DSM060 Coursework 1: Literature Review

Integrating Error-Related Potentials in Reinforcement Learning for BCI Systems

Submitted for partial fulfilment for the Data Science Research Topics course.

By

Hendrik Matthys van Rooyen

University of London

November 2023

CONTENTS

CHAPTER 1 ABSTRACT 2

CHAPTER 2 KEYWORDS 3

CHAPTER 3 BACKGROUND 4

3.1 PURPOSE AND SCOPE 4
3.2 RATIONALE 4
3.3 CONTEXT 4
3.4 METHODOLOGY 4
3.4.1 SEARCH STRATEGY 4
3.4.2 SOURCE EVALUATION 4
3.5 BASIC CONCEPTS 4
3.5.1 REINFORCEMENT LEARNING 4
3.5.2 ERROR-RELATED POTENTIALS 4

CHAPTER 4 REVIEW OF THE LITERATURE 5

4.1 INTEGRATION OF ERRPS IN REINFORCEMENT LEARNING: 5
4.1.1 MECHANISM OF ERRPS AS REWARD SIGNALS 5
4.1.2 CHALLENGES AND LIMITATIONS 6
4.2 COMPARATIVE ANALYSIS - ERRPS VS. TRADITIONAL REWARD SIGNALS 7

CHAPTER 5 CONCLUSION AND DISCUSSION 8

5.1 IDENTIFICATION OF GAPS 8
5.2 RECOMMENDATIONS FOR FUTURE STUDIES 8
5.3 KEY FINDINGS 8
5.4 CLOSING REMARKS 8

CHAPTER 6 REFERENCES 9

CHAPTER 1 ABSTRACT

This literature review explores the use of error-related potentials (ErrPs) as a reward mechanism in reinforcement learning (RL) for brain-computer interface (BCI) technology. It primarily focuses on how ErrPs, as brain-derived signals, can enhance the learning and adaptability of RL algorithms in BCI systems. The review analyses current literature, including academic papers and comprehensive studies, with a special emphasis on the integration of non-invasive EEG techniques, and their implementation. It highlights the significant advantages of utilizing ErrPs in RL, such as providing a more intuitive and direct feedback loop that aligns closely with the user's cognitive states and intentions. The review also discusses the challenges and limitations inherent in this approach. It concludes by identifying potential areas for future research, particularly in refining the detection and use of ErrPs in diverse BCI applications. This review underscores the importance of ErrPs in enhancing the effectiveness and user responsiveness of BCI systems, presenting a promising direction for future advancements in the field.

CHAPTER 2 KEYWORDS

Term	Definition
Error-Related Potentials (ErrPs)	Brain signals that occur when an individual (human) monitors the performance of an action, their own or external, particularly noticeable during errors. They are studied by comparing brain activity during correct and incorrect actions, focusing on specific signals related to errors.
Reinforcement Learning (RL)	A type of machine learning where a model learns to make decisions by performing actions and receiving feedback in the form of rewards or penalties. The model is not explicitly told which actions to take but must discover which actions yield the most reward through trial and error.
Brain-Computer Interface (BCI)	A technology that enables direct communication between the brain and external devices. BCIs interpret brain signals, allowing individuals to control external devices or computers with their thoughts, often used for assisting individuals with disabilities.
Electroencephalogram (EEG)	A method for recording electrical activity of the brain using electrodes placed along the scalp. EEG is widely used in BCI technology due to its non-invasive nature and is particularly useful for capturing ErrPs.

CHAPTER 3 BACKGROUND

3.1 PURPOSE AND SCOPE

This review explores the application of reinforcement learning in classifying EEG signals for use as computer inputs, a promising area in brain-computer interface (BCI) technology.

3.2 RATIONALE

Given the complexities of EEG signals and the potential of reinforcement learning to adapt and optimize performance, this study could offer significant advancements in BCI.

3.3 CONTEXT

BCIs facilitate communication between the brain and external devices, which may be especially useful for individuals with certain disabilities. EEG (opposed to iEEG and MRI), a non-invasive method, is widely used in BCI due to its portability and cost-effectiveness.

3.4 METHODOLOGY

3.4.1 SEARCH STRATEGY

Journals and other online resources (such as data sets) were gathered using key terms to search, as specified in Chapter 2. From resulting journals some select references were then considered as well.

3.4.2 SOURCE EVALUATION

Priority focus was given to recent, peer-reviewed papers and systematic reviews. As well as papers and documents referenced within the previously selected papers.

3.5 BASIC CONCEPTS

3.5.1 REINFORCEMENT LEARNING

According to Barto and Sutton (2018), reinforcement learning is the process through which the learning agent is not told which actions to take but is instead encouraged to explore available options to identify those most rewarding.

Per Liu *et al.* (2023), Reinforcement learning does have successful applications in the medical field and describes the RL agent as following an optimization strategy to identify the appropriate action for a given task.

3.5.2 ERROR-RELATED POTENTIALS

According to Xavier Fidêncio *et al.* (2022) ErrPs, short for error-related potentials, are brain signals that occurs when a human monitors performance of an action. To study these, scientists look at the difference in brain activity when people make mistakes compared to when they do things correctly. They average these differences over all subjects and trials. This method helps to focus on the specific brain signals related to errors. The characteristics of ErrPs are identified by looking at these differences in brain activity.

CHAPTER 4 REVIEW OF THE LITERATURE

4.1 INTEGRATION OF ERRPS IN REINFORCEMENT LEARNING:

4.1.1 MECHANISM OF ERRPS AS REWARD SIGNALS

The identification of ErrPs, as with many other non-invasive BCI inputs makes use of an electroencephalogram (EEG) machine Millan (2005) describes the process they followed to identify ErrPs evoked from erroneous BCI responses. They describe the process as follows:

1. EEG signals are acquired non-invasively from the scalp using electrodes. These signals are then spatially filtered and bandpass filtered (1-10 Hz) to focus on the ErrP, which are known to be slow cortical potentials.
2. The EEG signals are subsampled, and the relevant window for classification is identified. This window starts 150 ms after the feedback and ends 650 ms after the feedback, focusing on channels Cz and Fz due to the fronto-central distribution of ErrP.
3. A Gaussian classifier is used to classify single trials as "correct" or "error." The classifier is trained using temporal features extracted from EEG signals.
4. The Gaussian classifier consists of several prototypes representing different classes. The classifier is trained using clustering algorithms and stochastic gradient descent to minimize error.

Ferrez and Millan (2008) used sLORETA to estimate intracranial activity and a Gaussian classifier for single-trial classification of ErrPs. Recognition rates for erroneous and correct responses were analyzed. They confirmed that

In a later study by Xu *et al.* (2021), the researchers tried a different approach where all waveforms were classified as either ErrP or non-ErrP. In this study the Riemannian Geometry framework was used for classification.

According to Xavier Fidêncio *et al.* (2022) evoked ErrPs classifications is found in literature:

- **Response Errors**: Occur when a subject responds quickly to a stimulus.
- **Feedback Errors**: Arise when feedback informs the subject about the outcome of their choice.
- **Target Errors**: Triggered by unexpected changes in a task, like during a reaching/aiming task.
- **Interaction Errors**: Happen when there is a mismatch between a subject's command and the system's execution.
- **Execution Errors**: Occur when the system executes a different action than the command given by the subject. Xavier Fidêncio *et al.* does note that this may be the same as Interaction errors.
- **Outcome Errors**: Present when the desired goal of a movement is not fulfilled.
- **Observation Errors**: Emerge when a subject observes an error made by an external system.

Xavier Fidêncio *et al.* (2022) also states that ErrPs are shown to be generated not only from self-made errors but also during interactions with or observations of external systems. They state that several experiments have explored ErrPs in tasks requiring high concentration, such as motor imagery or cognitive tasks. In addition to detection, they discuss how some research has explored the practical application of ErrPs in improving BMI performance. This includes using online ErrPs detection to enhance BMIs, correcting human response errors, stopping cursor movements, and integrating error-correction mechanisms in BCI spellers.

Xu *et al.* (2021) explored two approaches to integrating the ErrP feedback into a Reinforcement Learning model. Firstly the “full access method”, in which every state-action pair is evaluated against the evoked or non-evoked ErrP signal, rewarding the function if no ErrP signal is present, thus optimizing the RL toward not committing actions that would evoke ErrPs.
The second, the researchers call this the “robust reward shaping framework”, is aimed at reducing the total number of ErrP queries and the load on the human in the loop. This is achieved by recording ErrP signals beforehand against a set dataset, which is then used to learn an auxiliary reward function that can compensate for the sparser environmental reward during RL training. The method focuses on learning the optimal policy in the human mind, improving the robustness to mistakes in ErrP decoding. Unlike the full access method, this approach only requires human feedback on initially generated trajectories.

Kim *et al.* (2017) approached controlling robot movement in simulation and reality by means of hand gesture through training a contextual bandit approach Reinforcement Learning model, while using the LinUCB algorithm. They chose this approach specifically for the requirement of mapping gesture to movement. Here they used the lack of ErrPs (NoErrPs) as focus in training the RL model, only updating on positive feedback. According the researchers, this differentiation between ErrPs and NoErrPs allows the RL system to adjust its policy based on the accuracy of its actions relative to the user's intentions.

4.1.2 CHALLENGES AND LIMITATIONS

Xavier Fidêncio *et al.* (2022) discusses some of the challenges in detecting and interpreting ErrPs. That ErrPs can be generated in the brain not only due to self-made errors but also during interaction or observation of external system operations. This variability is discussed above in the ErrPs classifications. Furthermore, the researchers found that the standard of ErrP measurement differs from study to study, making it difficult to establish a baseline.

Sentiment regarding existing uncertainty introduced by the ErrP classification exists in multiple studies (Liu *et al.*, 2023) (Xu *et al.*, 2021) (Blankertz *et al.*, 2003) although this is a recurrent theme in papers studying other brain signal classification as well. (Abenna *et al.*, 2022) (Wan *et al.*, 2023)

Xavier Fidêncio *et al.* (2022) also finds that ErrPs have been applied in various contexts such as BCI spellers, robotic arm, and software control, each presenting unique challenges in error signal interpretation.

Xavier Fidêncio *et al.* (2022) poses that, while there are cases in which ErrPs has been integrated with some benefit, it may not always be the case. They refer to Iturrate *et al.* (2015b) as showing that the overall online decoding accuracy was not significantly different from standard training approaches. However, it is noteworthy that in a case such as Xu *et al.*, (2021), the goal isn’t to achieve a higher classification accuracy through the use or ErrPs, but to alleviate strain on the human participant, therefor making the training process easier.

4.2 COMPARATIVE ANALYSIS - ERRPS VS. TRADITIONAL REWARD SIGNALS

Xu *et al.*, (2021) demonstrates that ErrPs provide a more efficient and less burdensome method of integrating human feedback into RL compared to traditional explicit feedback methods. This is achieved through innovative approaches like zero-shot learning, reward shaping frameworks, and effective combination of different types of rewards.

Xavier Fidêncio *et al.* (2022) found that ErrPs show considerable promise in enhancing the learning efficiency of RL models, particularly in brain-machine interfaces. They provide a direct, brain-derived feedback mechanism that can dynamically adjust learning strategies based on the user's neural responses to errors, offering a personalized and potentially more effective approach compared to traditional reward signals in RL.

Kim *et al.*, (2017) showed the use of ErrPs as intrinsic feedback in RL demonstrates promising results, especially in terms of learning efficiency and adaptability in real-world scenarios. The study highlights the potential of integrating human cognitive states, as reflected in EEG signals, into robotic control systems for enhanced human-robot interaction.

CHAPTER 5 CONCLUSION AND DISCUSSION

5.1 IDENTIFICATION OF GAPS

The review has identified certain gaps in the current state of research. While ErrPs have been successfully integrated into RL, there remains a lack of standardization in their measurement and interpretation across different studies. This variability hinders the establishment of a consistent baseline for comparing results and developing universal applications.

The review also failed to identify a multiple/complex input study beyond that of the BCI typing that BCI spellers (Margaux et al., 2012; Bevilacqua et al., 2020), from the Xavier Fidêncio *et al.* (2022) review which attempted to use ErrPs in the reward function of the RL model.

5.2 RECOMMENDATIONS FOR FUTURE STUDIES

Future research should focus on standardizing ErrP measurements and developing more robust algorithms capable of handling the variability and noise inherent in EEG signals. Additionally, exploring the integration of ErrPs in more diverse BCI applications, such as those involving complex motor tasks or cognitive functions, could yield valuable insights.

5.3 KEY FINDINGS

The key findings from the literature suggest that ErrPs offer a promising avenue for enhancing the efficiency of RL in BCIs. Studies have shown that ErrPs can serve as a direct, brain-derived reward mechanism, potentially offering a more personalized and effective approach compared to traditional reward signals in RL.

5.4 CLOSING REMARKS

In conclusion, the integration of ErrPs in RL presents a significant opportunity to advance BCI technology. While challenges persist, the potential benefits of a more focused and direct brain-machine interface are considerable. By addressing the identified gaps and building upon the current knowledge base, future research can pave the way for more sophisticated and user centred BCI systems.

Wordcount: 2068

CHAPTER 6 REFERENCES

Abenna, S. *et al.* (2022) 'EEG-based BCI: A novel improvement for EEG signals classification based on real-time preprocessing', *Computers in Biology and Medicine*, 148, p. 105931. Available at: https://doi.org/10.1016/j.compbiomed.2022.105931.

Al-Saegh, A., Dawwd, S.A. and Abdul-Jabbar, J.M. (2021) 'Deep learning for motor imagery EEG-based classification: A review', *Biomedical Signal Processing and Control*, 63, p. 102172. Available at: https://doi.org/10.1016/j.bspc.2020.102172.

Angelakis, D., Ventouras, E. and Asvestas, P. (2023) 'Brain Signals to Actions Using Machine Learning', *Engineering Proceedings*, 50(1), p. 7. Available at: https://doi.org/10.3390/engproc2023050007.

Blankertz, B. *et al.* (2003) 'Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), pp. 127–131. Available at: https://doi.org/10.1109/TNSRE.2003.814456.

Delisle-Rodriguez, D. *et al.* (2023) 'Multi-channel EEG-based BCI using regression and classification methods for attention training by serious game', *Biomedical Signal Processing and Control*, 85, p. 104937. Available at: https://doi.org/10.1016/j.bspc.2023.104937.

Ferrez, P.W. and del R. Millan, J. (2008) 'Error-Related EEG Potentials Generated During Simulated Brain–Computer Interaction', *IEEE Transactions on Biomedical Engineering*, 55(3), pp. 923–929. Available at: https://doi.org/10.1109/TBME.2007.908083.

Gil Ávila, C. *et al.* (2023) 'DISCOVER-EEG: an open, fully automated EEG pipeline for biomarker discovery in clinical neuroscience', *Scientific Data*, 10(1), p. 613. Available at: https://doi.org/10.1038/s41597-023-02525-0.

Huang, W. *et al.* (2023) 'EEG-based classification combining Bayesian convolutional neural networks with recurrence plot for motor movement/imagery', *Pattern Recognition*, 144, p. 109838. Available at: https://doi.org/10.1016/j.patcog.2023.109838.

info@sagebase.org, S.B. (no date) *Synapse | Sage Bionetworks*. Available at: https://www.synapse.org (Accessed: 11 November 2023).

Kim, S.-K. *et al.* (2017) 'Intrinsic interactive reinforcement learning – Using error-related potentials for real world human-robot interaction', *Scientific Reports*, 7. Available at: https://doi.org/10.1038/s41598-017-17682-7.

Liu, X. *et al.* (2023) 'Automatic focal EEG identification based on deep reinforcement learning', *Biomedical Signal Processing and Control*, 83, p. 104693. Available at: https://doi.org/10.1016/j.bspc.2023.104693.

Max Planck Institut Leipzig Mind-Brain-Body Dataset - LEMON (no date). Available at: http://fcon_1000.projects.nitrc.org/indi/retro/MPI_LEMON.html (Accessed: 11 November 2023).

Millan, J. del R. (2005) *You Are Wrong!---Automatic Detection of Interaction Errors from Brain Waves, Proceedings of the 19th International Joint Conference on Artificial Intelligence*, p. 1418.

Parra, L.C. *et al.* (2003) 'Response error correction-a demonstration of improved human-machine performance using real-time EEG monitoring', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), pp. 173–177. Available at: https://doi.org/10.1109/TNSRE.2003.814446.

Publicly Available EEG Datasets (2020) *OpenBCI Community*. Available at: https://openbci.com/community/publicly-available-eeg-datasets/ (Accessed: 11 November 2023).

Schalk, G. *et al.* (2000) 'EEG-based communication: presence of an error potential', *Clinical Neurophysiology*, 111(12), pp. 2138–2144. Available at: https://doi.org/10.1016/S1388-2457(00)00457-0.

Schalk, G. *et al.* (2009) 'EEG Motor Movement/Imagery Dataset'. physionet.org. Available at: https://doi.org/10.13026/C28G6P.

Spüler, M. and Niethammer, C. (2015) 'Error-related potentials during continuous feedback: using EEG to detect errors of different type and severity', *Frontiers in Human Neuroscience*, 9, p. 155. Available at: https://doi.org/10.3389/fnhum.2015.00155.

Sun, C. and Mou, C. (2023) 'Survey on the research direction of EEG-based signal processing', *Frontiers in Neuroscience*, 17. Available at: https://www.frontiersin.org/articles/10.3389/fnins.2023.1203059 (Accessed: 11 November 2023).

Sutton & Barto Book: Reinforcement Learning: An Introduction (no date). Available at: http://incompleteideas.net/book/the-book-2nd.html (Accessed: 19 December 2023).

Wan, Z. *et al.* (2023) 'EEGformer: A transformer–based brain activity classification method using EEG signal', *Frontiers in Neuroscience*, 17. Available at: https://www.frontiersin.org/articles/10.3389/fnins.2023.1148855 (Accessed: 11 November 2023).

Wang, X. *et al.* (2023) 'Deep learning-based EEG emotion recognition: Current trends and future perspectives', *Frontiers in Psychology*, 14. Available at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1126994 (Accessed: 11 November 2023).

Xavier Fidêncio, A., Klaes, C. and Iossifidis, I. (2022) 'Error-Related Potentials in Reinforcement Learning-Based Brain-Machine Interfaces', *Frontiers in Human Neuroscience*, 16. Available at: https://www.frontiersin.org/articles/10.3389/fnhum.2022.806517 (Accessed: 16 December 2023).

Xu, D. *et al.* (2021) 'Accelerating Reinforcement Learning using EEG-based implicit human feedback', *Neurocomputing*, 460, pp. 139–153. Available at: https://doi.org/10.1016/j.neucom.2021.06.064.