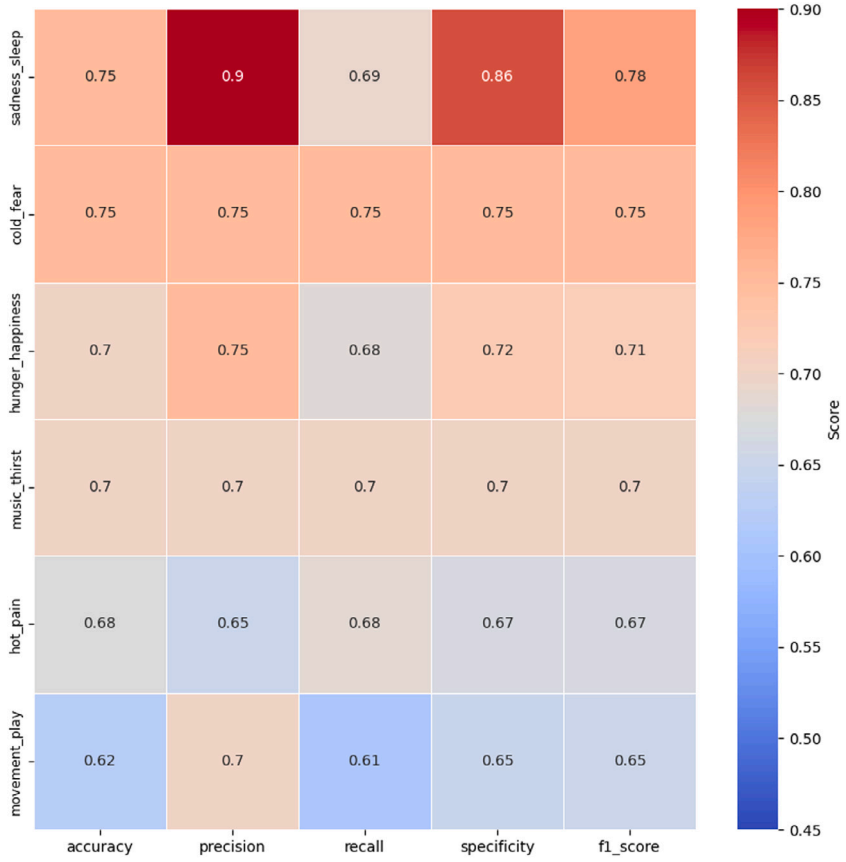


Fig. 12. Performance resulting from the best pairs of classifiers for the imagery stimuli for 18 channels.

Table 2
Accuracy of motivational states for perception and imagery conditions with different EEG configurations. The table displays only the performance results from the best pairs of classifiers for the two types of stimuli (perception and imagery) and the two EEG configurations. Blank spaces indicate configurations or stimulus pairs that did not yield top-performance results.

Category Pair	State Pair	Perception (14)	Imagery (14)	Perception (18)	Imagery (18)
Primary Needs - Primary Needs	hunger_sleep	0.80			
	thirst_sleep			0.82	
Primary Needs - Somato-Sensory	thirst_pain	0.80	0.65		
Somato-Sensory - Somato-Sensory	hot_pain			0.85	0.68
Primary Needs - Affective States	hunger_sadness			0.72	
	hunger_happiness		0.80	0.70	0.70
	sleep_sadness				0.75
Primary Needs - Secondary Needs	sleep_movement		0.70		
	thirst_music				0.70
Somato-Sensory - Affective States	cold_fear		0.82		0.75
	hot_happiness	0.80			
Somato-Sensory - Secondary Needs	hot_music		0.47		
	music_cold	0.88		0.88	
Affective States - Secondary Needs	happiness_movement			0.65	
	fear_play	0.82		0.80	
	sadness_play		0.75		
	sadness_movement	0.65			
Secondary Needs - Secondary Needs	movement_play				0.62

represented. All the other pairs of states show high accuracies of 0.80, suggesting strong neural activations for these states.

Perception (18 Channels)

The perception condition with an 18-channel EEG configuration shows similar robust performance. The highest accuracy is observed for the pair 'cold_music' with an accuracy of 0.88, indicating clear neural activations for these perceived states. All the other pairs show high accuracies between 0.72 and 0.85, except for the pair 'movement_happiness', which achieves an accuracy of 0.65, suggesting that the neural patterns for these states are less distinctly represented in the perception condition.

8.2.2. RQ2: Inter-conditions analysis

Perception (14) vs Imagery (14)

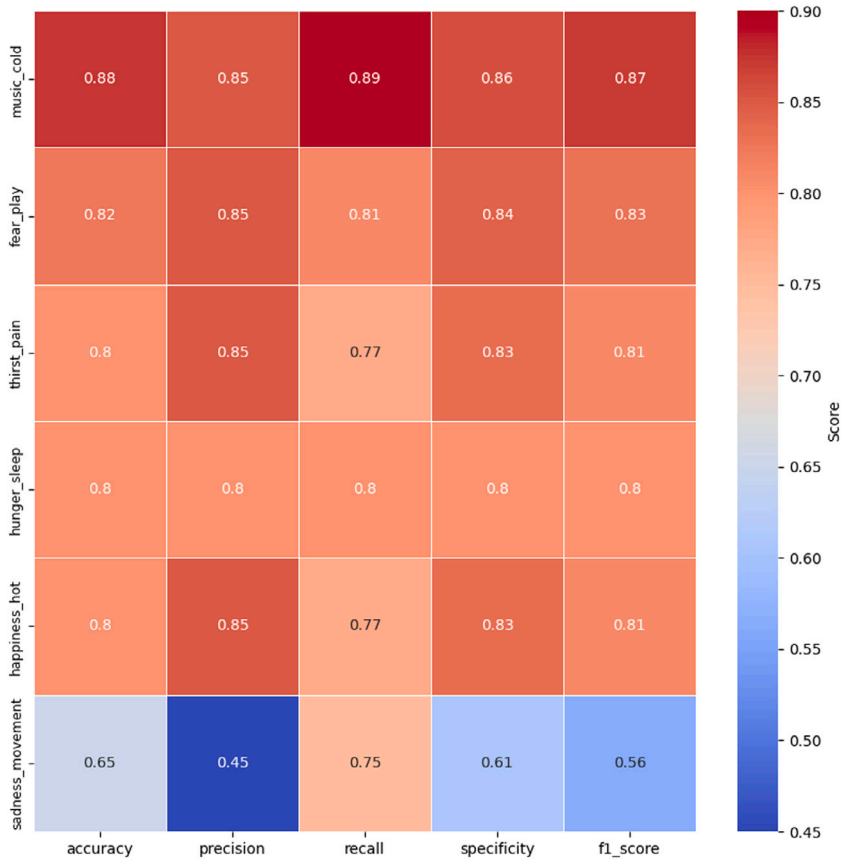


Fig. 13. Performance resulting from the best pairs of classifiers for the perception stimuli for 14 channels.

Comparing the perception and imagery conditions with 14-channel EEG configurations, significant differences emerge. The perception condition generally yields higher accuracy, suggesting clearer and more robust neural patterns for perceived states.

Perception (18) vs Imagery (18)

When using the 18-channel EEG setup, the perception condition continues to show higher accuracy compared to the imagery condition.

8.2.3. RQ3: Inter electrode configuration analysis

Perception (14) vs Perception (18) Comparing the 14-channel and 18-channel configurations under the perception condition reveals that increasing electrode density does not significantly enhance performance. The pair 'cold_music' shows consistent accuracy in both configurations (0.88 for both configurations).

Imagery (14) vs Imagery (18) It is interesting to note that while the 14 and 18-electrode configurations do not show significant differences in overall comparisons for each condition, a focused comparison of the best classifiers for the imagery condition reveals that the 18-electrode configuration performs better in discriminating secondary needs, such as listening to music.

9. Discussion

This study examines the neural representations of various motivational states under different conditions and EEG configurations. By analyzing the data within and across conditions, as well as between different EEG setups, we gain a comprehensive understanding of how the brain processes both perceived and imagined experiences.

9.1. RQ1 intra-condition analysis: perception (14 channels)

The perception condition using a 14-channel EEG configuration demonstrates a high level of accuracy in distinguishing between various motivational states. For primary needs, the combination of hunger and sleep achieved a notable accuracy of 0.80, indicating that the neural patterns associated with these states are distinctly represented when perceived. This suggests that primary physiological needs, such as hunger and sleep, elicit strong and clear neural responses that are easily detectable with the EEG setup.

The pair of thirst and pain, which bridges primary needs and somato-sensory experiences, also showed high accuracy (0.80). This consistency across different types of needs suggests that the EEG configuration effectively captures the neural activations associated with both physiological and sensory states.

The highest accuracy was observed for the somato-sensory and secondary needs pair, cold and music, with a score of 0.88. This finding underscores the robust neural patterns elicited by the perception of sensory experiences combined with secondary needs, such as enjoyment of music.

9.2. RQ1 intra-condition analysis: imagery (14 channels)

In the imagery condition using a 14-channel EEG configuration, the neural patterns associated with motivational states appear less distinct than in the perception condition. For the primary needs and somato-sensory pair, thirst and pain, the accuracy drops to 0.65, indicating that the neural responses for imagined states are less distinct. This

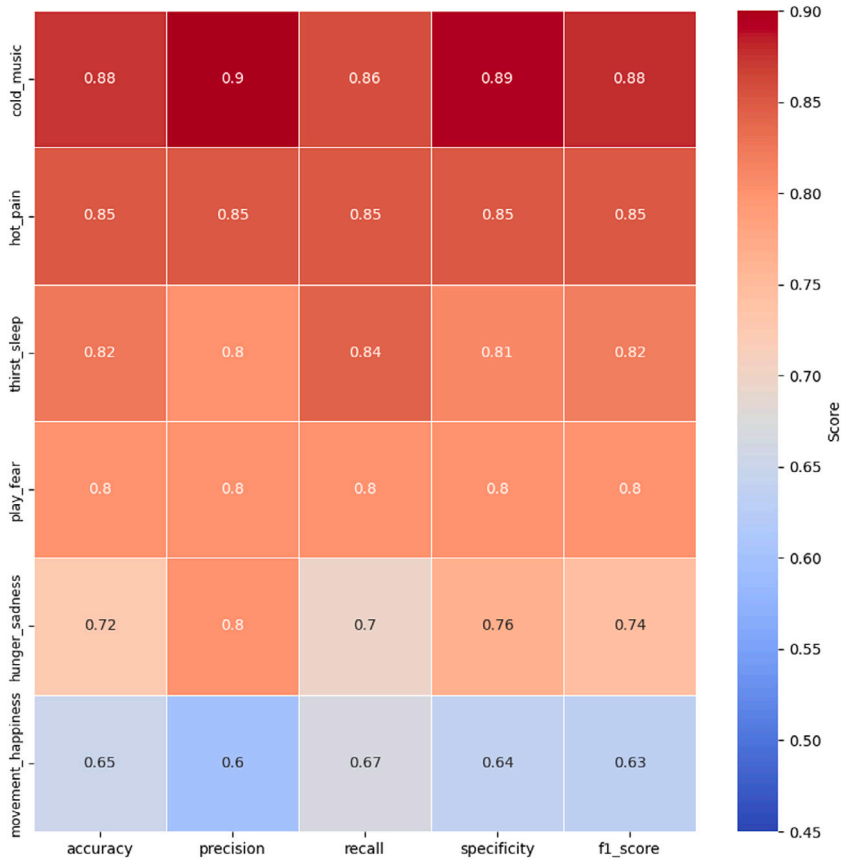


Fig. 14. Performance resulting from the best pairs of classifiers for the perception stimuli for 18 channels.

suggests a challenge in vividly imagining these states, which leads to more variable neural activations.

The pair of sleep and movement, which combines primary needs with secondary needs, shows an accuracy of 0.70, reflecting relatively reliable neural responses for these imagined states.

The somato-sensory and affective states pair, cold and fear, demonstrates high accuracy (0.82), suggesting that some imagined states can still generate strong neural patterns. However, the pair of hot and music shows low accuracy (0.47), highlighting the difficulty in vividly imagining this combination.

9.3. RQ2 inter-conditions analysis: perception (14) vs imagery (14)

Comparing the perception and imagery conditions with 14-channel EEG, several significant differences emerge. The perception condition generally yields higher accuracy, suggesting clearer and more robust neural patterns for perceived states.

The pair 'thirst_pain', which spans primary needs and somato-sensory experiences, shows higher accuracy in perception (0.80) compared to imagery (0.65). This suggests clearer neural activations for perceived states, indicating that the brain's response to actual sensory input is more distinct than to imagined input.

The somato-sensory and secondary needs pair maintains the highest accuracy in perception (0.88), reflecting more consistent and robust neural patterns than imagery.

This pattern of results is similar to what found (Lee, Jang, & Jun, 2022), although with generally lower accuracy performance (yielding approximately 57% (max 63.6%) for a 3-category classification in visual perception and approximately 46% (max 71%) accuracy for a 3-category classification in visual imagery). The difference in the vividness of neural signals generated by perceptual vs. imaginative states is well-known to neuroscientists. It has been explained by invoking the involvement of different neural mechanisms, the absence of precise sensory input signals in imagery, and the different directionality of the signals (from anterior to posterior regions of the brain, rather than vice versa) in the two cases (Dijkstra, Bosch, & van Gerven, 2017; Fulford et al., 2018; Koenig-Robert & Pearson, 2021). In our study, statistical analyses confirmed significant performance differences between perception and imagery conditions, with perception yielding consistently higher classification accuracy across electrode configurations ($p < 0.0001$, Bonferroni-corrected).This finding aligns with the existing literature and further highlights how perceptual states benefit from direct sensory input, which activates early sensory areas in the brain, such as the primary visual, auditory, and somatosensory cortices. Imagination, by contrast, relies more on higher-order regions, particularly the fronto-parietal network, to generate internal images without external stimuli.Consistent with previous findings, perception generally involves more straightforward sensory processing, resulting in stronger neural signals than the more subjective and variable signals associated with imagination, which integrates memory and abstract thought. Interestingly, in the study by Proverbio and Pischedda (2023a), in which N400 components were quantified in the 400–600 ms latency range at anterior/frontal, prefrontal, and fronto/central sites, ERPs were larger in the perception condition than in the imagination condition (except for somatosensory sensations). Consistent evidence for smaller potentials during imagination than perception has been reported in previous ERP studies of perceived and imagined emotional and food photographs (Marmolejo-Ramos,

Hellemans, Comeau, Heenan, Faulkner, Abizaid, & D'Angiulli, 2015)) and perceived or imagined auditory and visual objects such as words, speech, or music (Proverbio & Pischedda, 2023b).

9.4. RQ2 inter-conditions analysis: perception (18) vs imagery (18)

When comparing perception and imagery with the 18-channel EEG setup, we observe several key differences. The somato-sensory pair 'hot_pain' shows high accuracy in perception (0.85) compared to lower accuracy in imagery (0.68), suggesting that sensory experiences are more distinctly captured during perception. This indicates that while some imagined states generate clear neural patterns, the actual sensory input results in more robust neural activations.

For affective states, the pair 'play_fear' shows consistent accuracy in perception (0.80), indicating stable neural patterns for perceived states. The pair 'sleep_sadness' in the imagery condition achieves the highest accuracy (0.75), suggesting that some emotional states, when vividly imagined, can elicit strong and distinct neural responses. Moreover, the states involving movement consistently show the worst performance in both conditions, suggesting that this state is particularly challenging to distinguish from other motivational states.

9.5. RQ3 electrode configuration analysis: perception (14) vs perception (18)

In this study, the 18-electrode configuration was included to test the hypothesis — supported by literature — that additional electrodes along the central midline might improve the detection of signal peaks relevant to motivational state classification. Indeed, Leoni et al. (2023) have demonstrated that the most important channels for the classification of imagery datasets were focussed over the midline and central scalp areas (namely,central, dorsolateral prefrontal/frontocentral and centro/parietal sites: C1, C2, and Cz, CCP1 h and CCP2 h, FC2, and FFC1 h, FFC2 h, and FFC4 h). This pattern also aligns with electrophysiological findings reported by Proverbio and Pischedda (2023a). It is possible that midline sites best capture brain activity coming from areas devoted to imagery, working memory and long term memory, that is the medial prefrontal cortex and hippocampus (Ludowig, Bien, Elger, & Rosburg, 2010). Since no commercial EEG headset currently supports this exact configuration, we tested it as a hypothetical setup to explore its potential benefits. While the average classification performance did not improve significantly with 18 electrodes, the results from the best-performing classifier pairs for both imagery and perception stimuli showed a slight improvement over the 14-electrode setup.

For affective states, the pair 'happiness_fear' shows consistent accuracy (0.82 in 14 channels and 0.80 in 18 channels), highlighting stable neural patterns. The pair 'cold_fear' shows consistent accuracy with 14 channels (0.82), slightly lower with 18 channels (0.75). Overall differences in performance were not substantial, suggesting that both configurations effectively capture essential neural patterns.

The better classification accuracy for cold and happiness states strongly fits with the nature of electrophysiological signals. Indeed, in the EEG study (Proverbio & Pischedda, 2023a), N400 was larger to imagined positive appetitive states (*e.g.*, happiness and play) than negative ones (sadness or fear). In addition, N400 was of greater amplitude during imagery of thermal and nociceptive sensations, such as cold, than other motivational or visceral states. While somatosensory experiences (such as cold and pain) and affective states (such as happiness), emerge as conscious sensations from the gray matter (*i.e.*, close to surface electrodes), it is reasonable to hypothesize that the generators of visceral sensations and cravings are positioned further from the sensory detectors, producing electrical potentials that are weaker or noisier, and thus more challenging to classify. For instance, drowsiness or the urge to sleep is a physiological state regulated by the suprachiasmatic and dorsomedial nuclei of the hypothalamus (Saper, Scammell, & Lu, 2005); hunger and the need for food are primarily associated with the hypothalamic arcuate nucleus (Augustine, Lee, & Oka, 2020), while thirst involves the medial preoptic nucleus of the hypothalamus, along with the cingulate cortex and insula. These structures, particularly the hypothalamus, are located deep within the brain and do not directly contribute to the generation of electrical impulses detected by EEG (Nunez & Srinivasan, 2006). This similarity suggests a close relationship between the distinctiveness of bioelectrical signals and the accuracy of ML classification, which in turn supports the idea that ML algorithms might be able to offer some reliable glimpses into the content of the patient's mind as far as their physiological and psychological needs are concerned.

9.6. RQ3 electrode configuration analysis: imagery (14) vs imagery (18)

In the somato-sensory category, the pair 'cold_fear' shows high accuracy with both configurations (0.82 with 14 channels and 0.75 with 18 channels), suggesting consistent neural representations across different electrode densities.

In general, the differences in performance are not significant, suggesting that both configurations can capture the essential neural patterns for these states.

10. Limitations

This study presents several important findings regarding the classification of motivational states from EEG data, but it also has some limitations that should be addressed in future research.

First, the choice of a binary classification approach between pairs of mental states was made to simplify the model development process and ensure reliable generalization across subjects. However, this resulted in 132 separate experiments, which may raise concerns about the scalability and practicality of such an approach in real-world applications. While our method proved effective in this exploratory setting, further research should explore multiclass classification techniques or methods that can handle larger datasets more efficiently. Generalizing the approach to include more complex models that classify multiple states simultaneously could improve its applicability to larger and more diverse datasets.

Second, the study tested three machine learning models: SVM, Random Forest, and a feedforward neural network. While SVM was the most effective classifier in terms of performance, we recognize the importance of exploring a wider range of models, including deep learning architectures and ensemble methods, in future work. The inclusion of additional classifiers highlights the potential benefits of adopting more complex approaches to improve generalizability and performance in similar applications.

Third, the relatively small participant sample size (N=20) presents a potential limitation, particularly when using LOSO cross-validation. A small sample size may reduce the generalizability of the results, and larger datasets would likely improve the robustness of the findings. Although LOSO cross-validation is appropriate for the sample size used, future studies with more participants could help strengthen the validity and generalization of the models developed here. Increasing the sample size would also enhance the reliability of the ML models, making the findings more applicable across a broader range of individual differences and settings.

Finally, while we found that the electrode density had only a marginal effect on performance, the relatively low density used in this study may limit the resolution of the EEG data. Future research could investigate the impact of higher-density electrode configurations and explore alternative neuroimaging modalities or different scenarios to improve the spatial and temporal resolution of the neural signals associated with motivational states.

11. Conclusion

This study set out to explore the neural correlates of various motivational states using EEG data and ML algorithms. By examining both perception and imagery conditions and comparing different EEG configurations, we aimed to uncover the distinct neural signatures associated with these states. Our findings highlight the distinct neural signatures associated with perceived and imagined motivational states, which can be leveraged to develop more effective BCIs. The results suggest that 14-channel EEG setups are sufficiently robust for many applications, making BCI systems more cost-effective and less cumbersome. Understanding the neural dynamics of perception versus imagination can improve models of brain function and inform both clinical and non-clinical interventions.

In the EEG study, the imagined mental states were recalled and not really felt by participants (*e.g.*, states of pain were vividly recalled but in the absence of external nociceptive sensations). The simulation might differ from the real patient's mental state, but this problem applies to most imagination studies. Future research should aim to enhance the vividness and consistency of mental imagery, investigate the stability of neural patterns over time, and expand the range of motivational states studied.

CRediT authorship contribution statement

Tommaso Colafiglio: Methodological idea for the study, Methodology, Performed the analysis, Software, Writing – original draft, Analyzed and interpreted the results, Edited the final version of the manuscript. **Angela Lombardi:** Methodological idea for the study, Methodology, Performed the analysis, Visualization, Writing – original draft, Analyzed and interpreted the results, Edited the final version of the manuscript. **Tommaso Di Noia:** Methodology, Validated the methodological framework, Supervision, Analyzed and interpreted the results, Edited the final version of the manuscript. **Maria Luigia Natalia De Bonis:** Analyzed and interpreted the results, Edited the final version of the manuscript. **Fedelucio Narducci:** Methodological idea for the study, Methodology, Writing – original draft, Supervision, Analyzed and interpreted the results, Edited the final version of the manuscript. **Alice Mado Proverbio:** Methodological idea for the study, Data curation, Writing – original draft, Validated the clinical findings, Supervision, Analyzed and interpreted the results, Edited the final version of the manuscript.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Alice Mado Proverbio reports financial support was provided by University of Milan- Bicocca. Tommaso Colafiglio reports financial support was provided by University of Rome La Sapienza. Maria Luigia Natalia De Bonis reports financial support was provided by Polytechnic University of Bari. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partial supported by the following projects: "LIFE: the itaLian system wIde Frailty nEtwork", "DEMETRA: Development of an ensemble learning-based, multidimensional sensory impairment score to predict cognitive impairment in an elderly cohort of Southern Italy" (CUP D99J22001970006) Missione 6/componente 2/Investimento: 2.1 "Rafforzamento e potenziamento della ricerca biomedica del SSN", funded by European Commission – NextGenerationEU, CTEMT - "Casa delle Tecnologie Emergenti di Matera". "IDENTITA - rete Integrata meDiterranea per l'osservazione ed Elaborazione di percorsi di Nutrizione". This work was also carried out while Tommaso Colafiglio was enrolled in the Italian National Doctorate on Artificial Intelligence run by Sapienza University of Rome in collaboration with Politecnico di Bari. EEG data were collected as part of the project "Reading Mental Representations through EEG Signals", funded by a University of Milano-Bicocca grant, Italy (ATE – Fondo di Ateneo No. 31159-2019-ATE-0064) awarded to AMP. The authors gratefully acknowledge the technical support of Francesca Pischedda and Alessandra Brusa.

All authors approved the final version of the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eswa.2025.127076.

Data availability

Data will be made available on request.

References

Aggarwal, S., & Chugh, N. (2022). Review of machine learning techniques for EEG based brain computer interface. *Archives of Computational Methods in Engineering*, 1–20.

Alho, J., Gotsopoulos, A., & Silvanto, J. (2023). Where in the brain do internally generated and externally presented visual information interact? *Brain Research*, *1821*, Article 148582.

Augustine, V., Lee, S., & Oka, Y. (2020). Neural control and modulation of thirst, sodium appetite, and hunger. *Cell*, *180*(1), 25–32.

Cai, Y., She, Q., Ji, J., Ma, Y., Zhang, J., & Zhang, Y. (2022). Motor imagery EEG decoding using manifold embedded transfer learning. *Journal of Neuroscience Methods*, *370*, Article 109489.

Carbine, K. A., Rodeback, R., Modersitzki, E., Miner, M., LeCheminant, J. D., & Larson, M. J. (2018). The utility of event-related potentials (ERPs) in understanding food-related cognition: A systematic review and recommendations. *Appetite*, *128*, 58–78.

Cecotti, H., & Graeser, A. (2011). Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*, 433–445. http://dx.doi.org/10.1109/TPAMI.2010.125.

Chang, W., Wang, H., Yan, G., & Liu, C. (2020). An EEG based familiar and unfamiliar person identification and classification system using feature extraction and directed functional brain network. *Expert Systems with Applications*, *158*, Article 113448.

Chattopadhyay, S., Zary, L., Quek, C., & Prasad, D. K. (2021). Motivation detection using EEG signal analysis by residual-in-residual convolutional neural network. *Expert Systems with Applications*, *184*, Article 115548.

Colafiglio, T., Lofù, D., Sorino, P., Lombardi, A., Narducci, F., Festa, F., et al. (2024). EmoSynth real time emotion-driven sound texture synthesis via brain-computer interface. In *Adjunct proceedings of the 32nd ACM conference on user modeling, adaptation and personalization* (pp. 616–621).

Colafiglio, T., Sorino, P., Lofu, D., Lombardi, A., Narducci, F., & Di Noia, T. (2023). Combining mental states recognition and machine learning for neurorehabilitation. In *2023 IEEE international conference on systems, man, and cybernetics* (pp. 3848–3853). IEEE.

Dijkstra, N. (2024). Nuancing the heterarchical theory of visual mental imagery. *Physics of Life Reviews*, *49*, 10–11.

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. (2017). Vividness of visual imagery depends on the neural overlap with perception in visual areas. *Journal of Neuroscience*, *37*(5), 1367–1373.

Dijkstra, N., & Fleming, S. M. (2023). Subjective signal strength distinguishes reality from imagination. *Nature Communications*, *14*(1), 1627.

Doyle, T. E., Kucerovsky, Z., & Ieta, A. (2006). Affective state control for neuroprostheses. In *28th international conference of the IEEE engineering in medicine and biology society, EMBC 2006, new york city, NY, USA, August 30 - September 3, 2006, main volume* (pp. 1248–1251). IEEE, http://dx.doi.org/10.1109/IEMBS.2006.260531.

Du, X., Tang, Y., Jiang, Y., & Tian, Y. (2022). Individuals attention bias in perceived loneliness: an ERP study. *Brain-Apparatus Communication: A Journal of Bacomics*, *1*(1), 50–65.

Echtioui, A., Zouch, W., Ghorbel, M., Mhiri, C., & Hamam, H. (2024). Classification of BCI multiclass motor imagery task based on artificial neural network. *Clinical EEG and Neuroscience*, *55*(4), 455–464.

Fan, J., Wade, J., Key, A., Warren, Z., & Sarkar, N. (2017). EEG-based affect and workload recognition in a virtual driving environment for ASD intervention. *IEEE Transactions on Biomedical Engineering, PP*, http://dx.doi.org/10.1109/TBME.2017.2693157, 1–1.

Fulford, J., Milton, F., Salas, D., Smith, A., Simler, A., Winlove, C., et al. (2018). The neural correlates of visual imagery vividness–an fMRI study and literature review. *Cortex*, *105*, 26–40.

Gembler, F., Stawicki, P., & Volosyak, I. (2017). How many electrodes are needed for multi-target SSVEP-BCI control: Exploring the minimum number of signal electrodes for CCA and MEC. In G. R. Müller-Putz, D. Steyrl, S. C. Wriessnegger, & R. Scherer (Eds.), *From vision to reality - proceedings of the 7th graz brain-computer interface conference, GBCIC 2017, graz, steiermark, Austria, September 18-22, 2017*. Verlag der Technischen Universitaet Graz, http://dx.doi.org/10.3217/978-3-85125-533-1-29.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*(267), 1–13. http://dx.doi.org/10.3389/fnins.2013.00267.

Habelt, B., Arvaneh, M., Bernhardt, N., & Minev, I. (2020). Biomarkers and neuromodulation techniques in substance use disorders. *Bioelectronic Medicine*, *6*(1), 4.

Hu, L., Gao, W., Lu, Z., Shan, C., Ma, H., Zhang, W., et al. (2024). Subject-independent wearable P300 brain-computer interface based on convolutional neural network and metric learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.

Ji, J. L., Heyes, S. B., MacLeod, C., & Holmes, E. A. (2016). Emotional mental imagery as simulation of reality: Fear and beyond—A tribute to peter lang. *Behavior Therapy*, *47*(5), 702–719.

Jochumsen, M., Hougaard, B. I., Kristensen, M. S., & Knoche, H. (2022). Implementing performance accommodation mechanisms in online BCI for stroke rehabilitation: A study on perceived control and frustration. *Sensors*, *22*(23), 9051.

Kanoh, S., Miyamoto, K., & Yoshinobu, T. (2011). A P300-based BCI system for controlling computer cursor movement. In *33rd annual international conference of the IEEE engineering in medicine and biology society, EMBC 2011, boston, MA, USA, August 30 - sept. 3, 2011* (pp. 6405–6408). IEEE, http://dx.doi.org/10.1109/IEMBS.2011.6091581.

Kavanagh, D. J., Andrade, J., & May, J. (2005). Imaginary relish and exquisite torture: The elaborated intrusion theory of desire.. *Psychological Review*, *112*(2), 446.

Khan, N. N., Sweet, T., Harvey, C. A., Warschausky, S., Huggins, J. E., & Thompson, D. E. (2023). P300-based brain-computer interface speller performance estimation with classifier-based latency estimation. *Journal of Visualized Experiments*, (199).

Kim, J.-M., Heo, K.-S., Shin, D.-H., Nam, H., Won, D.-O., Jeong, J.-H., et al. (2024). A learnable continuous wavelet-based multi-branch attentive convolutional neural network for spatio–spectral–temporal EEG signal decoding. *Expert Systems with Applications*, *251*, Article 123975.

Koenig-Robert, R., & Pearson, J. (2021). Why do imagery and perception look and feel so different? *Philosophical Transactions of the Royal Society B*, *376*(1817), Article 20190703.

Kunjan, S., Grummett, T. S., Pope, K. J., Powers, D. M., Fitzgibbon, S. P., Bastiampillai, T., et al. (2021). The necessity of leave one subject out (LOSO) cross validation for EEG disease diagnosis. In *Brain informatics: 14th international conference, BI 2021, virtual event, September 17–19, 2021, proceedings 14* (pp. 558–567). Springer.

Kuś, R., Duszyk, A., Milanowski, P., Łabęcki, M., Bierzyńska, M., Radzikowska, Z., et al. (2013). On the quantification of SSVEP frequency responses in human EEG in realistic BCI conditions. *PloS One*, *8*(10), Article e77536.

Lee, S., Jang, S., & Jun, S. C. (2022). Exploring the ability to classify visual perception and visual imagery EEG data: toward an intuitive BCI system. *Electronics*, *11*(17), 2706.

Leoni, J., Strada, S. C., Tanelli, M., Brusa, A., & Proverbio, A. M. (2022). Single-trial stimuli classification from detected P300 for augmented brain–computer interface: A deep learning approach. *Machine Learning with Applications*, *9*, Article 100393.

Leoni, J., Strada, S. C., Tanelli, M., Jiang, K., Brusa, A., & Proverbio, A. M. (2021). Automatic stimuli classification from ERP data for augmented communication via brain–computer interfaces. *Expert Systems with Applications*, *184*, Article 115572.

Leoni, J., Strada, S. C., Tanelli, M., & Proverbio, A. M. (2023). MIRACLE: Mind ReAding classification engine. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*.

Llorella, F. R., Patow, G., & Azorín, J. M. (2020). Convolutional neural networks and genetic algorithm for visual imagery classification. *Physical and Engineering Sciences in Medicine*, *43*, 973–983.

Lombardi, A., Diacono, D., Amoroso, N., Biecek, P., Monaco, A., Bellantuono, L., et al. (2022). A robust framework to investigate the reliability and stability of explainable artificial intelligence markers of mild cognitive impairment and Alzheimer's disease. *Brain Informatics*, *9*(1), 17.

Ludowig, E., Bien, C. G., Elger, C. E., & Rosburg, T. (2010). Two P300 generators in the hippocampal formation. *Hippocampus*, *20*(1), 186–195.

Marmolejo-Ramos, F., Hellemans, K., Comeau, A., Heenan, A., Faulkner, A., Abizaid, A., et al. (2015). Event-related potential signatures of perceived and imagined emotional and food real-life photos. *Neuroscience Bulletin*, *31*, 317–330.

Maslova, O., Komarova, Y., Shusharina, N., Kolsanov, A., Zakharov, A., Garina, E., et al. (2023). Non-invasive EEG-based BCI spellers from the beginning to today: a mini-review. *Frontiers in Human Neuroscience*, *17*, Article 1216648.

Mattioli, F., Porcaro, C., & Baldassarre, G. (2022). A 1D CNN for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface. *Journal of Neural Engineering*, *18*(6), Article 066053.

Morozova, M., Nasibullina, A., Yakovlev, L., Syrov, N., Kaplan, A., & Lebedev, M. (2024). Tactile versus motor imagery: differences in corticospinal excitability assessed with single-pulse TMS. *Scientific Reports*, *14*(1), 14862.

Motoyama, H., & Hishitani, S. (2024). The neural basis of a cognitive function that suppresses the generation of mental imagery: Evidence from a functional magnetic resonance imaging study. *Vision*, *8*(2), 18.

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: the neurophysics of EEG*. USA: Oxford University Press.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719.

Pan, J., & Li, Y. (2017). A gaze-independent audiovisual brain-computer interface and its application in awareness detection. In G. R. Müller-Putz, D. Steyrl, S. C. Wriessnegger, & R. Scherer (Eds.), *From vision to reality - proceedings of the 7th graz brain-computer interface conference, GBCIC 2017, graz, steiermark, Austria, September 18-22, 2017*. Verlag der Technischen Universitaet Graz, http://dx.doi.org/10.3217/978-3-85125-533-1-71.

Pan, J., Yang, F., Qiu, L., & Huang, H. (2022). Fusion of EEG-based activation, spatial, and connection patterns for fear emotion recognition. *Computational Intelligence and Neuroscience*, *2022*(1), Article 3854513.

Proverbio, A., & Pischedda, F. (2023a). Measuring brain potentials of imagination linked to physiological needs and motivational states. *Frontiers in Human Neuroscience*, *17*, Article 1146789.

Proverbio, A. M., & Pischedda, F. (2023b). Validation of a pictionary-based communication tool for assessing physiological needs and motivational states: The PAIN set. *Frontiers in Cognition*, *2*, Article 1112877.

Proverbio, A. M., Tacchini, M., & Jiang, K. (2022). Event-related brain potential markers of visual and auditory perception: A useful tool for brain computer interface systems. *Frontiers in Behavioral Neuroscience*, *16*, Article 1025870.

Proverbio, A. M., Tacchini, M., & Jiang, K. (2023). What do you have in mind? ERP markers of visual and auditory imagery. *Brain and Cognition*, *166*, Article 105954.

Rakotomamonjy, A., Guigue, V., Mallet, G., & Alvarado, V. (2005). Ensemble of SVMs for improving brain computer interface P300 speller performances. In W. Duch, J. Kacprzyk, E. Oja, & S. Zadrożny (Eds.), *Artificial neural networks: biological inspirations – ICANN 2005* (pp. 45–50). Berlin, Heidelberg: Springer Berlin Heidelberg.

Rodriguez, J., Del-Valle-Soto, C., & Gonzalez-Sanchez, J. (2022). Affective states and virtual reality to improve gait rehabilitation: A preliminary study. *International Journal of Environmental Research and Public Health*, *19*(15), 9523.

Rogers, P. J. (2017). Food and drug addictions: Similarities and differences. *Pharmacology Biochemistry and Behavior*, *153*, 182–190.

Sairamya, N., Subathra, M., & George, S. T. (2022). Automatic identification of schizophrenia using EEG signals based on discrete wavelet transform and RLNDiP technique with ANN. *Expert Systems with Applications*, *192*, Article 116230.

Salvaris, M., & Sepulveda, F. (2009). Visual modifications on the P300 speller BCI paradigm. *Journal of Neural Engineering*, *6*(4), Article 046011.

Saper, C. B., Scammell, T. E., & Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature*, *437*(7063), 1257–1263.

Sharma, R., Kim, M., & Gupta, A. (2022). Motor imagery classification in brain-machine interface with machine learning algorithms: Classical approach to multi-layer perceptron model. *Biomedical Signal Processing and Control*, *71*, Article 103101.

Shen, G., Dwivedi, K., Majima, K., Horikawa, T., & Kamitani, Y. (2019). End-to-end deep image reconstruction from human brain activity. *Frontiers in Computational Neuroscience*, *13*, 21.

Sitaram, R., Lee, S., Ruiz, S., Rana, M., Veit, R., & Birbaumer, N. (2011). Real-time support vector classification and feedback of multiple emotional brain states. *Neuroimage*, *56*(2), 753–765.

Spagna, A., Hajhajate, D., Liu, J., & Bartolomeo, P. (2021). Visual mental imagery engages the left fusiform gyrus, but not the early visual cortex: A meta-analysis of neuroimaging evidence. *Neuroscience & Biobehavioral Reviews*, *122*, 201–217.

Syrov, N., Yakovlev, L., Nikolaeva, V., Kaplan, A., & Lebedev, M. (2022). Mental strategies in a P300-BCI: Visuomotor transformation is an option. *Diagnostics*, *12*(11), 2607.

Tang, X., Zhang, J., Qi, Y., Liu, K., Li, R., & Wang, H. (2024). A spatial filter temporal graph convolutional network for decoding motor imagery EEG signals. *Expert Systems with Applications*, *238*, Article 121915.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 99–114.

Vanhaudenhuyse, A., Schnakers, C., Boly, M., Bruno, M., Gosseries, O., Cologan, V., et al. (2007). Behavioural assessment and functional neuro-imaing in vegetative state patients. *Revue Medicale de Liege*, *62*, 15–20.

Walsh, B. E., & Schlauch, R. C. (2024). Differential impact of emotional and social loneliness on daily alcohol consumption in individuals with alcohol use disorder. *Drug and Alcohol Dependence*, *264*, Article 112433.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, *113*(6), 767–791.

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, *72*, 2031–2046.

Yao, L., Mrachacz-Kersting, N., Sheng, X., Zhu, X., Farina, D., & Jiang, N. (2018). A multi-class BCI based on somatosensory imagery. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*, 1508–1515, URL https://api.semanticscholar.org/CorpusID:51610953.

Zhang, J., Gao, S., Zhou, K., Cheng, Y., & Mao, S. (2023). An online hybrid BCI combining SSVEP and EOG-based eye movements. *Frontiers in Human Neuroscience*, *17*, Article 1103935.

Zhang, B., Zhou, Z., & Jiang, J. (2020). A 36-class bimodal ERP brain-computer interface using location-congruent auditory-tactile stimuli. *Brain Sciences*, *10*, 524. http://dx.doi.org/10.3390/brainsci10080524.

Zhang, Y., Zhou, G., Zhao, Q., Jin, J., Wang, X., & Cichocki, A. (2013). Spatial-temporal discriminant analysis for ERP-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society*, *21*, 233–243. http://dx.doi.org/10.1109/TNSRE.2013.2243471.