IDIAP RESEARCH REPORT

To Err Is Human: Learning from Error Potentials in Brain-Computer Interfaces

Ricardo Chavarriaga [a] Pierre W. Ferrez [a b]
José del R. Millán [a b]

IDIAP–RR 07-37

August 2007

first revision : June 07
second revision : August 07

to appear in
Proc of. 1st International Conference on Cognitive Neurodynamics - 2007 (ICCN'07)). Shanghai, China. November 2007

[a] IDIAP Research Institute, Martigny, Switzerland
[b] Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

IDIAP Research Report 07-37

To Err Is Human: Learning from Error Potentials in Brain-Computer Interfaces

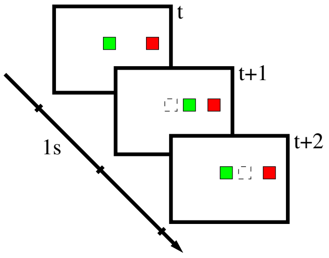
Ricardo Chavarriaga Pierre W. Ferrez José del R. Millán

August 2007

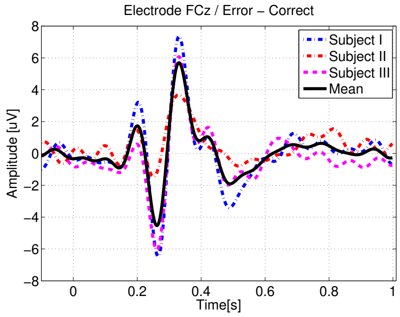
first revision : June 07
second revision : August 07

to appear in
Proc of. 1st International Conference on Cognitive Neurodynamics - 2007 (ICCN'07)). Shanghai, China. November 2007

Abstract. Several studies describe evoked EEG potentials elicited when a subject is aware of an erroneous decision either taken by him or by an external interface. This paper studies *Error-related potentials* (ErrP) elicited when a human user monitors an external system upon which he has no control whatsoever. In addition, the possibility of using the ErrPs as a learning signals to infer the user's intended strategy is also addressed. Experimental results show that single-trial recognition of correct and error trials can be achieved, allowing the fast learning of the user's strategy. These results may constitute the basis of a new kind of human-computer interaction where the former provides monitoring signals that can be used to modify the performance of the latter.This work has been supported by the Swiss National Science Foundation NCCR-IM2 and by the EC-contract number BACS FP6-IST-027140. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.



(a)



(b)

Figure 1: (a) Experimental protocol. *Green square*, moving cursor. *Red square*, target location. *Dotted square*, cursor location at the previous time step. Correct and erroneous movements are shown at times *t+1* and *t+2* respectively. (b) Grand average error related potential—Error minus Correct condition—on the FCz electrode (feedback is given at time t=0).

1 Introduction

The error monitoring process is crucial to improve performance for both humans and artificial cognitive systems. Upon identification of erroneous decisions, the likelihood of repeating such actions in the same context should be decreased in order to improve the performance. This mechanism is the base of the reinforcement learning theory [1]. Several studies on human EEG have identified event-related potentials elicited by erroneous decisions or error feedback; i.e. *error-related negativity* (ERN) and *feedback-related negativity* (FRN), respectively [2]. These potentials are characterized by a midline frontal negative deflection of the EEG around 80 ms and 250 ms, respectively. Moreover, evoked potentials have also been described when errors are generated by a human-computer interface and not by the user himself —e.g. keyboard or brain-computer interfaces (BCI) [3].

Although these signals convey valuable information about the user's evaluation of performance, they have seldom been used in the field of non-invasive BCIs. Parra and colleagues have proposed the use of ERNs to correct user's erroneous decisions on speed response tasks [4], while Ferrez et al. use *error-related potentials* (ErrP) to improve the information transfer rate of a BCI system [3]. However, these studies use ErrPs only to correct the response that generated the error, but they do not include a learning mechanism to prevent that error to be repeated in the future.

This paper addresses this issue by testing the use of error-related potentials as a teaching signal to learn the user's strategy when solving a simple task. In particular, *(i)* we test whether these potentials can be detected on single trial when a human subject monitors the performance of an external system upon which he has no control whatsoever; and *(ii)* we assess the feasibility of using such signals to learn the user's intention, i.e. the strategy the user expects the system to perform.

2 Experimental Protocol

Subjects seat in front of a computer screen where a moving cursor (a green square) is displayed. A red square at either the left or right of the cursor signals the target location, as shown in Figure 1a. At each time step the cursor moves horizontally depending on the location of the target. The user has no control over the cursor's movement and is asked only to monitor the performance of the system, knowing that the goal is to reach the red target. Cursor movement law is defined in a suboptimal way in order to study signals elicited by the system's errors. Specifically, at each time step there is a probability of 80% for the cursor to move towards the target. One experimental session consists of 10 blocks of 3 min each (approx. 75 trials per block).

Three healthy male subjects performed two sessions of the experiment. Data from the first session was used to characterize the evoked potentials for both error and correct trials and to train the classifier parameters (see

Table 1: Single trial recognition rates (%) for the three subjects on the test set (i.e. session 2).

	Subject I		Subject II		Subject III	
	Correct	Error	Correct	Error	Correct	Error
Correct	92.01	7.99	83.82	16.18	86.86	13.14
Error	26.50	73.50	41.08	58.91	33.71	66.29

section 3 below). The second session, recorded seven weeks after the first one, was used to test ErrP single-trial classification. EEG was recorded for all subjects with a sampling rate of 512 Hz using 64 electrodes according to the standard 10/20 International system. Data was re-referenced to the common average reference (CAR) and a 1-10Hz band-pass filter was applied. The average feedback-locked potential at channel FCz—difference between error and correct grand averages—for the three subjects is shown in Figure 1b.

To emulate experimental conditions of a BCI application, no artifact rejection was applied and all trials were used in the analysis. Moreover classification accuracy was assessed on a single-trial basis, as opposite to batch performance evaluation techniques such as cross-validation.

3 Single-trial Classification

Following previous studies, we classify the signals using a Gaussian classifier [3]. The activity on the FCz and Cz electrodes on the [200 ms, 450 ms] time window after the feedback presentation was used as input to the classifier. The statistical classifier estimates the posterior probability of a single trial corresponding to each of the two classes *error*, and *correct*. Classifier parameters were tuned using a stochastic gradient descent on the mean square error.

Table 1 shows the recognition rates on single-trial for the three subjects. The same classifier parameters were used in all cases. Learning rates were 10^{-2} and 10^{-4} for the center and covariances of the Gaussian prototypes. A total of 6 prototypes were used for each class. These results show classification above chance for both correct and error trials, with higher recognition rates for the correct trials.

4 Learning from Error-related Potentials

We test then whether the detection of ErrPs can be used as to infer what the optimal behavior should be. The rationale of this approach is that, given the system decisions and the user's evaluation of such decisions—indicated by the presence or not of ErrPs—, it is possible to infer what strategy is considered as correct by the human user.

Considering the experiment described above, the optimal strategy is to move the cursor towards the target. Let define $T_t, A_t \in [L, R]$ the target location and the cursor's direction of movement at time t, where $[L, R]$, stand for left and right respectively. Let $P_{A,T}$ be the probability of taking action A given the target location T, and a strategy Π, i.e. $P_{A_t,T_t} = P(Action = A_t|Target = T_t, \Pi)$. The optimal strategy Π^* can be expressed in terms of probabilities $P^*_{L,L} = P^*_{R,R} = 1$; respectively, $P^*_{L,R} = P^*_{R,L} = 0$.

Alike to reward signals in reinforcement learning, ErrP detection can be used to decrease the likelihood of performing actions considered as erroneous and, in the opposite case, to encourage correct actions. Let define $\Pi^t = \{P^t_{L,L}, P^t_{L,R}, P^t_{R,L}, P^t_{R,R}\}$ the strategy at time t, if an ErrP is detected, the probability of repeating the action A_t given the target location T_t must be decreased, i.e. $P^{t+1}_{A_t,T_t} = P^t_{A_t,T_t} - \Delta P^t_{A_t,T_t}$. Conversely, if the trial is considered as correct, $P^t_{A_t,T_t}$ is increased for the next time step. The probabilities of other actions given T_t are updated so that $\sum_i P_{A_i,T_t} = 1$. Note that we keep separate models for each possible target location.

We choose a variable step size $\Delta P^t_{A_t,T_t}$ such that probabilities close to chance level are penalized (i.e. $P^t_{A_t,T_t} = 0.5$, for two-action problems). In the current implementation $\Delta P^t_{A_t,T_t} = \eta H(P^t_{A_t,T_t})$, where $\eta = 0.3$ is a constant scaling factor, and $H(x) = -\sum_i^n P(x_i) log_2 P(x_i)$ is the binary entropy function.

Figure 2 shows how the optimal strategy is learned using error related potentials. For the three subjects, based on the recognition of error and correct trials the probability of performing the correct action increases for

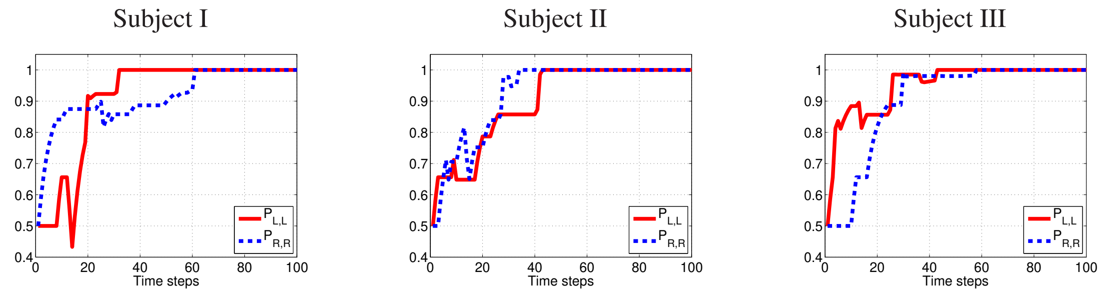


Figure 2: ErrP-based learning of the user's intended strategy for the three subjects. The X-axis represents the time steps and the Y-axis represents the probability of performing the correct action given the current strategy (i.e. $\Pi^t = \{P_{L,L}^t, P_{R,R}^t\}$).

both possible target locations. On average, the optimal strategy Π^* is acquired in about 40 trials.

5 Conclusions

This study shows a novel use of error-related potentials in the frame of brain-computer interfaces. In the proposed approach, the human user acts as a critic who observes the behavior of an acting agent and emits monitoring signals about its performance. Successful learning of the optimal strategy is achieved by *(i)* single trial recognition of ErrPs as monitoring signals (c.f. Table 1), as well as *(ii)* an efficient strategy update mechanism based on these signals (c.f. Figure 2).

Recent studies have linked ErrPs to the theory of reinforcement learning in humans [5]. Similarly, this study exploits ErrPs as a reinforcement signal for an external system to learn the user's intended strategy. Notice that we have reported results on the off-line learning of the optimal strategy. New experiments are currently undergoing to test this approach when the ErrP signals modify the performance of the actual system being monitored.

References

[1] Sutton, R., Barto, A.G.: Reinforcement Learning—An Introduction. MIT Press, Cambridge, Ma, USA (1998)

[2] Falkenstein, M., Hoormann, J., Christ, S., Hohnsbein, J.: ERP components on reaction errors and their functional significance: A tutorial. Biol Psychol **51**(2-3) (Jan 2000) 87–107

[3] 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, Edinburgh, UK (August 2005)

[4] Parra, L.C., Spence, C.D., Gerson, A.D., Sajda, P.: Response error correction—A demonstration of improved human-machine performance using real-time eeg monitoring. IEEE Trans Neural Syst Rehabil Eng **11**(2) (June 2003) 173–7

[5] Holroyd, C.B., Coles, M.G.H.: The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. Psychol Rev **109**(4) (Oct 2002) 679–709