A new error-monitoring brain–computer interface based on reinforcement learning for people with autism spectrum disorders

Journal of Neural Engineering

PAPER

A new error-monitoring brain–computer interface based on reinforcement learning for people with autism spectrum disorders

Gabriel Pires[1,3,*], Aniana Cruz[1], Diogo Jesus[1], Mine Yasemin[1], Urbano J Nunes[1,4], Teresa Sousa[2] and Miguel Castelo-Branco[2,5]

[1] Institute of Systems and Robotics of the University of Coimbra, Coimbra, Portugal
[2] Coimbra Institute for Biomedical Imaging and Translational Research of the University of Coimbra, Coimbra, Portugal
[3] Engineering Department, Polytechnic Institute of Tomar, Tomar, Portugal
[4] Department of Electrical and Computer Engineering, University of Coimbra, Coimbra, Portugal
[5] Faculty of Medicine, University of Coimbra, Coimbra, Portugal
* Author to whom any correspondence should be addressed.

E-mail: gpires@isr.uc.pt

Keywords: brain–computer interface, observation error-related potential (ErrP), reinforcement learning, autism spectrum disorders

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
26 June 2022

REVISED
18 October 2022

ACCEPTED FOR PUBLICATION
30 November 2022

PUBLISHED
12 December 2022

Abstract

Objective. Brain–computer interfaces (BCIs) are emerging as promising cognitive training tools in neurodevelopmental disorders, as they combine the advantages of traditional computerized interventions with real-time tailored feedback. We propose a gamified BCI based on non-volitional neurofeedback for cognitive training, aiming at reaching a neurorehabilitation tool for application in autism spectrum disorders (ASDs). *Approach.* The BCI consists of an emotional facial expression paradigm controlled by an intelligent agent that makes correct and wrong actions, while the user observes and judges the agent's actions. The agent learns through reinforcement learning (RL) an optimal strategy if the participant generates error-related potentials (ErrPs) upon incorrect agent actions. We hypothesize that this training approach will allow not only the agent to learn but also the BCI user, by participating through implicit error scrutiny in the process of learning through operant conditioning, making it of particular interest for disorders where error monitoring processes are altered/compromised such as in ASD. In this paper, the main goal is to validate the whole methodological BCI approach and assess whether it is feasible enough to move on to clinical experiments. A control group of ten neurotypical participants and one participant with ASD tested the proposed BCI approach. *Main results.* We achieved an online balanced-accuracy in ErrPs detection of 81.6% and 77.1%, respectively for two different game modes. Additionally, all participants achieved an optimal RL strategy for the agent at least in one of the test sessions. *Significance.* The ErrP classification results and the possibility of successfully achieving an optimal learning strategy, show the feasibility of the proposed methodology, which allows to move towards clinical experimentation with ASD participants to assess the effectiveness of the approach as hypothesized.

1. Introduction

Initially developed to create an artificial channel of communication for the user, brain–computer interface (BCI) systems are increasingly being considered for assisted cognitive training. BCI training combines the advantages of traditional computerized interventions with real-time tailored feedback. It provides personalized cognitive training based on specific neuronal patterns known to be of interest to achieve a behavioral result [1, 2].

BCI cognitive training has been mainly applied as an assistive tool in neurorehabilitation, aging, and neurodevelopmental conditions, such as attention-deficit/hyperactivity and autism spectrum disorders (ADHD and ASD) [3–7]. In the specific domain of ASD, a perturbation characterized by persistent social deficits related to impaired neurobehavioral

responses to social attention cues [8], it has been suggested in [9] that cognitive BCI training could lead to substantial improvements in behavior, cognition, and in particular social and affective processing. Their study presents a social mirroring game for encouraging social interactions, where the feedback is based on the participants' ability to control their mu rhythm. More recently, in [10] we have applied a combined virtual-reality P300-based BCI to train ASD social skills, particularly joint attention. A training paradigm was developed where social cues (eye gaze of a virtual avatar) were used to direct the focus of attention. A clinical trial following this approach has demonstrated the feasibility of BCI for cognitive training in ASD [11].

In addition to the social interaction impairments related to early difficulties in the expression and recognition of emotions [12], and disrupted communication abilities, ASD is characterized by repetitive and stereotyped behaviors [13]. Impairments in error-monitoring processes have been linked to these recurring and restricted patterns of behaviors. Increased responses in the anterior medial prefrontal cortex, left superior temporal gyrus, and the insula have been found during error commissions in high functioning autism [14]. These patterns suggest greater attentional deployment towards the internally driven emotional state associated with making an error in ASD. Since error-monitoring processes are recruited across different cognitive tasks throughout daily activities, increased emotional reaction to errors may have important consequences regarding the efficiency of learning processes. Individuals with ASD also present functional and structural abnormalities in the anterior cingulate cortex (ACC) [15]. Besides, there is evidence of altered electroencephalographic (EEG) components related to the neural correlates of error-monitoring in this condition [14, 16–18]. Such neuronal patterns are thought to compromise error awareness leading to impaired behavioral adjustments.

EEG research has revealed particular types of event-related potentials (ERPs) which are signatures of error-monitoring, namely the medial-frontal error-related negativity (ERN) that appears after errors have been committed. It has been suggested that ERN is elicited whenever an erroneous outcome emerges and it is followed by a centro-parietal error positivity (Pe), which seems to reflect the conscious recognition of errors [19–21]. Error-monitoring processes are mediated by a network that includes the prefrontal cortex, the ACC, and the striatum. Moreover, the ACC has been indicated as the neuronal source of the ERN and Pe, playing a central role in error/conflict monitoring within social contexts [22]. The ERN has shown to be robust across different types of responses and has been also linked to error-related midfrontal theta rhythm, which during error commission shares the same neuronal source and appears to be modulated already during response preparation [23]. Such frequency component has shown to be an important complementary signature of error detection, particularly in complex social contexts [24].

The ERPs occurring in response to errors have been generically called error-related potentials (ErrPs). It has been shown that the waveform morphology of ErrPs varies depending on the monitoring process. The first type of ErrPs that has been reported was termed 'Response-ErrP', appearing when a mistake or an incorrect action is made by the user and he/she recognizes it immediately (typically in speeded motor reaction tasks) [19]. The authors reported an ERN at 80 ms and a Pe ranging from 200 to 500 ms. When the subject is not aware of the error until external feedback, it is called 'Feedback-ErrP' and the main component is a negative deflection between 200 and 300 ms after feedback [25]. 'Observation-ErrP' or 'Recognition-ErrP' occurs when the subject is observing an operator or agent performing a task and the operator/agent makes an error. This ErrP, similarly to the 'Feedback-ErrP' is characterized by a negative deflection between 200 and 300 ms after the error, but presents a different morphology [26]. The last known type of error potentials is the 'Interaction-ErrP', which is elicited in BCI tasks, i.e. when the BCI provides an output or feedback differing from the subject intention. It is characterized by a small positive peak around 200 ms, a negative peak around 250 ms, a positive peak around 350 ms, and a second negative peak around 450 ms after the BCI feedback [27].

In the context of BCI research, error potentials have been mainly used to increase the BCI reliability through the correction of incorrect outputs given by the system [28–30]. More recently, interest has increased in using ErrPs to teach autonomous systems the perception and/or intention of the user who observes their actions [31–34] and in human-agent co-adaptation [35]. Chavarriaga and Millan [31] were the first proposing the combination of artificial learning of an intelligent agent with information stemming from real automatic responses of a human. They simulated in a pseudo-online way (i.e. performed offline, but mimicking the online process) the possibility of an agent to learn through reinforcement learning (RL) using rewards defined by ErrPs automatically provided by the human observing the agent.

In the current study, we further developed this concept to propose a BCI based on non-volitional neurofeedback for cognitive training, aiming at reaching a potential neurorehabilitation tool for ASD. In particular, we propose a BCI training approach in the form of a game where the participant has to detect whether the actions of an intelligent agent are correct or erroneous taking into account various facial cues,

expressing emotions, as instructions. We hypothesize that this training approach will allow not only the intelligent agent to learn but also the BCI user through operant conditioning. The neurofeedback is non-volitional in the sense that it does not result from participant's volitional EEG modulation of error signals, but rather indirectly from his active observation and strategy to assess the correctness of the agent's actions. If the intelligent agent learning progress is fed back, the subject will sense whether his teaching process is having the desired effect, and is able to adjust and improve his attentional strategy in response to that feeling. This can lead to reduced anxiety while increasing motivated behavior. Different from the traditional neurofeedback, where the participant is asked to consciously control a feedback signal (for example, associated to a specific frequency band), in non-volitional neurofeedback, also known as covert or implicit neurofeedback, the participant is simply asked to maximize a reward [36, 37]. This type of intervention is emerging as a promising approach capable of changing behavioral or physiological outcomes in several clinical conditions such as anxiety disorders or emotion perception [38]. The proposed BCI approach is expected not only to allow participants to unconsciously train the error-monitoring neuronal network but also to improve the recognition of the type of facial expressions used as instructions. Finally, collected brain error-related signals may also be used as potential quantitative measures (biomarkers) that can aid in the profiling of participant's characteristics in relation to psychological assessment variables.

The main goal of this paper is to validate the feasibility of the whole BCI methodological approach, so that it can move to clinical validation. This was done by testing the BCI with a control group of ten neurotypical people, and also an ASD participant who was included to gain preliminary experimental insight into the possibility for clinical translation. To our knowledge, this is the first RL-ErrP BCI targeted at the ASD population.

2. Methods

2.1. Participants and data Acquisition

Ten healthy participants (mean age: 28.1 ± 4.43 years, 5 women) and one high-functioning autistic male (33 years; full-scale intelligent quotient = 119) recruited from the Coimbra APPDA association (Associação Portuguesa para as Perturbações do Desenvolvimento e Autismo) participated in the study. The ASD diagnosis was assigned on the basis of the gold standard instruments: parental or caregiver interview (autism diagnostic interview—revised, ADI-R [39]), direct structured subject assessment (autism diagnostic observation schedule, ADOS [40]), and the current diagnostic criteria for ASD according to the diagnostic and statistical manual of mental disorders 5, DSM-5 [13]. The characterization details care as follows: ADOS-Communication: 2, ADOS-Socialization: 7, ADOS-Total: 9, ADI-R-Socialization: 20, ADI-R-Communication: 14, ADI-R-Restricted and stereotyped behaviors: 9. The study was approved by the Ethical Committee of the Faculty of Medicine of the University of Coimbra and was conducted complying with the code of Ethics of the Declaration of Helsinki. Informed consent was signed by participants.

EEG data were recorded with a g.USBamp bioamplifier from 16 electrodes: Fpz, Fz, FCz, FC1, FC2, Cz, C3, C4, CPz, Pz, P3, P4, PO7, PO8, POz and Oz. The electrodes were referenced to one of the earlobes and the ground was located at AFz. The EEG signals, sampled at 256 Hz, were acquired with active Ag/AgCl electrodes and pre-processed using a notch filter at 50 Hz and a band-pass filter with lower cutoff frequency of 0.5 Hz and a higher cutoff frequency of 30 Hz.

2.2. Gamified task and experimental procedures

In this study we examine the ability to evoke ErrPs in participants monitoring an intelligent agent whose actions are coupled to faces expressing sadness or happiness emotions. The agent performs autonomously correct or incorrect actions with a pre-defined probability, i.e. the user has no direct control over the agent. The overall system is schematically represented in figure 1. The system was designed as a gamified task to explore its viability for error-monitoring training in subjects with ASD using non-volitional neurofeedback. It comprises two main modules, a visual paradigm called 'Emotional Facial Expression Paradigm' (EFP) and an 'Intelligent Agent' that learns through RL from participant's judgments to the actions of the Intelligent Agent. These two modules are explained in sections 2.2.1 and 2.2.2 respectively.

2.2.1. Emotional facial expression paradigm (EFP)

An illustrative example of the EFP and respective timeline is shown in figure 2. Participants were seated in front of a computer screen and asked to observe the movements related by a given rule to a particular facial expression (referred to as cue image) expressing either happiness or sadness. The initial position of the facial expression is randomly chosen from predefined positions in the center horizontal axis of the screen. The movements follow two predefined rules: (a) if the image is showing an expression of happiness it should move to the center; (b) if the image shows an expression of sadness, it should move outwards. However, the movements have a predefined probability of not being correct (e.g. happy face moving outwards instead of inwards). The participants were instructed to mentally judge the movement of the image as correct or incorrect and were informed

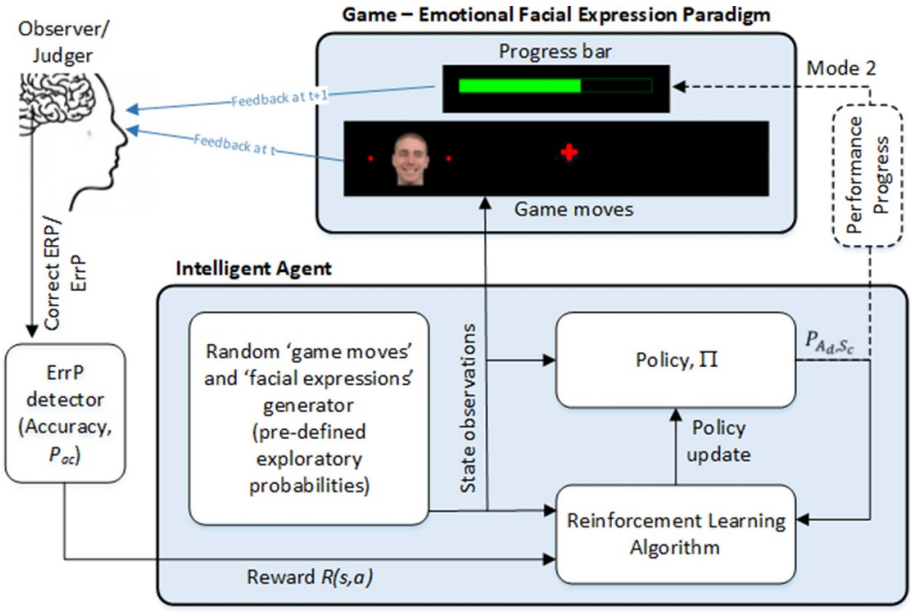


Figure 1. Diagram of the entire gamified task. It comprises two main modules: (a) the visual emotional facial expression paradigm (EFP); and (b) the Intelligent Agent that generates the states and the actions of the game, and that learns a policy through reinforcement learning. The player observes the EFP and judges its behavior providing positive or negative rewards that feed the RL algorithm. Dashed lines and blocks are used only in mode 2 (performance feedback) of the game. Reproduced with permission from [41].

that it was the users' function to mentally identify when the game was performing incorrectly. We followed the hypothesis of an observation-ErrP (oErrP) being elicited by the observation of wrong movement and a correct event-related potential (cERP) being elicited otherwise. Participants were told that the BCI system could learn the two rules if they correctly identified the correct and wrong movements of the cue image, i.e. they had the 'power' of teaching the BCI system just by attending to the occurrence of errors. The images expressing emotions, obtained from the Radboud faces database [41], are listed in figure 2(a). Given the deficits in communication and social interaction in autism, these images provide an appropriate social context for the gamified paradigm [42, 43]. Additionally, the emotional expressions of the faces are complemented by left or right gazing, which add an important social cue related to joint attention given that it has been suggested that children with autism are relatively insensitive to communication information based on gaze direction [11]. In the images, the gaze direction is always congruent with inward or outward movement (e.g. a happy face is always gazing to the center and a sad face outward, regardless of whether its position is to the left or right of the center of the screen).

Figure 2(b) shows a screenshot of the EFP at a given time of a trial. The initial cue image can appear in any of the six possible positions indicated with dashed rectangles. It is flanked by two red spots indicating the two possible target positions of the image after the movement. Participants should recognize the emotional expression as soon as it is presented. Then, they should direct their gaze to the target destination they believed was the correct one (one of the red spots). The inward or outward movement of the image occurs 1.25 s after the cue presentation. Importantly, the change of participants' gaze occurs before the inward or outward game movement to avoid contamination of the ERPs by ocular artifacts. Note that in the example of figure 2(b), the blue arrows and the dashed rectangles are not part of the paradigm, they were added here for a better understanding. The red cross at the center is also softer than the one here illustrated.

The timeline of a trial is represented in figure 2(c). Each trial starts with a black screen with a cross at the center. After 0.5 s, the cue image appears, the participant has to recognize the emotional expression, and gaze at the red spot expected as the target location after movement. The cue image moves at 1.75 s and the participant should judge the movement as correct or incorrect. The probability of wrong movements was set to 30% and was unknown to the participants. After image movement, it remains in that position for 1 s, and the trial ends. Depending on the game mode, the participant could receive feedback on his/her performance at the end of each trial (see section 2.4). A demonstrative video of the paradigm is provided as supplementary material.

2.2.2. Gamified RL

A gamified RL procedure was implemented to infer: (a) the possibility of the agent learning through

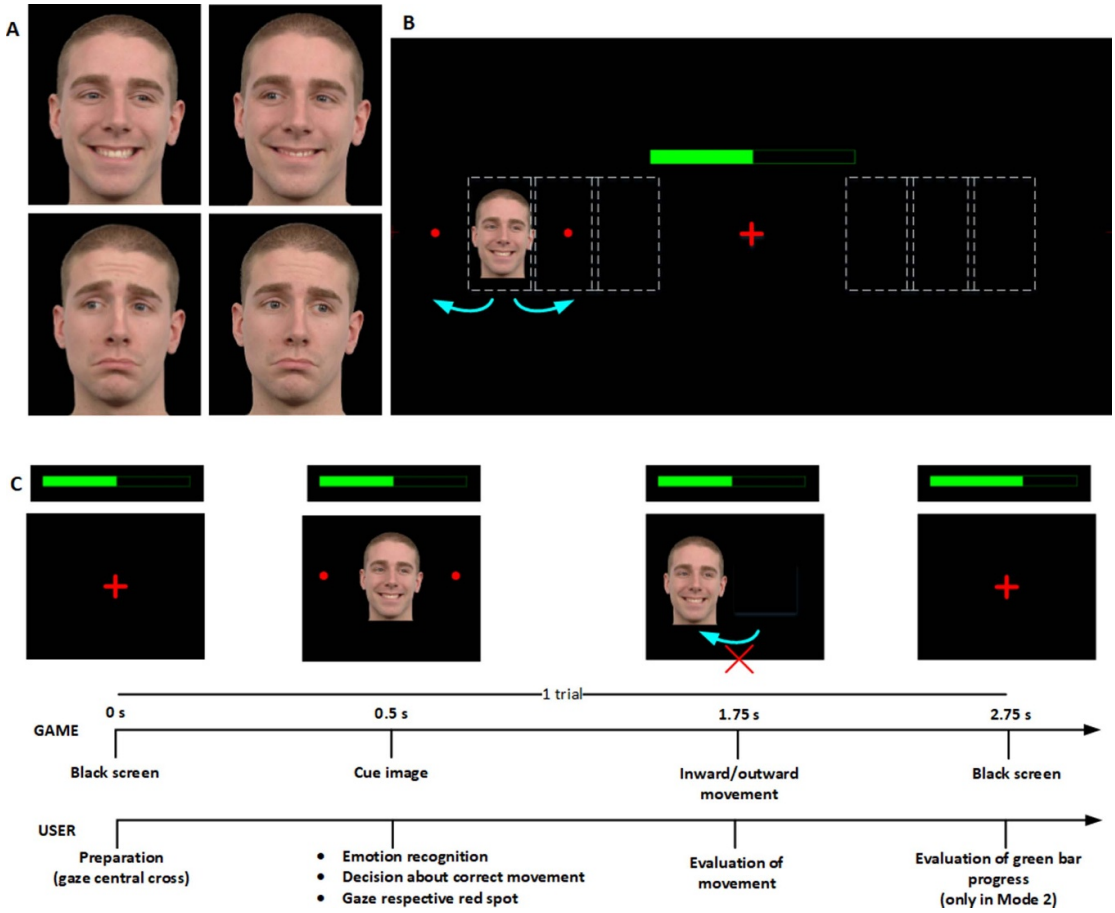


Figure 2. Emotional facial expression paradigm. (a) Close-up of the four facial expressions used (from the Radboud Faces Database https://www.ru.nl/bsi/research/facilities/radboud-faces-database/) (b) screenshot of the EFP during a testing session. The white dashed rectangles and the blue arrows are not part of the paradigm. They illustrate here respectively the possible initial positions of the cue image (illustrating an emotional expression) and the possible movements of the cue image (left or right). The red dots flanking the image define the possible target positions after image movement. The green bar increases or decreases according to the user's performance (only in game's mode 2). (c) Timeline with associated sequence of events and user tasks.

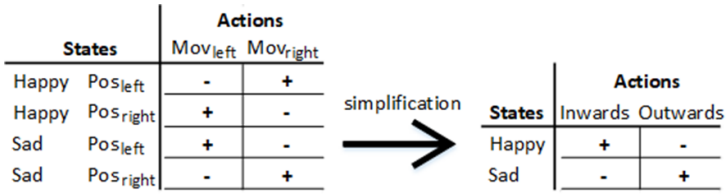


Figure 3. Simplified mapping of game states and actions. The position of the image on the left or right side of the screen can be disregarded as we can model the movements as inwards and outwards regardless of position, thus simplifying the policy function. '+' and '−' symbols stand for correct and wrong movement, respectively.

single-trial oErrP detection, and (b) how this learning could be used in the interaction with the participant. The intelligent agent module is schematically illustrated in figure 1. The agent randomly generates the facial expressions (happy and sad) as well as its initial position. This represents the state observations of the agent (expression, position). The possible actions are a movement to the left or right of the cue image. There is an optimal strategy for the agent, which consists on the movement of the happy image to the left if the image appears at the right side of the screen and a right movement if the happy image appears at left. The opposite movements should occur for sad images. This strategy can be modeled in a simplified way, as happy images should always move inward and sad faces outward independently of the position. The original and simplified state-action mappings are represented in figure 3.

Hence, considering S_c as the emotion expressed in the cue images, with $c \in [\mathrm{H}, \mathrm{S}]$ (happy and sad, respectively), and A_d as the direction to which the face moves, with $d \in [\mathrm{I}, \mathrm{O}]$ (inwards and outwards), the conditional probability of taking action A_d, given the cue image c under the strategy Π is represented as:

$$P_{A_d,S_c} = P(A_d|S_c,\Pi). \quad (1)$$

The optimal strategy (policy) is then defined as:

$$\Pi^* = [P_{A_O,S_S} = 1, P_{A_I,S_S} = 0, P_{A_I,S_H} = 1, P_{A_O,S_H} = 0]. \quad (2)$$

The goal of the agent is to learn this optimal strategy internally. At the beginning, the agent does not know the rules, therefore all probabilities are equal to 0.5. At each trial, the agent updates the probabilities based on the judgment of the observer (the participant). If the agent performs a wrong movement it is expected that an oErrP is elicited in the brain, otherwise a cERP is expected. The brain response is classified as oErrP or cERP in block 'ErrP-detector' represented in figure 1 (explained in detail in section 2.3). The participant's judgment provides the rewards to the learning algorithm. The reward depends on the reliability of the participant (the critic), that is, on the participant's ability to recognize the facial expression and associated movement, but also on the reliability of the ErrP-detector. As we are interested in assessing the participants' reliability, it would be desirable to have an ErrP-detector 100% reliable, but this may be difficult to obtain. However, it is important to make sure that the classification rate is acceptable so as not to derail the objective and gameplay of the task. Given the simplicity of the game rules and assuming a high reliability of the participant it may be expected a very fast learning of the optimal policy. However, for gamification purposes it is desired a slower and accurate convergence of the policy, that can show the performance of the participants but not very susceptible to the detector uncertainty. Let us consider a positive or negative reward $R(s,a)$ given by

$$R(s,a) = \pm 1 \times P_{\text{ac}} \quad (3)$$

where P_{ac} is the accuracy of the ErrP-detector that is used to adjust the reward weight according to the reliability of the detector. The policy is updated at instant $t+1$ according to

$$P^{t+1}_{A_d,S_c} = \begin{cases} P^t_{A_d,S_c} - P_{\text{ac}}\Delta P^t_{A_d,S_c} & \text{if oErrP detected} \\ P^t_{A_d,S_c} + P_{\text{ac}}\Delta P^t_{A_d,S_c} & \text{if cERP detected} \end{cases}. \quad (4)$$

Whenever the probability for one movement is updated, the opposite movement is updated accordingly. Let suppose as an example that a positive reward is obtained for a happy expression movement, then

$$\begin{cases} P^{t+1}_{A_I,S_H} = P^t_{A_I,S_H} + P_{ac}\Delta P^{t+1}_{A_I,S_H} \\ P^{t+1}_{A_O,S_H} = 1 - P^{t+1}_{A_I,S_H} \end{cases}. \quad (5)$$

The RL ΔP is variable, as proposed in [31], given by

$$\Delta P^t_{A_d,S_c} = \eta H(P^t_{A_d,S_c}) \quad (6)$$

where η is a learning rate factor (here set to 0.1) and $H(P^t_{A_d,S_c})$ is the binary entropy function. This way, update steps are larger at the beginning (probabilities close to chance level) and smaller as the rule certainty increases. The binary entropy function is defined by $H(x) = -\sum_i^n P(x_i)\log_2 P(x_i)$. For example, for Happy model

$$H(P_{A_d,S_H}) = -(P_{A_I,S_H}\log_2(P_{A_I,S_H}) + P_{A_O,S_H}\log_2(P_{A_O,S_H})). \quad (7)$$

Although the policy is internally updated, the agent continues to explore the world generating wrong movements with a 30% probability as set at the beginning.

It is noteworthy that although there are two states, happy and sad, they are independent of each other, and therefore there are two separate models that are being updated, one for happy and the other for sad. To provide feedback to the user of his/her performance in game mode 2, a progress bar (green bar) is filled according to average of the probabilities of correct movements of the two models

$$\text{progress}_{\text{bar}} = \frac{P^t_{A_O,S_S} + P^t_{A_I,S_H}}{2} \quad (8)$$

2.3. ErrP-detector

In order to classify each EEG epoch as erroneous or not, the EEG signals were segmented into 1-second epochs with onset on each stimulus (movement of the face). The signal was pre-processed applying a [0.5 10] Hz band-pass filter. Afterwards, a statistical spatial filter, Fisher criterion beamformer [44], was used to obtain the two most discriminative projections from the 16 EEG channels. Considering a spatio-temporal matrix $E_{N\times L}$, representing the epochs of N EEG channels with L time samples (in this case, as the time window is 1 s long, $L = 256$), the first and second projections are obtained respectively from the first and second spatial filters and then concatenated forming $y = [w_1^T E \;\; w_2^T E]$, where w_1 and w_2 and the two best spatial filters, and T denotes the transpose operator (see more details in [44]). After this operation, y is a vector of dimension $1 \times 2L$. In order to select the most relevant features (time samples of the projection), the r-squared correlation was used to select 200 features. This feature vector was then classified by a Bayes classifier which returns a target probability for each class (error or correct). The class with the highest probability is selected as the classification prediction. This is a classification pipeline that was already successfully validated in the context of ErrP prediction [28]. The statistical spatial filtering approach has the advantages of simultaneously increasing the signal-to-noise ratio (SNR) of the ERPs and reducing the feature vector dimensionality for classification. Moreover, this kind of approach is less sensitive to the classifier choice ([45] section 5.5.1).

2.4. Experimental procedure

During the experiment, the participant was sitting at a distance of 60 cm in front of a 24" computer screen. Each participant performed two sessions, a calibration session and a testing session. Before beginning these two sessions, the task was carefully explained to the participants, who were given the opportunity to test the gamified task, adjust the screen to participant's height and clarify any issue they might have. Each session consisted of 10 blocks and each block consisted of 30 trials (21 correct movements and 9 wrong movements). Each trial was 2.75 s long and the entire block was less than 1.5 min. There was an interval between blocks to rest, whose time depended on participants' will, but usually not exceeding 1 min. This interval was also used for debriefings. The calibration dataset comprised 90 error epochs and 210 correct epochs, that were used to obtain the classification models for the ErrP-detector.

Before starting the testing session, it was explained what would be different compared to the calibration session. The EFP of the testing session was similar to the calibration session except that oErrPs were now detected online and there was the possibility of providing feedback (green bar) based on the RL agent output. Participants tested two modes. In mode 1, participants did not receive any feedback, and in mode 2, participants received as feedback the progress of the green bar measuring their performance (as presented in figure 2). Each mode was tested in 5 blocks. Half of the participants started with mode 1 and the other half with mode 2. In both modes, at the end of the 5 blocks, the plot of the historical progress of their performance was shown.

2.5. Statistical analysis and metrics

2.5.1. Classification performance metrics

Given the imbalanced number of correct and error events, we used as classification performance metrics the sensitivity, the specificity, and the balanced accuracy which is given by the average between sensitivity and specificity as follows

$$bAcc = 0.5 \times \frac{TP}{TP + FN} + 0.5 \times \frac{TN}{TN + FP} \qquad (9)$$

where TP, TN, FN, and FP correspond to the number of true positives, true negatives, false negatives, and false positives, respectively. A true positive is an error-event well classified, and a true negative is a correct-event well classified. The results of the calibration session were obtained offline through leave-one-out ten-fold cross-validation (each block corresponding to one fold).

2.5.2. Statistical analysis

Paired sample t-tests (one-tailed) were used to compare the online classification results obtained in mode 1 and mode 2. To determine the differences statistically relevant between error and correct ERPs, a pointwise Welch's 2-sample t-test (two-tailed) was applied. The similarity of waveform morphology between ERPs was measured using the Pearson's correlation coefficient.

3. Results

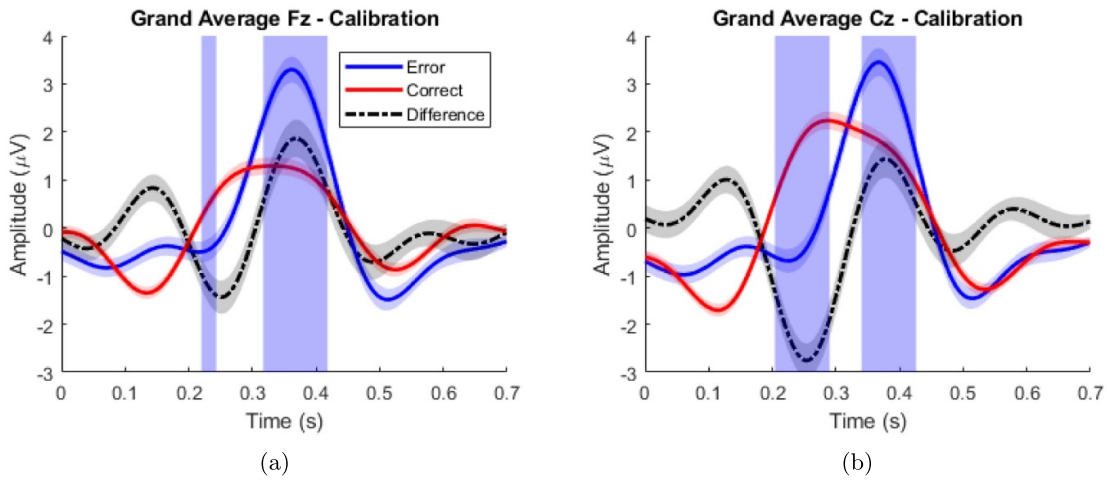
3.1. Neurophysiological analysis

The grand average (across all subjects) of ERPs for erroneous and correct trials in the calibration sessions of both paradigms are shown in figure 4, as well as the difference between error and correct trials at electrodes Fz and Cz, where differences related to error monitoring are expected. The plots were obtained from the calibration sessions of the ten participants, corresponding to 2100 correct epochs and 900 error epochs. At channel Fz, the oErrPs present a negative deflection (ERN) at 215 ms with a mean amplitude of $-0.499\ \mu$V and a positive deflection (Pe) at 359 ms with a mean amplitude of $3.290\ \mu$V. The cERP has a negative deflection at 132 ms with a mean amplitude of $-1.348\ \mu$V and a positive deflection at 328 ms with a mean amplitude of $1.29\ \mu$V.

To evaluate the discriminative time points, we computed pointwise Welch's t-tests comparing ErrPs and correct ERPs across all healthy participants. Statistically significant differences are highlighted in blue rectangles for a p-value $\leqslant 0.05$. It is clear that there are two discriminative time windows around the expected negative and positive deflections of oErrPs (ERN and Pe).

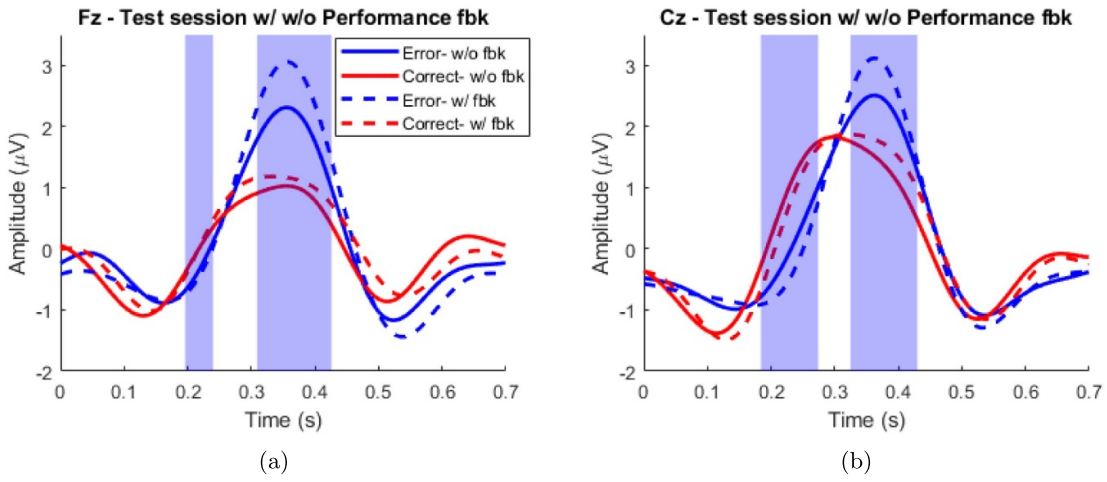
Figure 5 compares ErrPs and correct-related potentials of trials when participants received feedback on their performance (game mode 2) and when they did not receive feedback (game mode 1). The time windows for statistical significance (p-value $\leqslant 0.05$) in performance feedback trials are very similar to those obtained for the calibration data (figure 4). The waveforms obtained with and without performance feedback are slightly different, but the statistical t-test showed no statistically significant differences between the waveforms in the two conditions for any time point (for both correct trials and error trials).

Figure 6 shows correct and error trials and their difference at channel Fz in the time domain (a) and their difference in the time-frequency domain (b) ranging from -0.2 to 0.7 s. EEG data were recorded during calibration sessions from all participants in the control group and analyzed without any further processing (i.e. they are filtered only at range [0.5 30] Hz). At the moment of cue image movement (0 s) participants are at rest focused on the red spot and evaluate the movement as correct or wrong. The time-frequency plot was obtained in the [4 30] Hz range for the difference between error and correct trials using a two-cycle wavelet with a Hanning-tapered



(a) (b)

Figure 4. Grand average of oErrP and cERPs and difference between them across all subjects of the control group recorded at Fz (a) and Cz (b). For better visualization, the signals were filtered between [1 6] Hz with a digital zero-phase filter to avoid distortion. The shaded areas represent the standard error of the mean and the blue rectangles highlight the time windows where the difference is statistical significant (point-wise t-tests for p-value $\leqslant 0.05$).



(a) (b)

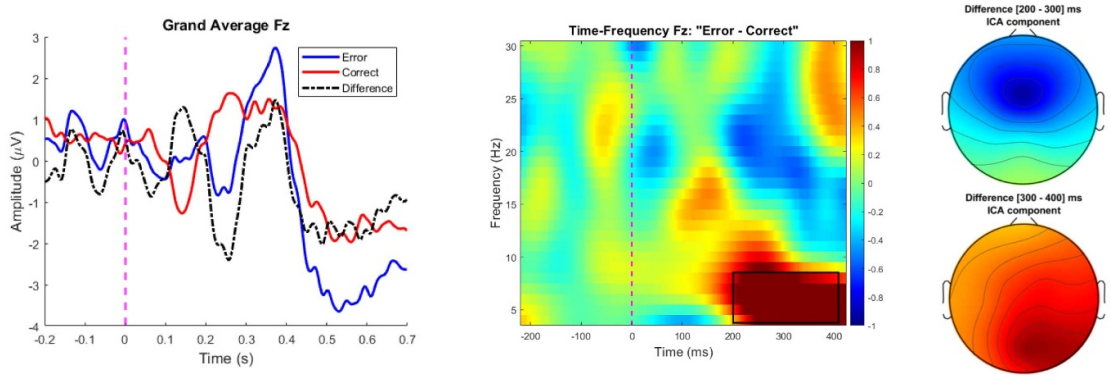
Figure 5. Grand average of oErrP and cERPs obtained during the test sessions with (w/) and without (w/o) providing feedback to participants of their performance (green bar progression) across all subjects of the control group recorded at Fz (a) and Cz (b). For better visualization, the signals were filtered between [1 6] Hz with a digital zero-phase filter to avoid distortion. The blue rectangles highlight the time windows where the difference of trials obtained with feedback performance are statistical significant (point-wise t-tests for p-value $\leqslant 0.05$).

window expanding with a 0.5 factor (obtained with EEGLab software). The signal was normalized by a baseline defined in the time window [−0.2 0] s, i.e. the period before image movement. The black rectangle corresponds to midfrontal theta band frequencies ([4 8] Hz), which are known to be related to error-monitoring processes, in the interval of interest where ERN and Pe peaks were previously identified ([200 400] ms). An increase in theta rhythms is visible for error trials.

Using EEGLab, topographic maps were also obtained from the difference between error and correct trials (figure 6(c)). First, the independent component analysis (ICA) [46] was applied selecting the time windows [200 300] and [300 400] ms associated to ERN and Pe respectively. The ICA components related to the cognitive processes were used to plot the topographic maps shown in figure 6(c). The scalp topography highlights the medial-frontal activity associated to ERN and the central-parietal activity associated to Pe.

3.2. Classification results

Table 1 shows the balanced classification accuracies obtained in the calibration session (10 blocks) and test session (5 blocks with performance feedback and 5 blocks without performance feedback). For each participant, the number of error trials and correct trials was 90 and 210, respectively in calibration and 45 and 105 for each game mode of the test session. The



(a) (b) (c)

Figure 6. (a) Grand average of time domain signal of error and correct trials and their difference, at channel Fz (signal filtered at [0.5 30] Hz); (b) time-frequency representation for the difference between error and correct trials (for the same data shown at (a)). Dashed vertical lines signalize the time of cue image movement. The black rectangle represents the interval where oErrPs were previously identified; (c) topographic maps obtained from ICA components related to the cognitive processes at [200 300] and [300 400] ms taking the difference between error and correct trials.

Table 1. Classification accuracy obtained offline (calibration session) and online (test session with and without feedback performance) for the control group (S1–S10) and ASD participant (A1). 'bAcc' and '♯Cvrg' stand respectively for balanced accuracy and number of sessions in which the optimal strategy was obtained. 'Order' indicates whether the respective test (with or without feedback) was made first (1) or later (2). 'ΔbAcc' is the bAcc difference between game mode 1 and game mode 2 and '*' refers to a statistically significant difference ($p = 0.009$). 'Avg' and 'Std' refer to average and standard deviation, respectively.

	Calibration session			Test with performance feedback					Test without performance feedback					
Subj	Sens	Spec	bAcc	Sens	Spec	bAcc	Order	♯Cvrg	Sens	Spec	bAcc	Order	♯Cvrg	ΔbAcc
S1	73.3	88.1	80.7	46.7	96.2	71.4	1	5	57.8	91.4	74.6	2	5	3.2
S2	82.2	79.0	80.6	86.7	71.4	79.0	2	3	75.6	81.9	78.7	1	5	−0.3
S3	74.4	77.6	76.0	73.3	75.2	74.3	1	2	82.2	78.1	80.2	2	3	5.9
S4	82.2	85.2	83.7	73.3	64.8	69.0	1	2	91.1	78.1	84.6	2	2	15.6
S5	70.0	81.0	75.5	57.8	90.5	74.1	2	5	64.4	83.8	74.1	1	4	0.0
S6	94.4	93.8	94.1	84.4	98.1	91.3	2	5	93.3	100.0	96.7	1	5	5.4
S7	87.8	86.2	87.0	71.1	82.9	77.0	1	4	75.6	77.1	76.3	2	4	−0.6
S8	73.3	67.1	70.2	84.4	65.7	75.1	1	0	84.4	68.6	76.5	2	2	1.4
S9	70.0	86.2	78.1	71.1	91.4	81.3	2	4	82.2	94.3	88.3	1	4	7.0
S10	84.4	88.6	86.5	77.8	80.0	78.9	2	4	86.7	85.7	86.2	1	4	7.3
Avg	**79.2**	**83.3**	**81.3**	**72.7**	**81.6**	**77.1**		**3.4**	**79.3**	**83.9**	**81.6**		**3.8**	**4.5***
Std	**8.2**	**7.4**	**6.9**	**12.5**	**12.2**	**6.2**		**1.6**	**11.3**	**9.3**	**7.2**		**1.1**	**4.9**
A1	87.8	83.8	85.8	86.7	61.0	73.8	1	1	95.6	62.9	79.2	2	1	5.4

results of the calibration session were obtained offline through leave-one-out ten-fold cross-validation (each block corresponding to one fold). The average accuracy obtained from all subjects in the control group (S1–S10) in the calibration was $81.3 \pm 6.9\%$. The online results obtained in the test session were respectively $81.6 \pm 7.2\%$ and $77.1 \pm 6.2\%$ for game mode 1 (without performance feedback) and mode 2 (with performance feedback). The accuracy difference of 4.5% was statistically significant (paired t-test, $p = 0.009$). The participant with ASD (A1) got performance results close to the control group average.

3.3. RL of agent

Regarding the hypothesis of the agent being able to learn using oErrPs and cERPs as biofeedback reward (negative/positive), the RL approach led to the optimal strategy in most blocks of the test sessions. Getting the optimal strategy means the rules have been learned. Table 1 shows that on average, the optimal strategy was obtained in 3.8 of the 5 blocks (game mode 1) and in 3.4 of the 5 blocks (game mode 2). Figure 7 shows representative examples of good (successful) and bad (unsuccessful) learning convergences obtained in two testing sessions by two participants. The curves are obtained directly from (8) starting from a probability 0.5. When the curves reach 1 (greenbar is 100% filled), the rules were learned (probabilities equal to 1), which happens in the example for subject S6.

4. Discussion

In this study, we aimed at exploring the use of oErrPs in a BCI cognitive application where the user is monitoring external agent actions with the intent to train error-monitoring through a gamified RL task. The user's detection of correct and wrong actions by

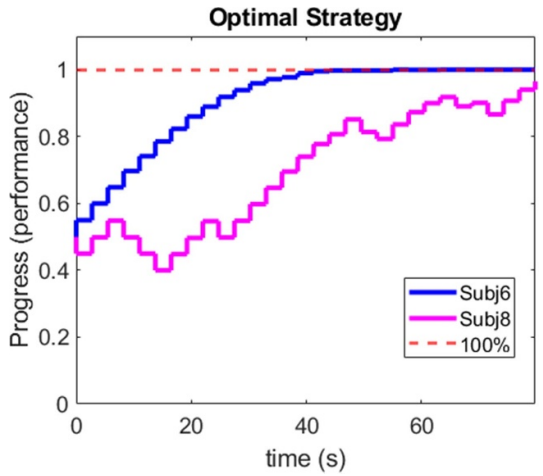


Figure 7. Two examples of green bar progression (performance feedback) representing agent's learning through user's performance on correctly detecting correct and wrong decisions of the agent. The bar progression is computed according to (8). These two curves represent examples of a session with good performance by participant S6 where the optimal strategy was achieved and a session with poor performance by participant S8 where there was a learning fluctuation and the optimal strategy was not reached.

the agent involves the processing of facial emotion expression cues. The implementation and validation of such a methodology, that detects in real-time oErrPs elicited in tasks associated with social/emotional cues, are of particular interest as an assistive training tool (as non-volitional neurofeedback) for ASD individuals. Here, we aimed to validate the methodology with a group of ten neurotypical individuals as a prior stage to clinical validation in ASD. It matters to understand whether the methodological approach, with emphasis on paradigm design and the accuracy of ErrP detection, is feasible enough to move on to clinical experiments. As a first exploratory step to help predict the feasibility of the experiment, a participant with ASD was enrolled in the experiment, but was not included in the group statistical analysis.

4.1. Observation ErrP

The first goal of the work was to infer whether the emotional facial paradigm would elicit an oErrP. The grand-average of ERPs evoked by wrong agent's decisions clearly show an oErrP with components distinct of cERPs. The ERN component at FCz occurred at about 215 ms after the cue image movement followed by the Pe component around 359 ms. The differences between oErrP and cERPs are statistically significant around the ERN and Pe, evidencing the two different neural responses associated to incorrect and correct actions identification. The oErrP waveforms are consistent with previous oErrP literature, exhibiting a negative deflection followed by a positive deflection, although they are known to typically occur with different waveforms, amplitudes, and latencies depending on the task. For example, in [47], one of the first studies researching error processing associated to observation, authors reported the ERN during an action observation task at around 250 ms and Pe at 375 ms. In [31], in a square movement paradigm, the peaks of the difference between error and correct trials (only this difference is provided) occur at 260 ms (negative peak) and 350 ms (positive peak). In a robot gesture learning task [33], the negative peak of the oErrP occurs around 350 ms and the positive peak around 500 ms. Pezzetta at al [48] reported ERN between 300 and 400 ms and Pe between 400 and 800 ms, in a reach-to-grasp task performed by an avatar in a virtual environment.

Besides the classically described ErrPs, theta rhythms ([4 8] Hz) have also been associated to error-monitoring processes. They have been reported to increase when the user commits errors or receive erroneous feedback [24, 49, 50] but also while observing erroneous actions [48]. They can be atypical in ASD [51] and other psychiatric conditions [52], and thus may contribute to a better profiling of ASD groups. The time-frequency analysis shows an increase of the theta rhythm when participants observe an error committed by the intelligent agent, which reinforces the idea that theta rhythms are also associated to observed errors. This can be used in a future BCI version as a complement measure to obtain neurophysiological biomarkers related to ASD but also to improve error classification [53].

4.2. oErrP classification

The oErrP balanced classification accuracies obtained online with and without performance feedback (green bar progression) were respectively 77.1% and 81.6% (recorded from only 16 channels). The online results are at the same level of those obtained offline from the calibration data, showing the effectiveness of the online methodological classification procedure. These good classification accuracies match the statistically significant differences obtained between oError and cERPs in the two time-windows of interest, as observed in figures 4 and 5, and show the viability of the approach to be used as an assistive rehabilitation tool in ASD. The results obtained with the ASD participant reached the average of the control group. The inclusion in the study of a patient with ASD was intended to help further consolidate the feasibility of the paradigm approach, namely, to infer whether a representative ASD patient could understand and control the RL-based paradigm, and at the same time note his motivation, reaction to the interface, and engagement with the gamified approach. The good performance and engagement achieved by this high functioning ASD participant does not show that other types of ASD participants with a lower intellectual level will have the same level of performance and will require generalization in future studies.

It does however provide increased confidence in the feasibility of the methodology and its possible use in ASD cohorts. The good overall results of the classification are an important achievement as they represent a primary goal of this study in order to be able to move on to the clinical phase. A low detection accuracy has a negative impact on the convergence rate of the intelligent agent learning and might lead to participants' frustration, derailing the playability of the gamified-RL approach. Moreover, the use of non-volitional neurofeedback approaches depends heavily on the detection accuracy, as also emphasized in [36].

It is important to note that although we have ensured that participants are already fixating on the target red spot when the image movement occurs, some gaze shifts may still happen, in particular when the image movement diverges of the target location. However, given the balance of targets appearing at left or right, the ocular movement patterns should not bias the classification for any of the classes.

oErrPs are classified at a single-trial level, which is a very challenging task due to their very low SNR. For example, the measured SNR of the ErrPs at Fz and Cz were respectively −12.7 and −12.9 dBs taking the [200 500] ms time window from calibration data. The classification results compare favorably to other studies using oErrPs such as in [31] where 69.5% was achieved pseudo-online with a group of six healthy participants, using downsampling and a Gaussian filter for classification. Higher classification accuracies were obtained on interaction-ErrPs (rounding 85%–90%) in our own previous studies [28, 54], applying the same approach used here, and in other studies [29] based on stepwise linear discriminant analysis regression and random forest. Yet, it is well known that results are very dependent on the type of ErrP and the task itself which may elicit different levels of cognitive processes, awareness and engagement [55]. For example, in [33], using 64 channels the authors attained a 91% accuracy detecting oErrPs in a human-robot interaction task, using the xDAWN spatial filter [56] combined with a data augmentation approach to which the authors attributed the very good classification accuracy. However, when replicating such data augmentation approach in our own data, the results did not match those obtained in [33].

Two modalities, with and without performance feedback, were tested in order to evaluate the impact of feedback. The classification accuracy was on average 4.5% ($p = 0.009$) higher when the green progress bar was not provided, although the differences in ERP waveforms between game mode 1 and mode 2 were not statistically different. This performance decrease suggests a negative effect of using the green bar feedback. In debriefings done with participants, some of them reported that when they believed they did everything well (i.e. correctly recognized the facial emotion and identified the correctness of the movement of the face) and the green bar decreased rather than increasing (i.e. wrong classifier detection), this caused some frustration and focus disruption during the experiment. In short, the green bar seems to be working as an attentional distracter. This was stated by some of the participants including the ASD participant. To further analyze this classification decrease, we computed the detection rate on the subsequent trial after a decrease in green bar and compared it with the detection rate obtained when the green bar increased. The classification error rates were respectively 33.3% and 22.8 %, which suggests a high influence of the negative performance feedback on the subsequent trial. Additionally, we compared the oErrP waveforms in trials after correct performance feedback and after wrong performance feedback as shown in figure 8. The plots show that the two waveforms substantially differ (Pearson correlation equal to 0.53), and that the negative and positive peaks are much less accentuated and with smaller amplitudes when the participant receives a negative feedback, which sustains the influence of the feedback as an attentional distractor. This result highlights the importance of carefully analyzing feedback, for example, assessing the balance between 'distractor vs engagement' effects. Some modifications to the paradigm may have to be considered as discussed in section 4.4. Under the conditions of our experiment, the green bar feedback does not seem to have a positive impact on our non-volitional neurofeedback paradigm. The effect and use of neurofeedback has been a topic of research [57, 58], though not focused on non-volitional neurofeedback.

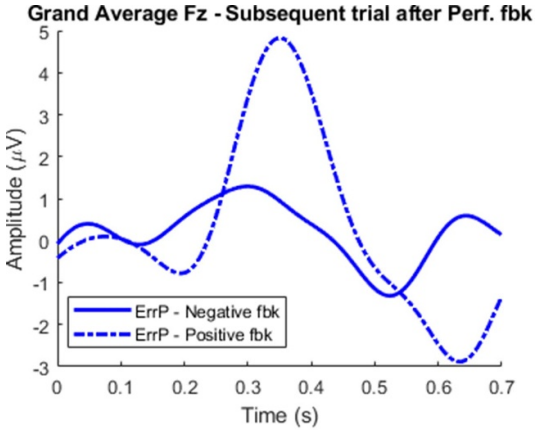


Figure 8. Comparison of oErrPs elicited on the subsequent trial after positive and negative performance feedback.

4.3. RL of the agent

Regarding the hypothesis of the intelligent agent learning using the ErrP biofeedback as reward, we found that all participants were able to reach the optimal strategy at least in one of the attempts. On average, the optimal strategy was reached and maintained in 3.8 session blocks over the 5 without performance feedback and 3.4 over the 5 session blocks

with performance feedback. A post-hoc analysis of the data showed that even when the optimal strategy was not reached, it was approximated in many blocks of the session. In particular, we verified that in 96% of the blocks the learning convergence curve exceeded 90%. In testing sessions debriefings, these RL convergence curves were shown to participants. At this stage, we only sought to informally: (a) make participants aware of their progression or fluctuation throughout the session block; (b) motivate them, showing, for example, that they had a good overall performance, despite the fact that they did not reach 100%; and (c) evaluate their response when they realized they have reached the optimal strategy. When considering this experimental approach as an assistive neurorehabilitation tool for ASD, these aspects are of paramount importance as they are expected to provide participants the sense of reward and the feeling of being able to 'teach' a machine on their own, aiming at increasing their motivation and sense of control [42, 59]. The use of RL in the form of a game plays here an important role as it gives participants a very intuitive way of showing that their incremental 'teaching' correlated to their performance. Error/performance monitoring is essential for making decisions, handling errors, adjustment for adaptive and goal-directed behavior, and learning [26, 60]. Particularly in ASD, impairments in error/performance monitoring have been linked to restricted patterns of behaviors and abnormal emotional reactions to errors. When participants are asked to judge the actions of the artificial agent to make the BCI system learn, they are training their neural circuits associated with error monitoring. Interestingly, they are doing this without explicit volitional control strategy, because their goal is to teach the BCI system just by attending to the occurrence of errors. Thus, there is an implicit (and unconscious) cognitive training despite the participant not being directly controlling the progression of the bar, which grounds the rational of the approach also as a non-volitional neurofeedback tool for cognitive/behavioral training. In [61] a non-volitional neurofeedback approach has shown the possibility of changing relevant connectivity patterns in participants with ASD. Although tested with functional magnetic resonance imaging (fMRI), there is a high interest in transferring similar approaches to EEG.

From the perspective of computational implementation of the RL, we added the possibility of adjusting the learning rate through a subject-dependent parameter related to the participant's classification performance (equation (3)—obtained from calibration data). This individual parameter increases the certainty of rules update, which is an add-on for the consistency of the practical implementation of the approach. Looking at the results in table 1, we observe that convergence was biased by the classifier specificity, as the number of correct events is higher. This can be further adjusted in the future during calibration. Overall, it is shown that the possibility for the agent to learn according to the participant's performance is displayed in a logical and intuitive way enough for the participant to understand it as a game, and thereby, that the developed framework is suitable for the next step of clinical trials with a group of ASD participants.

4.4. Improvement perspectives for the next experimental stage

The hypothesis that the training approach will allow not only the intelligent agent to learn but also that it will be possible to train error-monitoring networks with a potential improvement in motivation and anxiety levels of ASD participants, is still to be validated. This will be carried out at the next experimental stage in a longitudinal clinical trial with several composite endpoints (with a primacy of behavioral measures as primary outcomes), conducted with a group of patients with ASD, taking advantage of the experience already gained in [11]. The effect of cognitive training will be assessed considering neurophysiological and behavioral outcomes. For example, improvement in classification accuracy is an indicator that participants are correctly identifying agent's errors, providing a direct, quantitative measure that cognitive processes (e.g. error monitoring and selective attention) are improving. Behavioral outcomes will be assessed through interviews and questionnaires measuring the efficacy of the intervention in terms of levels of anxiety, motivation, physiological arousal, and adaptive behavior. This will be rated using the Hamilton Anxiety Rating Scale, the Intrinsic Motivation Inventory, the Autism Treatment Evaluation Checklist, the Vineland Adaptive Behavior Scale and the Facial Expressions of Emotion: Stimuli and Tests.

Despite the very promising results achieved in this study with the current experiment design, adjustment of some parameters of the gamified-RL approach are likely necessary. Namely, one should consider the possibility of providing participants with only positive rewards or making reward decisions only when the classification confidence is higher than a given threshold, thus avoiding possible causes of participant frustration. The visual information to provide performance feedback can also be changed to be more engaging, for example, providing as feedback the score level of the classifier in addition to the greenbar. In addition, the inclusion of eye-tracking can be considered to be used as ground-truth to disambiguate whether the error came from the classifier or from the participant, or to infer whether there may have been participant hesitation when making the decision.

Eventually, we may find necessary to reduce the calibration time or to obtain session-independent models. The calibration time can be reduced by increasing the error probability to obtain more error

samples for training. It has been shown in one of our previous studies [62] that the classification models built from a high error-rate calibration generalize to sessions with lower error-rates. In [48], the authors also showed that observation ErrPs were triggered even when the probability of wrong actions was higher than correct actions. In the context of session-independent models for ErrPs and P300 ERPs, several promising transfer learning solutions have been proposed based on Riemannian geometry [54] and deep learning approaches using convolutional neural networks (CNNs) and variants such as EEGNet [63] and interpretable CNN [64] which have shown to outperform classical approaches [65].

5. Conclusion

We proposed a gamified BCI combining an EFP and machine RL, which was very successful in eliciting ErrP signals which were detected with a high classification accuracy and showing the effective possibility of making an intelligent agent learn. The BCI, based on non-volitional neurofeedback, can be used simultaneously for error-monitoring and as a promising tool for cognitive training in ASD while increasing motivated behavior, facial expression recognition, with positive effects on anxiety. These achievements were of utmost importance, allowing to prepare a future clinical testing phase with ASD population. Further conclusions about its effective use as a training tool, as well as enhancements that can be applied regarding the user experience and engagement during the task and its role in a co-adaptation neurofeedback scenario, will be possible at a clinical intervention research stage.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

We want to thank Susana Mouga for her kind help in the ASD participant's characterization. This work has been financially supported by Portuguese foundation for science and technology (FCT) under Grants B-RELIABLE: PTDC/EEI-AUT/30935/2017 and BCI-CONNECT: PTDC/PSI-GER/30852/2017, and by ISR-UC UIDB/00048/2020 and CIBIT UIDB &P /4950/2020 Grants.

CRediT author statement

G P: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing—original draft, Writing—review & editing, Supervision. A C: Conceptualization, Methodology, Software, Investigation. D J: Conceptualization, Methodology, Software, Investigation, Formal analysis, Writing—original draft, Writing—review & editing. M Y: Methodology, Software. U N: Writing—review & editing, Supervision. T S: Conceptualization, Methodology, Investigation, Formal analysis, Writing—original draft, Writing—review & editing. M C-B: Conceptualization, Methodology, Formal analysis, Writing—review & editing, Supervision.

Conflict of interest

The authors declare no competing interests.

ORCID iDs

Gabriel Pires https://orcid.org/0000-0001-9967-845X
Mine Yasemin https://orcid.org/0000-0002-4142-7487
Urbano J Nunes https://orcid.org/0000-0002-7750-5221
Teresa Sousa https://orcid.org/0000-0003-2652-3152
Miguel Castelo-Branco https://orcid.org/0000-0003-4364-6373

References

[1] Chaudhary U, Birbaumer N and Ramos-Murguialday A 2016 Brain-computer interfaces for communication and rehabilitation *Nat. Rev. Neurol.* **12** 513–25
[2] Lebedev M A and Nicolelis M A L 2017 Brain-machine interfaces: from basic science to neuroprostheses and neurorehabilitation *Physiol. Rev.* **97** 767–837
[3] Lim C G, Lee T S, Guan C, Sheng Fung D S, Zhao Y, Wei Teng S S, Zhang H and Rama Krishnan K R 10 2012 A brain-computer interface based attention training program for treating attention deficit hyperactivity disorder *PLoS One* **7** 1–8
[4] Badia S, Morgade A, Samaha H and Verschure P 2013 Using a hybrid brain computer interface and virtual reality system to monitor and promote cortical reorganization through motor activity and motor imagery training *IEEE Trans. Neural Syst. Rehabil. Eng.* **21** 174–81
[5] Lee T-S *et al* 11 2013 A brain-computer interface based cognitive training system for healthy elderly: a randomized control pilot study for usability and preliminary efficacy *PLoS One* **8** 1–8
[6] Si Ning Y *et al* 2018 Effectiveness of a personalized brain-computer interface system for cognitive training in healthy elderly: a randomized controlled trial *J. Alzheimer's Dis.* **66** 127–38
[7] Vourvopoulos A, Jorge C, Abreu R, Figueiredo P, Fernandes J C, Badia I Bermúdez S 2019 Efficacy and brain imaging correlates of an immersive motor imagery BCI-driven VR system for upper limb motor rehabilitation: a clinical case report *Front. Hum. Neurosci.* **13** 1
[8] Greene D J, Colich N, Iacoboni M, Zaidel E, Bookheimer S Y and Dapretto M 2011 A typical neural networks for social orienting in autism spectrum disorders *NeuroImage* **56** 354–62
[9] Friedrich E V C, Suttie N, Sivanathan A, Lim T, Louchart S and Pineda J A 2014 Brain-computer interface game applications for combined neurofeedback and biofeedback treatment for children on the autism spectrum *Front. Neuroeng.* **7** 21

[10] Amaral C P, Simões M A, Mouga S, Andrade J and Castelo-Branco M 2017 A novel brain computer interface for classification of social joint attention in autism and comparison of 3 experimental setups: a feasibility study *J. Neurosci. Methods* **290** 105–15

[11] Amaral C, Mouga S, Simões M, Pereira H C, Bernardino I, Quental H, Playle R, McNamara R, Oliveira G and Castelo-Branco M 2018 A feasibility clinical trial to improve social attention in autistic spectrum disorder (ASD) using a brain computer interface *Front. Neurosci.* **12** 477

[12] Monteiro R, Simões M, Andrade J and Castelo-Branco M 2017 Processing of facial expressions in autism: a systematic review of EEG/ERP evidence *Rev. J. Autism Deve. Disorders* **4** 1–22

[13] American Psychiatric Association 2013 *Diagnostic and Statistical Manual of Mental Disorders* 5th edn (Arlington, VA: American Psychiatric Publishing) (https://doi.org/10.1176/appi.books.9780890425596)

[14] Goldberg M C, Spinelli S, Joel S, Pekar J J, Denckla M B and Mostofsky S H 2011 Children with high functioning autism show increased prefrontal and temporal cortex activity during error monitoring *Dev. Cogn. Neurosci.* **1** 47–56

[15] Thakkar K N, Polli F E, Joseph R M, Tuch D S, Hadjikhani N, Barton J J S and Manoach D S 2008 Response monitoring, repetitive behaviour and anterior cingulate abnormalities in autism spectrum disorders (ASD) *Brain* **131** 2464–78

[16] Vlamings P H J M, Jonkman L M, Hoeksma M R, Engeland H V and Kemner C 2008 Reduced error monitoring in children with autism spectrum disorder: an ERP study *Eur. J. Neurosci.* **28** 399–406

[17] South M, Larson M J, Krauskopf E and Clawson A 2010 Error processing in high-functioning autism spectrum disorders *Biol. Psychol.* **85** 242–51

[18] McMahon C M and Henderson H A 2015 Error-monitoring in response to social stimuli in individuals with higher-functioning autism spectrum disorder *Dev. sci.* **18** 389–403

[19] Falkenstein M, Hoorman J, Christ S and Hohnsbein J 2000 ERP components on reaction errors and their functional significance: a tutorial *Biol. Psychol.* **51** 87–107

[20] Wessel J R 2012 Error awareness and the error-related negativity: evaluating the first decade of evidence *Front. Hum. Neurosci.* **6** 88

[21] Pavone E F, Tieri G, Rizza G, Tidoni E, Grisoni L and Aglioti S M 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** 268–79

[22] Santesso D L, Drmic I E, Jetha M K, Bryson S E, Goldberg J O and Hall G B 2011 An event-related source localization study of response monitoring and social impairments in autism spectrum disorder *Psychophysiology* **48** 241–51

[23] Estiveira J, Dias C, Costa D, Castelhano J, Castelo-Branco M and Sousa T 2022 An action-independent role for midfrontal theta activity prior to error commission *Front. Hum. Neurosci.* **16** 805080

[24] Dias C, Costa D, Sousa T, Castelhano J, Figueiredo V, Pereira A C and Castelo-Branco M 2022 A neuronal theta band signature of error monitoring during integration of facial expression cues *PeerJ* **10** e12627

[25] Holroyd C B and Coles M G H 2002 The neural basis of human error processing: reinforcement learning, dopamine and the error-related negativity *Psychol Rev.* **109** 679–709

[26] Schie H, Mars R B, Coles M G H and Bekkering H 2004 Modulation of activity in medial frontal and motor cortices during error observation *Nat. Neurosci.* **7** 549–54

[27] Ferrez P W and Millan J del R 2008 Error-related EEG potentials generated during simulated brain-computer interaction *IEEE Trans. Biomed. Eng.* **55** 923–9

[28] Cruz A, Pires G and Nunes U J 2018 Double ErrP detection for automatic error correction in an ERP-based BCI speller *IEEE Trans. on Neur. Syst. and Rehab. Eng.* **26** 26–36

[29] Zeyl T, Yin E, Keightley M and Chau T 2016 Adding real-time Bayesian ranks to error-related potential scores improves error detection and auto-correction in a p300 speller *IEEE Trans. on Neur. Syst. and Rehab. Eng.* **24** 46–56

[30] Spüler M, Bensch M, Kleih S, Rosenstiel W, Bogdan M and Kübler A 2012 Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-BCI *Clin. Neurophysiol.* **123** 1328–37

[31] Chavarriaga R and Millan J del R 2010 Learning from EEG error-related potentials in noninvasive brain-computer interfaces *IEEE Trans. on Neur. Syst. and Rehab. Eng.* **18** 381–8

[32] Iturrate Iñaki, Chavarriaga R, Montesano L, Minguez J and Millán J del R 2015 Teaching brain-machine interfaces as an alternative paradigm to neuroprosthetics control *Sci. Rep.* **5** 13893

[33] 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** 17562

[34] Batzianoulis I, Iwane F, Wei S, Ramos Correia C G P, Chavarriaga R, Millán J del R and Billard A 2021 Customizing skills for assistive robotic manipulators, an inverse reinforcement learning approach with error-related potentials *Commun. Biol.* **4** 2399–3642

[35] Ehrlich S K and Cheng G 2018 Human-agent co-adaptation using error-related potentials *J. Neural Eng.* **15** 066014

[36] Taschereau-Dumouchel V, Cortese A, Lau H and Kawato M 2020 Conducting decoded neurofeedback studies *Soc. Cogn. Affect. Neurosci.* **16** 838–48

[37] Ramot M and Martin A 2022 Closed-loop neuromodulation for studying spontaneous activity and causality *Trends Cogn. Sci.* **26** 290–9

[38] Shibata K, Watanabe T, Kawato M and Sasaki Y 2016 Differential activation patterns in the same brain region led to opposite emotional states *PLoS Biol.* **14** 1–27

[39] Lord C, Rutter M and Ann L C 1994 Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders *J. Autism Dev. Disorders* **24** 659–85

[40] Lord C *et al* 1989 Austism diagnostic observation schedule: a standardized observation of communicative and social behavior *J. Autism Dev. Disorders* **19** 185–212

[41] Langner O, Dotsch R, Bijlstra G, Wigboldus D H, Hawk S T and Van Knippenberg A D 2010 Presentation and validation of the Radboud faces database *Cogn. Emotion* **24** 1377–88

[42] Direito B *et al* 2021 Training the social brain: clinical and neural effects of an 8-week real-time functional magnetic resonance imaging neurofeedback phase IIA clinical trial in autism *Autism* **25** 1746–60

[43] Simões M, Monteiro R, Andrade J, Mouga S, França F, OLiveira F, Carvalho P and Castelo-Branco M 2018 A novel biomarker of compensatory recruitment of face emotional imagery networks in autism spectrum disorder *Front. Neurosci.* **12** 791

[44] Pires G, Nunes U and Castelo-Branco M 2011 Statistical spatial filtering for a P300-based BCI: tests in able-bodied and patients with cerebral palsy and amyotrophic lateral sclerosis *J. Neurosci. Methods* **195** 270–81

[45] Pires G 2012 Biosignal classification for human interface with devices and surrounding environment *PhD Thesis* Department of Electrical and Computer Engineering University of Coimbra

[46] Delorme A, Palmer J, Onton J, Oostenveld R and Makeig S 2012 Independent EEG sources are dipolar *PLoS One* **7** 1–14

[47] Schie H T van, Mars R B, Coles M G H and Bekkering H 2004 Modulation of activity in medial frontal and motor cortices during error observation *Nat. Neurosci.* **7** 549–554

[48] Pezzetta R, Nicolardi V, Tidoni E and Aglioti S M 2018 Error, rather than its probability, elicits specific electrocortical signatures: a combined EEG-immersive virtual reality study of action observation *J. Neurophysiol.* **120** 1107–18

[49] 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
[50] Cavanagh J F and Frank M J 2014 Frontal theta as a mechanism for cognitive control *Trends Cogn. Sci.* **18** 414–21
[51] Buzzell G A, Thomas H R, Choi Y B and Kim S H 2021 Atypical mediofrontal theta oscillations underlying cognitive control in kindergarteners with autism spectrum disorder *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **7** 566–75
[52] McLoughlin G, Gyurkovics Mé, Palmer J and Makeig S 2022 Midfrontal theta activity in psychiatric illness: an index of cognitive vulnerabilities across disorders *Biol. Psychiatry* **91** 173–82
[53] Dias C, Costa D M, Sousa T, Castelhano J, Figueiredo V, Pereira A C and Castelo-Branco M 2021 Classification of erroneous actions using EEG frequency features: implications for BCI performance *2021 43rd Annual Int. Conf. IEEE Engineering in Medicine Biology Society (EMBC)* pp 629–32
[54] Cruz A, Pires G and Nunes U J 2022 Spatial filtering based on Riemannian distance to improve the generalization of ErrP classification *Neurocomputing* **470** 236–46
[55] Chavarriaga R, Sobolewski A and Millán J del R 2014 Errare machinale est: the use of error-related potentials in brain-machine interfaces *Front. Neurosci.* **8** 208
[56] 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** 2035–43
[57] Mladenovic J, Frey J, Pramij S, Mattout J and Lotte F 2022 Towards identifying optimal biased feedback for various user states and traits in motor imagery BCI *IEEE Trans. Biomed. Eng.* **69** 1101–10
[58] Direito B, Ramos M, Pereira J, Sayal A, Sousa T and Castelo-Branco M 2021 Directly exploring the neural correlates of feedback-related reward saliency and valence during real-time fMRI-based neurofeedback *Front. Hum. Neurosci.* **14** 578119
[59] Cramer S C 2011 Harnessing neuroplasticity for clinical applications *Brain* **134** 1591–609
[60] Ullsperger M, Danielmeier C and Jocham G 2014 Neurophysiology of performance monitoring and adaptive behavior *Physiol. Rev.* **94** 35–79
[61] Ramot M, Kimmich S, Gonzalez-Castillo J, Roopchansingh V, Popal H, White E, Gotts S J and Martin A 2017 Direct modulation of aberrant brain network connectivity through real-time neurofeedback *eLife* **6** e28974
[62] Cruz A, Pires G and Nunes U J 2018 Generalization of ErrP-calibration for different error-rates in P300-based BCIs *2018 IEEE Int. Conf. on Systems, Man and Cybernetics (SMC)* pp 644–9
[63] Lawhern V J, Solon A J, Waytowich N R, Gordon S M, Hung C P and Lance B J 2018 EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces *J. Neural Eng.* **15** 056013
[64] Borra D, Magosso E, Castelo-Branco M and Simões M 2022 A Bayesian-optimized design for an interpretable convolutional neural network to decode and analyze the P300 response in autism *J. Neural Eng.* **19** 046010
[65] Simões M 2020 BCIAUT-P300: a multi-session and multi-subject benchmark dataset on autism for P300-based brain-computer-interfaces *Front. Neurosci.* **14** 568104