OPEN

Online asynchronous decoding of error-related potentials during the continuous control of a robot

Catarina Lopes-Dias [ID], Andreea I. Sburlea [ID] & Gernot R. Müller-Putz [ID]*

Error-related potentials (ErrPs) are the neural signature of error processing. Therefore, the detection of ErrPs is an intuitive approach to improve the performance of brain-computer interfaces (BCIs). The incorporation of ErrPs in discrete BCIs is well established but the study of asynchronous detection of ErrPs is still in its early stages. Here we show the feasibility of asynchronously decoding ErrPs in an online scenario. For that, we measured EEG in 15 participants while they controlled a robotic arm towards a target using their right hand. In 30% of the trials, the control of the robotic arm was halted at an unexpected moment (error onset) in order to trigger error-related potentials. When an ErrP was detected after the error onset, participants regained the control of the robot and could finish the trial. Regarding the asynchronous classification in the online scenario, we obtained an average true positive rate (TPR) of 70% and an average true negative rate (TNR) of 86.8%. These results indicate that the online asynchronous decoding of ErrPs was, on average, reliable, showing the feasibility of the asynchronous decoding of ErrPs in an online scenario.

Brain-computer interfaces (BCIs) are systems that measure brain activity, often using electroencephalography (EEG), and convert it into actions of an external device[1]. As BCIs enable communication without movement, they are a valuable tool to provide more independence to people with severe motor disabilities[2–4].

The main obstacle to the widespread use of BCIs is their non-optimal performance, which sometimes leads to a misinterpretation of the user's intention and a consequent execution of a wrong action. The user's experience with the BCI can be spoiled by occurrence of many mistakes or by the effort to correct them.

The user's awareness of the committed mistake is associated with a neural pattern named error-related potential (ErrP). ErrPs occur both in humans and in monkeys and can be measured using several imaging techniques[5–12]. Additionally, ErrPs morphology is comparable in humans with and without spinal cord injury[13]. ErrPs are related with conflict monitoring[14] and have been reported in association with the awareness of self-committed mistakes, observed mistakes of another person or agent, and BCI's mistakes[14–17].

The use of ErrPs is an intuitive approach to improve BCIs' performance, either in a corrective manner, by allowing the BCI to take a corrective action, or in an adaptive manner, by reducing the possibility of future errors[18–20].

The detection of ErrPs in a time-locked manner is well established[21–23], and it has been extensively applied in discrete BCIs, whose actions occur in a discrete manner, allowing users to interact with a computer or with a robot[24–31].

Recently, an effort has been made to develop BCIs that provide a more intuitive control to the user, by e.g., providing continuous control to the user[32–34]. In this situation, the user can perceive, at any moment, that an error occurred. This possibility triggered the research on the asynchronous detection of ErrPs[35–39].

In the current study, we investigate the feasibility of the online asynchronous ErrPs' detection, while participants continuously controlled a robotic arm towards a target, using their right hand. In 30% of the trials, the user's control of the robot was halted at random point. Participants could regain the robot's control if an ErrP was detected after the error onset. To our knowledge, this is the first report of online asynchronous detection of ErrPs.

Materials and Methods

Participants. 15 right-handed volunteers (5 women) participated in the experiment. All participants had normal or corrected-to-normal vision and had no history of brain disorders. The participants were, on average, 23.5 ± 2.5 years old (mean ± std). Participants were paid 7.50 euros per hour, were explained the experimental

Institute of Neural Engineering, Graz University of Technology, Graz, Austria. *email: gernot.mueller@tugraz.at

Figure 1. Experimental setup. In this figure, the robot is at its home position. The squares on the screen represent the physical targets (violet cuboids) on the wooden structure. The rectangle on the bottom part of the screen represents the participant's hand home position and the text above it states 'Bring your hand to the home position'. Inside the wooden structure, there is a Leap Motion device (not visible here) used to track the participants' right hand movement.

protocol and signed an informed consent form that had been previously approved by the local ethics committee of the Medical University of Graz (Ethical approval number 30-275 17/18). The experiment was performed in accordance with the Declaration of Helsinki.

Hardware and measuring layout. We recorded EEG and EOG data at a samplig rate of 500 Hz using BrainAmp amplifiers and an ActiCap cap (Brain Products, Munich, Germany). We used 61 EEG electrodes and 3 EOG electrodes. The EOG electrodes were placed above the nasion and below the outer canthi of the eyes. The ground electrode was placed at position AFz and the reference electrode was placed on the right mastoid. The layout of the EEG electrodes is described in Fig. 1 of the Supplementary Material.

Experiment layout. Figure 1 depicts the physical layout of the experiment. Participants sat on a chair in front of a table. On the table was a wooden structure: 4-sided box, with open sides towards the participants and the tabletop. On the ceiling of the structure was a Leap Motion device (Leap Motion, San Francisco, United States), used to track the participants' right hand (not visible in Fig. 1). The participants kept their right hand lying on the table, inside the wooden structure. This setup occluded the participants' hand from their field of view. On the right side of the participants, we placed a robotic arm (Jaco Assistive robotic arm - Kinova Robotics, Bonn, Germany). On the wooden structure, were placed two physical targets: violet cuboids with a square base of 14 cm side. The centres of the targets were 35 cm apart and their mid-point was located 30 cm in front of the home position of the robot's hand, as shown in Fig. 1. Behind the structure, within the participant's line of sight to the targets, was a monitor that displayed information regarding the experiment. The participant shown in Fig. 1 gave her informed consent for the photo to be made available in an open-access publication.

Controlling the robot. During the *trials*, participants could control the position of the robot's hand on a horizontal plane, by moving their hand on the table, which was tracked with the Leap Motion. To reduce the range of the participants' movements, we considered the robot's hand displacement to be three times larger than the participants' hand displacement.

Experiment overview. Before the experiment, two blocks in which participants performed eye movements were recorded. The experiment then consisted of 12 blocks of 30 trials each. 70% of the trials of each block were *correct trials* (21 trials) and the remaining 30% were *error trials* (9 trials).

The sequence of correct and error trials within each block was randomly generated using a uniform distribution to place the error trials. We defined a maximum of 2 consecutive error trials in each block and repeated the randomization procedure until the sequence of trials satisfied this condition.

Half of the trials in each block were associated with the left target and the remaining trials with the right target. The sequence of targets within each block was randomly assigned using a uniform distribution. We defined a maximum of 3 consecutive trials with the same target in each block and repeated the randomization procedure until the targets' sequence satisfied this requirement.

Pre-trial. During the pre-trial period, the monitor displayed information regarding the coming trial. As depicted in Figs. 1 and 2, on the top part of the screen were displayed two squares representing the targets lying on the wooden structure. One of the squares was filled in white and the other had no fill. The filled square indicated the selected target for the coming trial. On the bottom part of the screen was a rectangle, representing the home
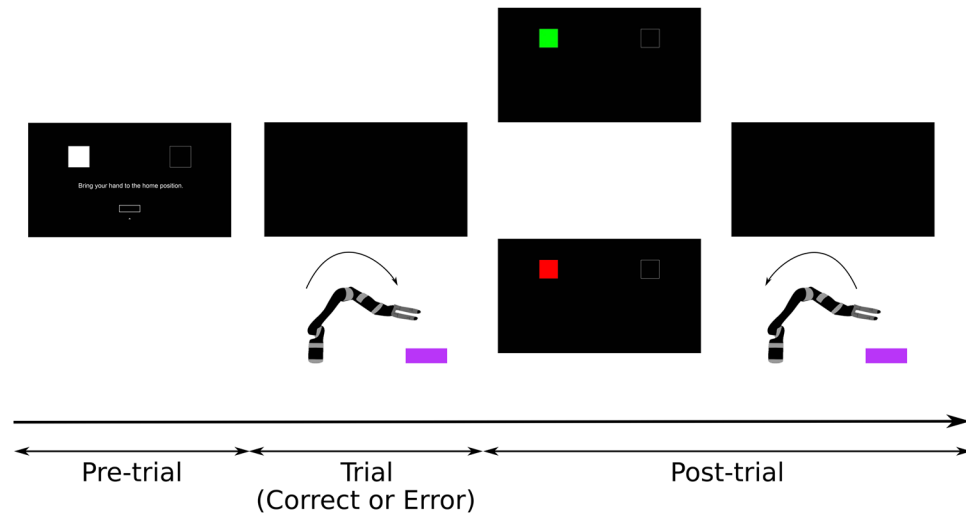
Figure 2. Experimental protocol. During the pre-trial period, participants could rest for as long as they wished. The pre-trial period ended and a new trial started when the participants moved their hand to its home position. During the trials, the screen was black. Participants were instructed to bring the robot's hand to the selected target. A trial finished either when the robot reached the target or after 6 seconds, in case target was not reached. Afterwards (post-trial period), the squares reappeared on the screen for 1.2 seconds and gave feedback regarding hitting the target: a green square indicated that the target was hit and a red square indicated that the target was not hit. Then, the screen turned black, the robot automatically returned to its home position and a new pre-trial period started.

position of the participant's hand. The position of the participant's hand was depicted by a dot on the screen. The trial would start when the dot entered the rectangle. This ensured that the participant's hand was at a comparable position at the beginning of each trial (within a 1×3 cm rectangle).

Participants could use the pre-trial period to rest for as long as they needed. When participants felt ready to start the trial, they had to bring their hand below the home position, fixate their gaze on the physical target and finally enter the rectangle from below. This final step ensured a forward movement of the robot. Participants were also instructed to keep their gaze fixed at the target during the entire trial duration, in order to prevent eye movements.

Trials. The aim of each trial was to bring the robot's hand from its home position to the selected target. During the trials, the screen was black. A trial ended when the robot's hand was above the intended target (hit) or after 6 seconds (no hit). Afterwards (post-trial period), as shown in Fig. 2, the two squares from the pre-trial period reappeared on the screen for 1.2 seconds and the filled square was now coloured in either green (hit) or red (no hit). This feedback was always in line with the behaviour of the robot. Then, the screen would turn black, the robot would automatically return to its home position and a new pre-trial period would start.

Error Trials. During these trials, the paradigm triggered an *error*. The error consisted in interrupting the participants' control of the robot and adding a 5 cm upwards displacement to the robot's hand. Participants perceived the error by noticing the robot stopping and lifting. The errors occurred randomly, when the robot's hand was within 6 to 15 cm in the forward direction from its home position. This represents approximately 25 to 65% of the minimal forward displacement necessary for the robot to hit the target. For every error trial, we drew a value d_e from a uniform distribution U([6, 15]). The error was triggered when the robot's hand reached the distance d_e cm, in the forward direction, from its home position.

Correct Trials. In these trials, the paradigm did not trigger any error. Participants reached the selected target in $99.75 \pm 0.14\%$ of the correct trials (mean ± std). Correct trials lasted on average 2.02 ± 0.14 s (mean ± std). Correct trials were comparable in the *calibration and online parts of the experiment*.

Calibration and online parts of the experiment. The *calibration part* of the experiment comprised the first 8 blocks and the *online part* comprised the last 4 blocks. The calibration part was used to collect data to *train an ErrP classifier* and to find a *threshold* for the classifier. In the online part of the experiment, we tested the ErrP classifier, tuned with the calculated threshold, for the asynchronous detection of ErrPs.

For a matter of fluidity of the experiment, we decided not to give participants feedback of the false positive detections, i.e., of the *ErrP detections* when no error had occurred. Thus, from the participants' perspective, the online ErrP classifier had no effect on the correct trials and affected only the error trials. However, false positive detections were taken into account when evaluating the classifier.

Calibration error trials. In the error trials during calibration, when the error happened, the participants lost control of the robot, which remained still for the rest of the trial. The total trial duration was 6 seconds. Participants were instructed not to move until the trial ended.

Online error trials. In the error trials during the online part of the experiment, the participants had the possibility of correcting the robot's errors. If, after the error onset, an ErrP was detected by the ErrP classifier (true positive detection), the robot's hand lowered 5 cm and the participants regained control of the robot. The downward movement informed the the participants of the ErrP detection. Participants were instructed to move the robot's hand to the selected target when regaining control of the robot. To accommodate the extra movement, we added 6 seconds to the maximal trial duration when the first true positive detection occurred. When no true positive detection occurred, the robot remained still, as in the error trials during calibration.

Correct trials. For the participants, correct trials were identical in both the calibration and the online parts of the experiment, due to our decision of not giving feedback of the false positive detections in the online part of the experiment.

Data preprocessing. Eye movements and blinks were removed from the EEG data, using the data recorded right before the beginning of the experiment and using the subspace subtraction algorithm[40]. The EEG signal was then filtered between $[1, 10]$ Hz using a causal Butterworth filter of order 4.

Defining events. For the calibration error trials, we defined the error onset as the moment in which the robot started its upwards displacement. The error onset was individually calculated for every error trial, based on the robot's position. The average delay between the error marker and the error onset was 0.210 ± 0.004 s (mean $\pm$ std).

For the online error trials, we considered an average error onset, by adding the average delay of the robot, calculated from the calibration data (0.210 s), to the time of the error marker in every error trial. This aimed to compensate the less reliable error onset estimation in case an ErrP occurred between the error marker and the start of the robot's upwards displacement (false positive detection).

Correct trials were not associated with any intrinsic event. Therefore, we defined a virtual onset, occurring one second after the start of every correct trial. The virtual onset was chosen at a time-point in which errors could already occur in the error trials, in order to assure a comparable expectation in the participants.

Electrophysiological analysis. For the electrophysiological analysis, we considered an EEG epoch of 1.5 s from every trial. For the correct trials, we considered the interval $[-0.5, 1.0]$ s, time-locked to the virtual onset (0 s). For the calibration error trials, we considered the interval $[-0.5, 1.0]$ s, time-locked to the error onset (0 s). For the online error trials, we considered the interval $[-0.5, 1.0]$ s, time-locked to the average error onset (0 s).

Detection of error-related potentials. We used the data from the calibration part of the experiment to build an ErrP classifier that was tested asynchronously in the online part of the experiment.

Train an ErrP classifier. For every participant, we considered all trials from the calibration part of the experiment. We took, as features, the amplitudes of all 61 EEG channels at every time-point within a 450 ms window of every trial. The window started 300 ms after the error onset of error trials and 300 ms after the virtual onset of correct trials.

Next, in order to reduce the number of features, we performed principal component analysis (PCA) on the features, keeping the components that explained 99% of the data' variability. These components were then used as features to train a shrinkage-LDA classifier with two classes: error and correct. After PCA we kept, on average, 139.5 ± 13.5 features per participant (mean $\pm$ std). Figure 2 of the Supplementary Material depicts the grand-average original feature space in the time-spatial domain as well as the grand-average projection into the time-spatial domain of the features kept after PCA.

ErrP detection. The classifier was constructed to be evaluated in an asynchronous manner, using a sliding window, with a leap of 18 ms. The classifier's evaluation of each window resulted in the probability of the analysed window to belong to either class (correct or error). We defined an *ErrP detection* when two consecutive windows had a probability of belonging to the error class above a certain threshold τ.

Threshold τ for the ErrP classifier. The threshold τ was obtained using the calibration data and used in the online part of the experiment to tune the ErrP classifier.

In order to find the threshold that best suited each participant, we performed a 2×5-fold asynchronous cross-validation in the participant's calibration data, where we tested 41 thresholds: from 0 to 1 in steps of 0.025. We used a low number of repetitions in the cross-validation to promote a shorter duration of the experiment.

As evaluation metric for the asynchronous ErrP detection in the cross-validated data, we defined the true negative trials (TN trials) as the correct trials in which no ErrP was detected during the entire trial duration. We defined the true positive trials (TP trials) as the error trials in which no ErrP was detected before the error onset and at least one ErrP was detected within 1.5 s of the error onset.

Then, we calculated the average true negative rate (TNR) and the average true positive rate (TPR) for all the tested thresholds, based on the 10 iterations. The average TNR and average TPR were further smoothed using a moving average with 7 samples. The smoothed curves were named moving average TPR and moving average TNR.
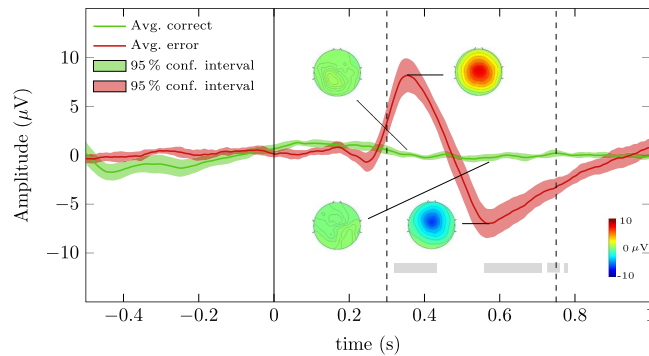
Figure 3. Grand average correct and error signals of the calibration part of the experiment at channel FCz (green and red solid lines, respectively). The green and red shaded areas represent the 95% confidence intervals of the grand average signals. The regions in which correct and error signals were significantly different are marked with a grey rectangle (Wilcoxon rank-sum tests, Bonferroni corrected, $p < 0.01$). The vertical line at $t = 0$ s represents the error onset of error trials and the virtual onset of correct trials. The dashed vertical lines at $t = 0.30$ s and $t = 0.75$ s delimit the window used to train the ErrP classifier.

For every participant, we considered the threshold that maximized performance to be the one that maximized the product of the moving average TPR and the moving average TNR. This threshold was then used in the online part of the experiment.

Online ErrP detection. The ErrP classifier was tested online in the last 4 blocks of the experiment. We decided to relax the evaluation metrics when testing the classifier online (in comparison with the metrics described for the cross-validated data) in order to consider the possible occurrence of secondary error-related potentials[28].

In the online evaluation, we defined the true negative trials (TN trials) as the correct trials in which no ErrP was detected (keeping the same definition used in the evaluation of the cross-validated data). Additionally we now defined the true positive trials (TP trials) as the error trials in which no ErrP was detected before the average error onset and at least one ErrP was detected after the average error onset.

A video of the online experiment can be seen in the Supplementary Material. The participant in the video gave her informed consent for it to be made available in an open-access publication.
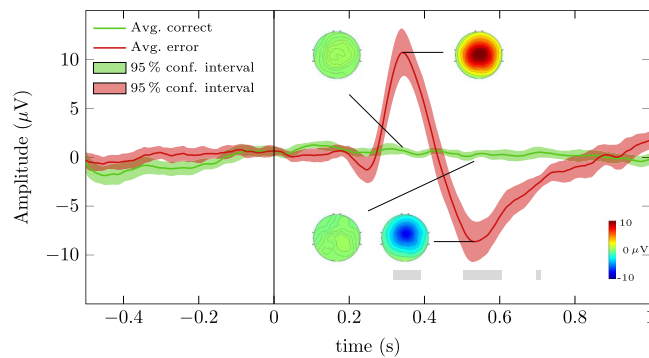
Results

Electrophysiology. *Calibration.* Figure 3 depicts the grand average correct and error signals of the calibration part of the experiment at channel FCz (green and red solid lines, respectively). The green and red shaded areas represent the 95% confidence intervals of the grand average signals. The time-intervals in which correct and error signals were significantly different ($t = [0.320, 0.432]$ s, $t = [0.558, 0.710]$ s, $t = [0.726, 0.760]$ s and $t = [0.770, 0.780]$ s) are represented by grey rectangles (Wilcoxon rank-sum tests, Bonferroni corrected, $p < 0.01$). The vertical line at $t = 0$ s represents the error onset for the error trials and the virtual onset for the correct trials. The error signal presents a small negativity with peak amplitude $-0.71\,\mu$V at 0.246 s, followed by a positivity with peak amplitude of 8.46 μV at 0.354 s, which is followed by a broader negativity with peak amplitude $-6.98\,\mu$V at 0.568 s. Figure 3 also depicts the topoplots of correct and error trials at the time-points $t = 0.354$ s and $t = 0.568$ s.

Online part. Figure 4 depicts the grand average correct and error signals of the online part of the experiment at channel FCz (green and red solid lines, respectively). The green and red shaded areas represent the 95% confidence intervals of the grand average signals. The time-intervals in which correct and error signals were significantly different ($t = [0.316, 0.390]$ s, $t = [0.504, 0.606]$ s and $t = [0.698, 0.710]$ s) are represented by grey shaded areas (Wilcoxon rank-sum tests, Bonferroni corrected, $p < 0.01$). The error signal presents a small negativity with peak amplitude $-1.29\,\mu$V at 0.246 s, followed by a positivity with peak amplitude 10.7 μV at 0.342 s and by a broader negativity with peak amplitude $-8.63\,\mu$V at 0.532 s. Figure 4 also depicts the topoplots of correct and error trials at the time-points $t = 0.342$ s and $t = 0.532$ s.

Asynchronous ErrP detection. *Offline asynchronous ErrP detection in the calibration data.* During the experiment, we performed asynchronous ErrP detection in the calibration data to find the threshold τ that was used online (using a 2×5-fold cross-validation to reduce the experiment duration, as described in section *Threshold τ for the ErrP classifier*).

For visualization purposes, here we present the asynchronous ErrP detection results, obtained using a 10×5-fold cross-validation in the calibration data, in which we tested 41 thresholds τ from 0 to 1, with steps of 0.025. The evaluation metric used to assess the results was the same as described in section *Threshold τ for the ErrP classifier*. Figure 5 displays the grand average TPR and TNR of the asynchronous classification performed using a 10×5-fold cross-validation in the calibration data (red and green solid lines, respectively), in function of the threshold τ. The chance-level TPR and TNR (red and green dashed lines, respectively) were obtained by performing the same classification procedure with randomly permuted training labels. The shaded green and red

Figure 4. Grand average correct and error signals of the online part of the experiment at channel FCz (green and red solid lines, respectively). The green and red shaded areas represent the 95% confidence interval for the grand average signals. The grey rectangles represent the time-intervals in which correct and error signals were significantly different (Wilcoxon rank-sum tests, Bonferroni corrected, $p < 0.01$). The vertical line represents the average error onset of the error trials and the virtual onset of the correct trials.
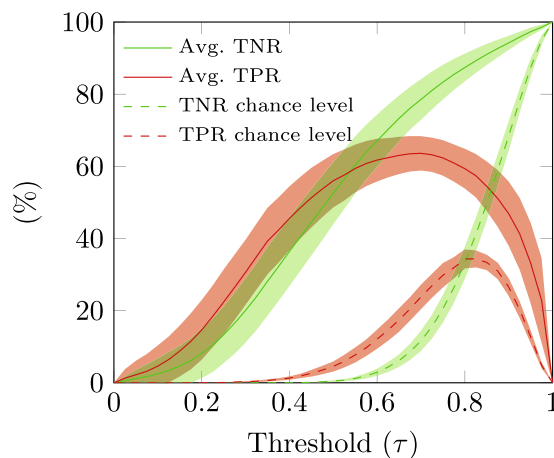
Figure 5. Asynchronous ErrP detection in the calibration data. Grand average TNR and TPR (solid green and red lines, respectively) in function of the threshold τ, calculated from the 10×5-fold cross-validation in the calibration data. The chance-level TPR and TNR are represented with red and green dashed lines. The shaded areas represent the 95% confidence intervals for the grand average curves.

areas represent the 95% confidence intervals of the grand average curves. The obtained TPR results were significantly higher than chance levels TPR results for thresholds $\tau \in [0.100, 0.975]$ (Wilcoxon rank-sum tests, one sided, Bonferroni corrected, $p < 0.01$). The obtained TNR results were significantly higher than chance level TNR results for thresholds $\tau \in [0.150, 0.975]$ (Wilcoxon rank-sum tests, one sided, Bonferroni corrected, $p < 0.01$).
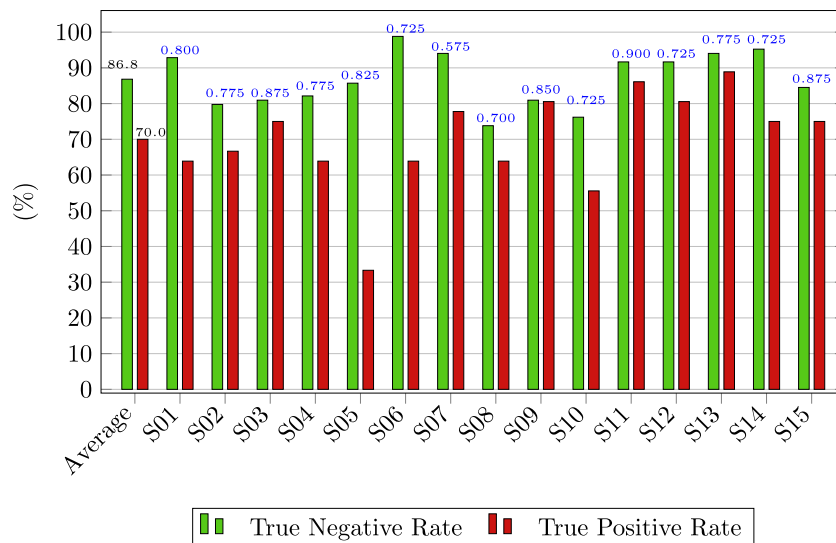
Online asynchronous ErrP detection. In the online part of the experiment, we used for the asynchronous ErrP detection, a subject specific-threshold τ, calculated as described in section *Threshold τ for the ErrP classifier*. The evaluation metric used to assess the results was described in section *Online ErrP detection*. Figure 6 depicts the TPR and TNR of the online asynchronous ErrP classification for every participant as well as the average results. We obtained an average TPR of 70.0% and average TNR of 86.8%. The blue numbers on top of the bars indicate the used threshold τ used for every participant.

Figure 7 depicts, for every participant, a violin plot of the time-points of all the ErrP detections in the error trials of the online part of the experiment, in relation to the average error onset ($t = 0$ s).

Discussion

In the described experiment, we asynchronously decoded ErrPs in an online scenario. Here, we showed the ErrPs' electrophysiology during the calibration and the online parts of the experiment. In both conditions, ErrP displayed similar shapes but the grand average ErrP in the online condition exhibited stronger peak amplitudes.

We chose to display, in both conditions, the ErrPs' electrophysiology using EEG signals filtered with a causal filter in order to match the ErrPs' appearance in the online scenario. The displayed results differ from standard

Figure 6. Online asynchronous ErrP detection. The green bars represent the TNR of every participant and their average. The red bars represent the TPR of every participant and their average. The average TPR was 70.0% and the average TNR was 86.8%. The blue numbers indicate the threshold τ, used for each participant.
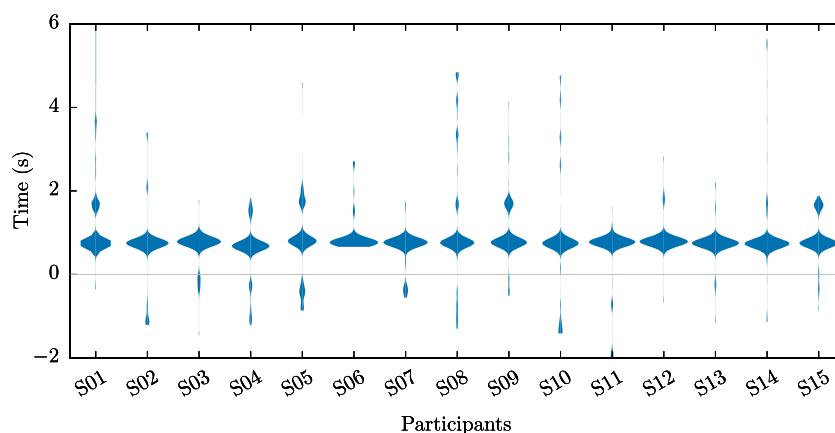
Figure 7. Time-points of all ErrP detections in the online scenario. Violin plots, for every participant, of the time-points of all ErrPs detections in the error trials of the online part of the experiment, in relation to the average error onset ($t = 0$ s).

state-of-the-art literature, in which it is commonly used a zero-phase filter. In our situation, the typical N200 component of ErrPs is shifted to after the ErrPs' P300 component. The difference is a direct consequence of using a causal filter and does not reflect any particularity of the neural activity.

We also showed results regarding the asynchronous ErrP detection in the calibration data using cross-validation, where different thresholds for the ErrP classifier, ranging from 0 to 1, could be tested.

Finally, we displayed the results of the asynchronous ErrP detection for the online part of the experiment, in which we obtained an average TNR of 86.8% of and an average TPR of 70%. In the online part of the experiment, all participants displayed a major cluster of ErrP detections within one second of the error onset, as shown in Fig. 7. Some participants displayed a secondary cluster of ErrP detections, which can possibly be associated with secondary ErrPs, as described by Salazar-Gomez and colleagues[28]. Alternatively, these later detections could also be possibly linked to an event-related potential associated with the robot resuming its movement (that the classifier erroneously classified as an ErrP).

We decided not to give participants feedback regarding false positive detections, neither in correct nor in error trials, to maintain the flow of the experiment and avoid interruptions. Still, from Figs. 6 and 7, we can infer that the majority of ErrP detections were not associated with false positive detections.

Literature supports that, in general, feedback improves BCIs performance and several feedback modalities have been tested[41–45]. But, to the best of our knowledge, the effect of ErrPs' feedback has not been studied yet. Nevertheless, we believe it can help participants to be more engaged and could possibly be associated with the increase in the peak amplitudes of the ErrP verified in the online scenario. Moreover, we believe that providing

feedback of the false positive detections could help participants to understand if they have any control over these detections and, if so, adapt their behaviour accordingly.

Therefore, we conclude that the asynchronous decoding of ErrPs in an online scenario is possible and reliable and we suggest that giving participants full feedback of the ErrP detections would not decrease and would possibly increase participants' performance.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Brunner, C. *et al.* BNCI Horizon 2020: towards a roadmap for the BCI community. *Brain-Computer Interfaces* **2**, 1–10, https://doi.org/10.1080/2326263X.2015.1008956 (2015).
2. Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G. & Vaughan, T. M. Brain–computer interfaces for communication and control. *Clinical Neurophysiology* **113**, 767–791, http://www.sciencedirect.com/science/article/pii/S1388245702000573 (2002).
3. Millán, J. D. R. *et al.* Combining brain–computer interfaces and assistive technologies: State-of-the-art and challenges. *Frontiers in Neuroscience* **4**, 161, https://doi.org/10.3389/fnins.2010.00161 (2010).
4. Müller-Putz, G. R. *et al.* Towards noninvasive hybrid brain-computer interfaces: Framework, practice, clinical application, and beyond. *Proceedings of the IEEE* **103**, 926–943 (2015).
5. Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E. & Donchin, E. A neural system for error detection and compensation. *Psychological Science* **4**, 385–390 (1993).
6. Falkenstein, M., Hohnsbein, J., Hoormann, J. & Blanke, L. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. *Psychophysiological Brain Research* 192–195 (1990).
7. Falkenstein, M., Hohnsbein, J., Hoormann, J. & Blanke, L. Effects of crossmodal divided attention on late ERP components. 2. error processing in choice reaction tasks. *Electroencephalography and Clinical Neuro-physiology* **78**, 447–455 (1991).
8. Mathalon, D. H., Whitfield, S. L. & Ford, J. M. Anatomy of an error: ERP and fMRI. *Biological Psychology* **64**, 119–141, http://www.sciencedirect.com/science/article/pii/S0301051103001054, Information Processing and Error Analysis: Retrospectives on a Career (2003).
9. Godlove, D. C. *et al.* Event-related potentials elicited by errors during the stop-signal task. i. macaque monkeys. *Journal of Neuroscience* **31**, 15640–15649, http://www.jneurosci.org/content/31/44/15640 (2011).
10. Milekovic, T., Ball, T., Schulze-Bonhage, A., Aertsen, A. & Mehring, C. Error-related electrocorticographic activity in humans during continuous movements. *Journal of Neural Engineering* **9**, 026007, http://stacks.iop.org/1741-2552/9/i=2/a=026007 (2012).
11. Milekovic, T., Ball, T., Schulze-Bonhage, A., Aertsen, A. & Mehring, C. Detection of error related neuronal responses recorded by electrocorticography in humans during continuous movements. *PLoS One* **8**, 1–20, https://doi.org/10.1371/journal.pone.0055235 (2013).
12. Völker, M. *et al.* Intracranial error detection via deep learning. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 568–575 (2018).
13. Keyl, P. *et al.* Differences in characteristics of error-related potentials between individuals with spinal cord injury and age- and sex-matched able-bodied controls. *Frontiers in Neurology* **9**, 1192, https://doi.org/10.3389/fneur.2018.01192 (2019).
14. Yeung, N., Botvinick, M. & D Cohen, J. The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological review* **111**, 931–59 (2004).
15. Schalk, G., Wolpaw, J. R., McFarland, D. J. & Pfurtscheller, G. EEG-based communication: presence of an error potential. *Clinical Neurophysiology* **111**, 2138–2144, http://www.sciencedirect.com/science/article/pii/S1388245700004570 (2000).
16. Ferrez, P. W. & Millán, J. d. R. You are wrong!: Automatic detection of interaction errors from brain waves. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, IJCAI'05, 1413–1418, http://dl.acm.org/citation.cfm?id=1642293.1642517 (2005).
17. van Schie, H. T., Mars, R. B., Coles, M. G. H. & Bekkering, H. Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience* **7**, 549–554 (2004).
18. Artusi, X., Niazi, I. K., Lucas, M.-F. & Farina, D. Accuracy of a BCI based on movement-related and error potentials. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3688–3691 (2011).
19. Llera, A., van Gerven, M. A., Gómez, V., Jensen, O. & Kappen, H. J. On the use of interaction error potentials for adaptive brain computer interfaces. *Neural Networks* **24**, 1120–1127 (2011).
20. Chavarriaga, R., Sobolewski, A. & Millán, J. D. R. Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in Neuroscience* **8**, 208, https://doi.org/10.3389/fnins.2014.00208 (2014).
21. Kim, S.-K. & Kirchner, E. A. Handling few training data: classifier transfer between different types of error-related potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **24**, 320–332 (2016).
22. Iturrate, I., Montesano, L. & Minguez, J. Single trial recognition of error-related potentials during observation of robot operation. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4181–4184 (2010).
23. Zhang, H. *et al.* EEG-based decoding of error-related brain activity in a real-world driving task. *Journal of Neural Engineering* **12**, 066028 (2015).
24. Ferrez, P. W. & Millán, É. D. R. Simultaneous real-time detection of motor imagery and error-related potentials for improved BCI accuracy. In *Proceedings of the 4th International Brain–Computer Interface Workshop and Training Course*, 197–202 (2008).
25. Chavarriaga, R. & Millán, Jd. R. Learning from EEG error-related potentials in noninvasive brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **18**, 381–388 (2010).
26. Kreilinger, A., Neuper, C. & Müller-Putz, G. R. Error potential detection during continuous movement of an artificial arm controlled by brain–computer interface. *Medical & Biological Engineering & Computing* **50**, 223–230 (2012).
27. Kreilinger, A., Hiebel, H. & Müller-Putz, G. R. Single versus multiple events error potential detection in a BCI-controlled car game with continuous and discrete feedback. *IEEE Transactions on Biomedical Engineering* **63**, 519–529 (2016).
28. Salazar-Gomez, A. F., DelPreto, J., Gil, S., Guenther, F. H. & Rus, D. Correcting robot mistakes in real time using EEG signals. *2017 IEEE International Conference on Robotics and Automation (ICRA)* 6570–6577 (2017).
29. Kim, S.-K., Kirchner, E. A., Stefes, A. & Kirchner, F. Intrinsic interactive reinforcement learning - using error-related potentials for real world human-robot interaction. *Scientific Reports (Sci Rep)* **7**, 17562 (2017).
30. Ehrlich, S. K. & Cheng, G. Human-agent co-adaptation using error-related potentials. *Journal of Neural Engineering* **15**, 066014, 10.1088%2F1741-2552%2Faae069 (2018).
31. Ehrlich, S. K. & Cheng, G. A feasibility study for validating robot actions using eeg-based error-related potentials. *International Journal of Social Robotics*, https://doi.org/10.1007/s12369-018-0501-8 (2018).

32. Millán, Jd. R. & Mouriño, J. Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **11**, 159–161 (2003).
33. Townsend, G., Graimann, B. & Pfurtscheller, G. Continuous EEG classification during motor imagery-simulation of an asynchronous BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **12**, 258–265 (2004).
34. Nooh, A. A., Yunus, J. & Daud, S. M. A review of asynchronous electroencephalogram-based brain computer interface systems. In *International Conference on Biomedical Engineering and Technology* (2011).
35. Spüler, M. & Niethammer, C. Error-related potentials during continuous feedback: using EEG to detect errors of different type and severity. *Frontiers in human neuroscience* **9** (2015).
36. Omedes, J., Iturrate, I., Minguez, J. & Montesano, L. Analysis and asynchronous detection of gradually unfolding errors during monitoring tasks. *Journal of Neural Engineering* **12**, 056001 (2015).
37. Omedes, J., Iturrate, I., Chavarriaga, R. & Montesano, L. Asynchronous decoding of error potentials during the monitoring of a reaching task. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 3116–3121 (2015).
38. Lopes-Dias, C., Sburlea, A. I. & Müller-Putz, G. R. Error-related potentials with masked and unmasked onset during continuous control and feedback. In *7th Graz Brain-Computer Interface Conference 2017*, 320–332 (2017).
39. Lopes-Dias, C., Sburlea, A. I. & Müller-Putz, G. R. Masked and unmasked error-related potentials during continuous control and feedback. *Journal of Neural Engineering* **15**, 036031, 10.1088%2F1741-2552%2Faab806 (2018).
40. Kobler, R. J., Sburlea, A. I. & Müller-Putz, G. R. A comparison of ocular artifact removal methods for block design based electroencephalography experiments. In *Proceedings of the 7th Graz Brain-Computer Interface Conference 2017* (2017).
41. Darvishi, S., Ridding, M. C., Abbott, D. & Baumert, M. Does feedback modality affect performance of brain computer interfaces? In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, 232–235 (2015).
42. Kauhanen, L. *et al.* Haptic feedback compared with visual feedback for BCI. In *Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course* (2006).
43. Ramos, A., Halder, S. & Birbaumer, N. Proprioceptive feedback in BCI. In *2009 4th International IEEE/EMBS Conference on Neural Engineering*, 279–282 (2009).
44. Alimardani, M., Nishio, S. & Ishiguro, H. Effect of biased feedback on motor imagery learning in BCI-teleoperation system. *Frontiers in Systems Neuroscience* **8**, 52, https://doi.org/10.3389/fnsys.2014.00052 (2014).
45. Barbero, Á. & Grosse-Wentrup, M. Biased feedback in brain-computer interfaces. *Journal of NeuroEngineering and Rehabilitation* **7**, 34, https://doi.org/10.1186/1743-0003-7-34 (2010).

Acknowledgements

Author contributions

C.L.D., A.I.S and G.R.M.P. conceived the study. C.L.D. implemented the paradigm and performed the acquisition. C.L.D. conducted the analysis. C.L.D., A.I.S and G.R.M.P. interpreted the data. C.L.D. created the figures and the video. C.L.D. wrote the draft of the manuscript. C.L.D., A.I.S and G.R.M.P. edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-54109-x.

Correspondence and requests for materials should be addressed to G.R.M.-P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.