International Journal of Psychophysiology 106 (2016) 97–105

Contents lists available at ScienceDirect

International Journal of Psychophysiology

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

ERP correlates of error processing during performance on the Halstead Category Test

I.M. Santos [a,*], A.R. Teixeira [b,c], A.M. Tomé [b,e], A.T. Pereira [a,d], P. Rodrigues [f,g], P. Vagos [h], J. Costa [f], M.L. Carrito [a,d], B. Oliveira [a,d], N.A. DeFilippis [i], C.F. Silva [a]

[a] Center for Health Technology and Services Research (CINTESIS), Department of Education and Psychology, University of Aveiro, Portugal
[b] IEETA, University of Aveiro, Portugal
[c] ESEC, Polytechnic Institute of Coimbra, Portugal
[d] IBILI, Faculty of Medicine, University of Coimbra, Portugal
[e] DETI, University of Aveiro, Portugal
[f] University of Beira Interior, Portugal
[g] CICS, Research Center in Health Sciences, Portugal
[h] CINEICC, Cognitive and Behavioural Center for Research and Intervention, Faculty of Psychology and Educational Sciences, University of Coimbra, Portugal
[i] Georgia School of Professional Psychology at Argosy University, Atlanta, USA

ARTICLE INFO

Article history:
Received 17 December 2015
Received in revised form 6 June 2016
Accepted 17 June 2016
Available online 19 June 2016

Keywords:
feedback-related negativity (FRN)
Halstead Category Test
Singular Spectrum Analysis (SSA)
event-related potentials (ERP)
feedback processing

ABSTRACT

The Halstead Category Test (HCT) is a neuropsychological test that measures a person's ability to formulate and apply abstract principles. Performance must be adjusted based on feedback after each trial and errors are common until the underlying rules are discovered. Event-related potential (ERP) studies associated with the HCT are lacking. This paper demonstrates the use of a methodology inspired on Singular Spectrum Analysis (SSA) applied to EEG signals, to remove high amplitude ocular and movement artifacts during performance on the test. This filtering technique introduces no phase or latency distortions, with minimum loss of relevant EEG information. Importantly, the test was applied in its original clinical format, without introducing adaptations to ERP recordings. After signal treatment, the feedback-related negativity (FRN) wave, which is related to error-processing, was identified. This component peaked around 250ms, after feedback, in fronto-central electrodes. As expected, errors elicited more negative amplitudes than correct responses. Results are discussed in terms of the increased clinical potential that coupling ERP information with behavioral performance data can bring to the specificity of the HCT in diagnosing different types of impairment in frontal brain function.

1. Introduction

The Halstead Category Test (HCT) (DeFilippis and McCampbell, 1997; DeFilippis, 2002), which is part of the Halstead-Reitan Neuropsychological Battery, is a neuropsychological test routinely used to assess abstract reasoning, concept formation and problem solving abilities in a variety of clinical contexts and populations (Allen et al., 2007, 1999; Choca et al., 1997). The test consists of 208 items, divided into 7 subtests, and the total number of errors is the score most commonly used to assess performance, which has proven to be highly sensitive for identifying brain dysfunction (Choca et al., 1997). However, as studies seem to demonstrate that the HCT is a multidimensional instrument which assesses diverse cognitive executive abilities (Allen et al., 1999), the total error score has also received some criticism, due to its lack of specificity in identifying which particular abilities are impaired (Allen et al., 2007). For this reason, a number of studies have attempted to develop more elaborate approaches to scoring the HCT and created scales based on individual subtests (e.g., McNally et al., 2015; Minassian et al., 2003). Results suggest that by looking at the scores on the various subscales, the HCT is able to provide information about a variety of cognitive functions that are usually assessed by different instruments, which is useful in informing clinicians about different domains where further evaluation should be focused in case of impairment. Nonetheless, studies demonstrating that the scales are able to discriminate neurologically impaired patients and healthy individuals are still lacking (McNally et al., 2015). The informative potential of this test regarding brain dysfunction could be considerably increased if brain activity was registered concomitantly to test performance. To the best of our knowledge, no studies so far have examined brain activity through event-related potentials (ERPs) during performance on the HCT. ERPs are a high temporal resolution electrophysiological technique, which is particularly suited for studying the time course of brain processes.

* Corresponding author at: University of Aveiro, Department of Education and Psychology, Campus Universitario de Santiago, 3810–193 Aveiro, Portugal.
E-mail address: isabel.m.b.santos@gmail.com (I.M. Santos).

http://dx.doi.org/10.1016/j.ijpsycho.2016.06.010

Therefore, ERP recordings during performance on the HCT could be helpful in increasing the diagnostic value of the test, providing a more direct link between test scores and underlying brain processes, and helping to disentangle the multitude of cognitive operations that are in play during HCT performance. This information could be used to better understand the brain processes and cognitive abilities that are impaired in association with a particular behavioral pattern, thus increasing the specificity of the test, an aspect of the HCT which has been pointed out as a limitation (Choca et al., 1997).

Although the electroencephalogram (EEG) can be easily recorded while participants perform on the HCT, the conditions are not ideal for the study of ERPs. Several characteristics of the test items differ from what is a classical ERP paradigm, namely the length of stimulus presentation, which in the HCT has no time limit and is typically long. This originates a high number of artifacts, both ocular and due to movement, since it is more difficult for the participant to exert self-control of these for the time of exposure. Another issue is the number of test items, which in total is 208, but which can be significantly reduced if specific analysis need to be carried out, namely analyzing only certain subtests, or separating correct from incorrect responses. Thus, it is necessary to employ a signal processing method that will allow for preservation of a high signal-to-noise ratio. The main aim of the present work was to demonstrate that it is possible to use the HCT in its current clinical format and simultaneously record an EEG for ERP analysis, in order to increase the informative value of the test in terms of identifying the underlying brain processes.

Another test commonly used in neuropsychological practice to assess executive function is the Wisconsin Card Sorting Test (WCST) (Heaton et al., 1993). Similarly to the HCT, the WCST is commonly regarded as a test of abstract reasoning and concept formation, and it is also generally adopted as an indicator of frontal lobe dysfunction. However, a lack of specificity to frontal damage has also been reported (e.g., Anderson et al., 1991), which has led to attempts to use ERPs as a brain activity measure to probe in detail the relation between brain dynamics and the cognitive processes underlying WCST performance (Barceló et al., 1997; Barceló, 2003). These studies have been successful in identifying frontal ERP components related to task performance, but also components that have a non-frontal origin (such as the P3b wave), which help clarify some reports of non-specificity of the WCST to prefrontal damage (Barceló and Rubia, 1998). However, these ERP studies have used a modified version of the WCST in order to adapt it to the recording of ERPs and to specifically allow the investigation of the electrophysiological dynamics related to attention set-shifting (Barceló, 2003). Our aim was to be able to explore the ERP correlates of performance on the HCT in its current clinical format, without introducing any adaptation to the task. In this way, we would be able to provide an additional tool that could be used in conjunction with behavioral performance. This would allow the use of standard norms and cut-off points for neuropsychological diagnosis, whilst simultaneously informing on the underlying brain dynamics, thus increasing the clinical potential of the test.

In the present work, we focused our analysis on the posterior P100 visual ERP component, which is elicited by visual stimuli independently of the task that the subject is doing (Luck, 2005) and is generally related to activity in extrastriate areas of the visual cortex (Pratt, 2011) and on the feedback-related negativity (FRN) wave, a frontocentral ERP component related to error processing (Miltner et al., 1997), which should be expected during performance on the HCT. This test consists of 208 items, divided into 7 subtests. The items consist of nonverbal stimuli representing geometric figures or designs, and the participant is asked to indicate the number between 1 and 4 that each stimulus suggests. After each response, visual feedback on whether the response was correct or incorrect is provided, which helps participants to adjust their strategy. The participant is informed that all items in a particular subtest have the same underlying abstract principle, and that this abstract principle may or may not change between subtests. Thus, an incorrect feedback after a response indicates the need to search for a different abstract principle. On the other hand, a correct feedback indicates that the same abstract principle should be maintained for the following items within that subtest. Every time a new subtest begins, the participant is informed that the underlying principle may be the same as in the last subtest or that it may change (McNally et al., 2015). Thus, this test measures concept learning, flexibility of thinking and ability to learn and apply new rules. Also, the test directly taps the ability to learn from experience, monitor the errors that are committed, and adjust one's response strategy as a function of feedback on the accuracy of a previous response, until a correct rule has been successfully identified which can be followed in the subsequent trials. Given the nature of the task, committing errors is common until the new rule is discovered.

An ERP component has been described, which follows the display of negative feedback, in tasks where errors are due to uncertainty regarding the correct response, and participants only become aware of the accuracy of their response after a feedback signal has been provided (Walsh and Anderson, 2012). This wave has been called feedback-related negativity (FRN). This component was first discovered in a time estimation task where participants had to push a button a second after a signal. A feedback stimulus told the participants whether the estimation was accurate or wrong, and a negative deflection appeared after negative feedback (Miltner et al., 1997). The FRN is measured maximally at midline fronto-central electrodes and is typically larger for erroneous responses than for correct responses, peaking between 200 and 250 ms after feedback. This component is believed to originate in a general purpose neural system for dealing with errors in different types of task, which contributes to the adjustment of ongoing behavior (Gehring and Willoughby, 2002; Miltner et al., 1997). Research has suggested that the FRN reflects the evaluation function of a neural system that determines whether an outcome was correct or incorrect relative to one's expectation (Gehring and Willoughby, 2002; Hajcak et al., 2006; Holroyd, 2004). Importantly, the FRN seems to reflect the processing of external cues about performance (Bernat et al., 2011). The most likely neural generator of the FRN has been localized in the dorsal area of the anterior circulate cortex (ACC), a brain region known to be involved in cognitive control and behavior regulation, and which is important for the ability to adapt behavior to different task demands and circumstances (Hauser et al., 2014; Walsh and Anderson, 2012).

As mentioned above, the Wisconsin Card Sorting Test (WCST) is another neuropsychological test commonly used to assess frontal lobe executive functions. Like the HCT, it also requires participants to infer sorting rules associated with simple geometrical stimuli (based on their color, shape, or number). Participants must adapt their responses based on the provided feedback regarding the accuracy of their responses (correct or incorrect feedback). A new discovered rule will be maintained for a number of trials, but after a while it will change again, requiring participants to discover the new sorting rule (Vilà-Balló et al., 2015). Studies using a modified version of the WCST, adapted to the recording of ERPs, have demonstrated the occurrence of the FRN component during performance on the test. In particular, (Kraus and Horowitz-Kraus, 2014) have shown that individuals with dyslexia exhibit decreased FRN amplitudes in the early phases of the task compared to normal readers, consistent with their difficulty in learning from previous mistakes. Another study, also using a modified ERP version of the WCST, has found larger FRN amplitude to positive feedback in a violent juvenile offender group compared to a control group, which suggested difficulty in using previous external feedback to accurately predict the negative outcomes of their behavior (Vilà-Balló et al., 2015). Thus, the FRN seems to be a reliable neurophysiological correlate of feedback-processing during performance on the WCST, a test that requires various executive abilities which are common to the HCT, as outlined above. Also consistent with these findings is evidence that performance on the WCST is positively correlated with metabolic activity in the cingulate region (Adams et al., 1995). So, since performance on the HCT also depends on the ability to monitor errors and adapt behavioral responses based on the success of previous performance, it is expected that the feedback stimuli in the HCT should elicit an FRN.

The main objective of this work was to demonstrate the possibility of meaningfully recording, analyzing and interpreting event-related potentials associated with performance on the HCT. To the best of our knowledge, this has never been done before. The main question addressed in this paper is whether it is possible to identify specific ERP waveforms (in this case, the FRN component after feedback, which should differentiate between right and wrong answers) without introducing changes in the structure of the test, and maintaining its clinical format and validity. As mentioned before, this is important because it allows the possibility of relating brain activity with clinical performance according to established norms in a clinical context. However, the identification of specific ERP components without modifying the structure of the HCT was a challenge, since the low number of trials and long stimulus duration are not typical from an ERP paradigm and originate a low signal-to-noise ratio and a large number of artifacts. The fact that participants had unlimited time to examine each stimulus, before providing a response, originated a large number of high amplitude artifacts, both due to blinks and other ocular movements, and also body movements. This compromises the analyzes of ERP components through more traditional methods, given that the elimination of artifact contaminated segments would lead to a very low signal-to-noise ratio, due to the relatively small number of trials available in this test, in its original format. Importantly, we demonstrate that the application of a specific filtering and signal processing technique successfully cleans the data, allowing the identification of the relevant ERP components.

This paper details the data processing methodology used to filter the signal, which guaranteed an artifact free signal with minimal loss of relevant information and no phase or latency distortions, which are common after the application of various types of filters (Rousselet, 2012). However it should be noticed that filtering manipulations follow the distributive property (Luck, 2005). Thus, averaging filtered signals is equivalent to averaging the raw signals and filtering the resulting average. Therefore, in this study, average ERP waves for each subject were computed first and then the filtering operation was applied. We demonstrate that, by extracting a high amplitude component of the average signal, it is possible to visualize characteristic transient events after the stimulus or the feedback. First, we show a clear P100 in response to the visual stimuli, which demonstrates the efficiency of the method (importantly showing that the filtering technique does not introduce distortions in the signal). Secondly, we also demonstrate that after extracting a high amplitude component of the signal, it is possible to identify the FRN component related to error processing, which should be elicited by feedback stimuli during the HCT. The extraction of the high-amplitude artifact component follows a methodology known as Singular Spectrum Analysis (SSA) that allows, with a filtering mechanism, the decomposition of a signal time-series into components that are in phase with the original signals (Tomé et al., 2010). Moreover, the SSA strategy was modified in order to compute the coefficients of the filter with a template of the signal to be extracted. This template is the average of the electro-oculogram signal. The advantage over the standard SSA methodology is that only one filter is computed and applied to all channels. After filtering the signal following this methodology, results show, as expected, larger negative amplitudes for negative feedback (incorrect responses) compared to positive feedback (correct responses) in the FRN component during performance on the HCT.

2. Methods

2.1. Participants

Fifty-eight graduate and undergraduate students were recruited from the University of Aveiro (39 females and 19 males). The participants mean age was 22.5 years (SD = 4.9; range: 19–48). Consent forms were signed prior to the experimental task and participants were paid for their participation. This study was conducted in accordance with the Declaration of Helsinki.

2.2. Materials and procedure

A computerized version of the Halstead Category Test (HCT) was used to assess cognitive executive frontal lobe function. This is a nonverbal test that measures a person's ability to formulate abstract principles. It consists of 208 items divided into 7 subtests: the first two are the training subtests (8 items in subtest I and 20 items in subtest II), the third and fourth measure spatial/positional reasoning, the fifth and sixth assess proportional reasoning (40 items in each of the subtests III to VI) and the last one is a memory subtest (20 items in subtest VII). Each test item shows one or more figures that, altogether, suggest a number ranging from one to four. In subtests I to VI, participants are instructed to determine or guess the correct number based on their conceptualization of the abstract principle represented by the stimulus, as shown in Fig. 1. Visual feedback is provided after each response, to indicate if the participant's response was right or wrong. Based on this feedback, the participant must maintain or change their response strategy accordingly, keeping the same abstract principle if the response was correct, or trying to guess a new one, in case of wrong feedback. The participant is informed that the same abstract principle is kept throughout each subtest and may change or not between subtests. Subtest VII consists in a memory task, where there is no unifying abstract principle underlying the correct responses to all the items. Instead, participants are informed that they will see items that they have seen before and that they must recall which was the correct answer the first time they saw each particular item and give that same answer again. All 20 items in subtest VII were presented before, distributed by the other 6 subtests. The scoring of the test consists of the total number of errors made.

Participants were seated in a sound-attenuated cabin during performance on the HCT. Specific instructions were given at the beginning of each subtest on the computer screen. The response consisted of pressing one of four keys numbered 1 to 4 on the computer keyboard, on each test trial. All stimuli remained on the computer screen until the participant responded. Visual feedback was provided 1500 ms after the response, indicating if the response was right (written in green) or wrong (written in red). The feedback remained on the screen for 750 ms, and there was an interval of 2000 ms before the next stimulus was displayed on the screen (see Fig. 2 for a schematic representation of a test trial).

Participants were instructed to rest at the end of each subtest. Stimuli presentation, response registration and synchronization with the EEG recording system were controlled by E-Prime software (Psychology Software Tools, Pittsburgh, PA).

2.3. EEG recordings

EEG signals were collected with a Neuroscan SynAmps2 amplifier through an Easy-Cap with 26 channels and recorded with the software Scan 4.3 (Neuroscan Systems). EEG was continuously recorded with Ag-AgCl sintered electrodes which were located according to the 10–10 system (FP1, FPz, FP2,F7, F3, Fz, F4, F8, FC3, FCz, FC4, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, PO7, PO8, O1, Oz and O2). The reference electrode was placed at the tip of the nose. Vertical EOG (VEOG) was recorded by two electrodes placed above and below the left eye and horizontal EOG (HEOG) was recorded from the outer canthi of both eyes. The impedance was kept below 5KΩ. A notch filter for 50 Hz was used during recordings.

2.4. EEG/ERP data analysis

In this work, we considered exclusively the EEG data associated with subtests III, IV, V and VI of the HCT, which tap directly spatial/positional and proportional reasoning. Subtests I and II were not included because

Fig. 1. Stimuli examples used in HCT test. More than one abstract principle can be formulated for each item, in order to correspond a number ranging from 1 to 4 to each stimulus. Participants must find out which is the correct abstract principle used in each subtest. Correct responses to the stimuli examples, from left to right: Top row 4, 2, 3; Bottom row 1, 2, 1.

they are training subtests, where participants are still learning the task. Subtest VII was also excluded because it consists of a memory subtest which relies more heavily on conceptual shifting and short term memory. Considering those subtests, the average number of right and wrong answers for each individual was 117 ± 24 and 34 ± 24, respectively.

We will here describe the signal processing steps that allowed us to obtain an artifact free signal. This signal was then firstly analyzed to identify the posterior P100 component in response to all the visual stimuli. Secondly, we analyzed the fronto-central FRN component as an index of feedback processing that differed between right and wrong responses.

2.4.1. Filtering using Singular Spectrum Analysis

Singular Spectrum Analysis (SSA) (Golyandina et al., 2001; Alexandrov, 2009) provides means to decompose a time series (signal) into components that are in-phase with the original signal (Tomé et al., 2010). It was shown that signal enhancement can be achieved by a bank of finite impulse response filters arranged as parallel pairs of analysis-synthesis filters (Tomé et al., 2010). It is particularly suitable to extract high-amplitude components by designing the analysis-synthesis pair related to the highest energy component $y[n]$. Then, the input signal $x[n]$ can be considered the sum of two components

$$x[n] = y[n] + \tilde{x}[n] \quad n = 0, 1 \ldots N-1 \tag{1}$$

where the $x\sim[n]$ represents the component of interest (Teixeira et al., 2006) which can be computed by subtracting the high- amplitude artifact $y[n]$ to $x[n]$. Assuming that the non-causal analysis-synthesis pair have an impulse response $c[n]$, with $2M-1$ amplitude values, the output signal can be expressed by the following convolution summation

$$\begin{aligned} y[n] &= \sum_{k=-(M-1)}^{M-1} c[k]x[n-k] \\ &= \sum_{k=0}^{M-1} c[k]x[n-k] + \sum_{k=1}^{M-1} c[-k]x[n+k] \end{aligned} \tag{2}$$

where $c[k] = c[-k]$ assures that the output signal is in-phase with the input. The first summation is called the causal contribution (depends on the past of the sample n-th sample) and the second summation is the anti-causal contribution (depends on the future of n-th sample) (Widmann and Schröger, 2012). The global strategy is called non-causal filtering scheme (Rousselet, 2012) and this strategy leads to a frequency response with zero-phase (Tomé et al., 2010). The output signal $y[n]$ can have the same time duration as the input. Then, assuming that input time indexes are $n = 0, 1, \ldots N-1$ the following issues need to be considered

- for the samples at $n = 0, \ldots, M-2$ the causal contribution depends on samples $x[n]$ that are unknown, for instance $\{x[-1], x[-2], \ldots\}$. Usually assumed as having value zero.
- for the samples at $n = N-M+1, \ldots, N-1$ the anti-causal contribution depends on samples of $x[n]$ that are unknown, for instance $\{x[N], x[N+1], \ldots\}$. As before usually assumed as zero.
- for the samples at $n = M-1, \ldots N-M$ are computed by applying Eq. (2) without any constraint. Those samples are usually called the stationary response of the filter while previous are called transient responses (Tomé et al., 2010). The stationary response should then correspond to segments of signals to be further processed or analyzed.

To get a further insight of the filtering operation the difference Eq. (2) can be transformed into a polynomial expression by substituting time-delayed sequence $x[n \pm k]$ by its equivalent *z-transform* $z^{\pm k}X(z)$, assuming that the transform *z-transform* of the input sequence $x[n]$ is $X(z)$. Applying a similar operation to the output sequence the transfer function $C(z) = \frac{Y(z)}{X(z)}$ is easily obtained. The frequency response of the system is then obtained by substituting $z = \exp(j2\pi f/F_s)$ into the transfer function, where $f \le F_s$ and F_s is the sampling rate in Hertz (number of samples per second). It can be shown that $C(e^{j2\pi f/F_s})$ is a weighted sum of cosine functions and therefore having only real values (Tomé et al., 2010).

2.4.2. Filter coefficients and filter design

The analysis-synthesis pair of SSA is formed by a causal filter (the analysis) whose input is the sequence $x[n]$ and the output is fed into the synthesis filter which is an anti-causal filter. The analysis- synthesis filter pairs of SSA (Tomé et al., 2010) are data-driven filters, e.g., the filter coefficients are computed based on the autocorrelation function of

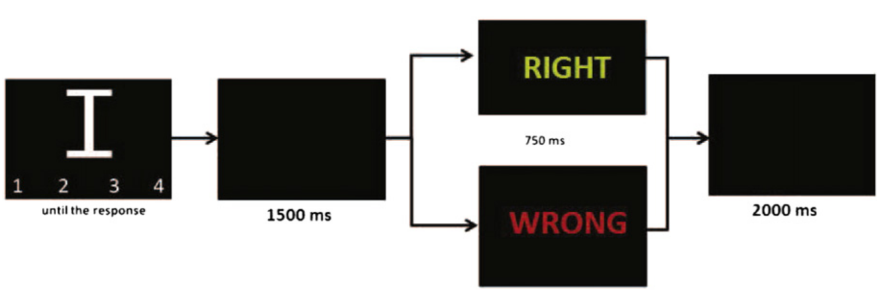

Fig. 2. Schematic representation of a test trial in the HCT.

the signal to be filtered. The autocorrelation function is the correlation of the signal with itself as a function of shifts in time, $r[m]=\mathbb{E}\{x[n]x[n+m]\}$. The values of the autocorrelation function are used to estimate the correlation matrix whose eigen-decomposition gives the necessary information to calculate the filter coefficients. The eigenvectors of the matrix are used to estimate the coefficients $c[n]$ and the corresponding eigenvalues are related with the energy of the filters output.

The original signals suffered from a strong interference of high amplitude artifacts, such as participant movements or ocular movements. The EOG signal from each ERP was used to calculate the filter coefficients ensuring a custom filtering. Therefore, for each participant, the EOG average was used to compute the correlation matrix with entries of the autocorrelation function computed with $M = 101$ time shifts. The eigenvector corresponding to the largest eigenvalue, of the eigendecomposition of the correlation matrix, was used to compute the analysis-synthesis pair as described before, originating a filter $c[n]$ with 201 coefficients. Thus, only the frequencies associated with high amplitude movements present in each EOG channel that usually propagate to the remaining channels will be eliminated.

Fig. 3 represents the frequency response and impulse response, respectively, of one filter designed to process the EEG segments. The frequency response shows that the pass-band is $f<3$Hz and the first zero around $f=10$Hz which corresponds to $\simeq F_s/M$. The filter stop-band shows that the frequency zone around ≈ 10Hz and their multiples are strongly attenuated. Therefore the frequency contents of the filter output $y[n]$ must have most of their energy concentrated up to 3*Hz* and in particular no energy around the 10*Hz* band (Fig. 3 - left).

The same filter coefficients $c[n]$ were applied to all channels in order to always have the same gain in amplitude. Finally, the corrected ERP signal was obtained by subtracting the extracted components to the original ERP $x\sim[n]=x[n]-y[n]$. Note that the corrected signal will preserve the alpha and beta bands of the original signal $x[n]$.

3. Results

In this work segments of signals time-locked either with visual stimulus (t_v) or with response (t_r) onsets were considered. Therefore the segments, with duration of 4*s*, were taken from t_*-1000ms to t_*+3000ms, where $t_*=\{t_v,t_r\}$, are the visual stimulus and the response onsets. Then the filter was applied to the average signal and the corrected version was obtained subtracting to the original signal. The filtering technique that was applied in the present work allowed to center the data. Therefore, no baseline removal was necessary. Note that the length of the segments ensures that the transient response of the filter does not include the time span of the ERP events under study.

3.1. P100

The segments centered in the visual stimulus, i.e., the first analysis window, were used to demonstrate the impact of the application of this method and its ability to clean the high amplitude artifacts in the signal. After filtering, in this time window it is possible to clearly identify the posterior P100 visual ERP component, which is associated with visual processing and is elicited by all visual stimuli, independently of the task (Luck, 2005). The segments were epoched time-locked to the onset of the visual stimuli, [0 1000] ms. Fig. 4 shows (a) the original average signal $x[n]$ of one participant and (b) the corresponding corrected version $x\sim[n]$. As can be observed, all high amplitude artifacts were adequately removed and the peak around 100 ms is clearly visible in occipital regions.

The grand average of the original and the corrected signal of all participants was also computed. Fig. 5 (a) presents the grand-average in channel O1 in the clean signal. To confirm the effect of the filtering operation described above, the head topography of the grand average waveforms considering a visualization window between [100 120] ms after the visual stimuli onset is presented, Fig. 5 (b) and (c), for original and corrected signal, respectively. As can be observed, whereas in the original signals (b), there is no clear evidence of the posterior positivity that characterizes the P100, this positivity is evident in the corrected signals (c).

3.2. FRN analysis

As before segments, with 4*s*, time locked to the response onset, were averaged according to the condition (right or wrong) and then filtered to obtain $y[n]$ and then subtracted to the original in order to have a corrected version $x\sim[n]$. For the FRN analysis, the EEG data segments centered on response onset were epoched between 100 ms prior to feedback onset to 400 ms after it. The more conspicuous FRN effects were observed in channels FCz and Cz, in agreement with the existing literature. Fig. 6 displays the grand-average ERP waveforms for these channels considering the two subsets of trials corresponding to the Right and Wrong answers, time-locked to the feedback. Thus, the value 0 ms corresponds to the moment when the feedback occurred. In the displayed waveforms, we can observe large negative components after the feedback, peaking around 250 ms, which are consistent with the feedback-related negativity. As can be observed, the FRN wave corresponding to the wrong answers (dashed line) is more negative than for the right answers (solid line).

Fig. 7 represents the topography of the grand-average waveforms considering a visualization time window centered at 250 ms. The results show that this negative deflection was strongest at the electrode FCz, and exhibited a fronto-central scalp distribution. Furthermore, an apparent difference between the Wrong and Right conditions is observed, with more negative amplitudes for the Wrong trials.

To assess the FRN differences between right and wrong trials, the average amplitudes and peak latencies in the window [220 260] ms after the feedback were analyzed. FRN latencies were estimated as the minimum peak value found in the window [220 260] ms after the feedback onset. Following other authors, we opted to run our analyses with the

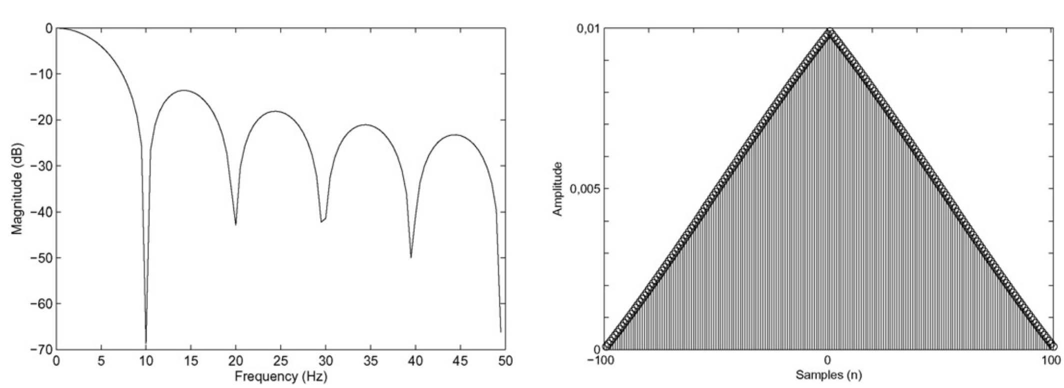

Fig. 3. Illustration of filter design for an ERPr segment. *Left* - Frequency response of the system ($0<f<50$Hz); *right* - Impulse response $c[n]$.

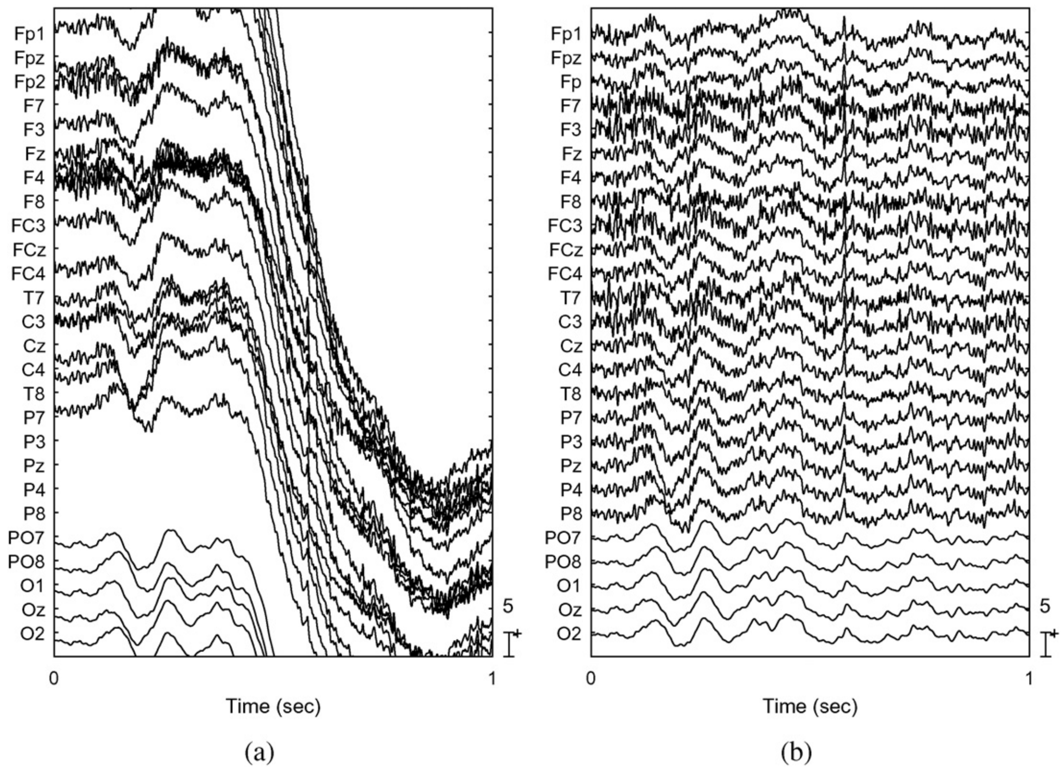

Fig. 4. Average signal of one participant, when epochs were time-locked to the onset of the visual stimuli, to reflect the filtering process in time: (a) Original signal; (b) Corrected signal.

average amplitudes because mean amplitude measures are less affected by differences in the number of trials between conditions (Bellebaum and Daum, 2008; Luck, 2005), which was the case in the present study. For the analysis, the channels FC3, FCz, FC4 and Cz were considered, as they were the channels where the FRN was more evident after visual inspection. Two 2×4 repeated measures ANOVAs were carried out, with error condition (right and wrong) and electrode (FC3, FCz, FC4 and Cz) as within-subjects factors, considering both average amplitude and peak latency as the dependent variables. The Greenhouse-Geisser correction was applied for violations of sphericity and the Bonferroni adjustment was applied to multiple comparisons. Regarding amplitude, there was a significant main effect of error condition, $F(1,57) = 16.29$, $p < .001$, partial $\eta^2 = .222$, where the wrong trials elicited significantly more negative amplitudes ($M_{amp} = -1.67\mu V$) than the right trials ($M_{amp} = -0.76\mu V$). There was also a significant main effect of electrode, $F(2.44,138.86) = 37.17$, $p < .001$, partial $\eta^2 = .40$. The midline electrodes FCz ($M_{amp} = -1.46\mu V$) and Cz ($M_{amp} = -1.58\mu V$) elicited significantly more negative amplitudes than FC3 ($M_{amp} = -1.04\mu V$) and FC4 ($M = -0.77\mu V$). The interaction between error condition and electrode was not significant, $F(2.68,152.99) = 1.004$, $p = .392$, partial $\eta^2 = .017$. Regarding peak latencies, there was a significant main effect of electrode, $F(3,171) = 6.31$, $p < .001$, partial $\eta^2 = .10$, where Cz

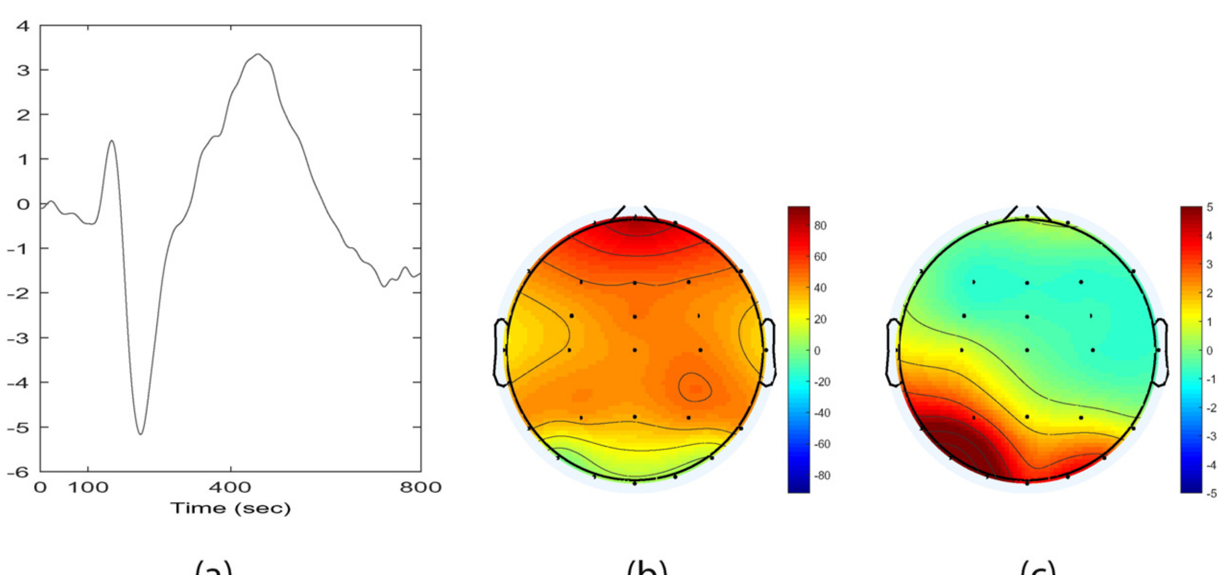

Fig. 5. (a) Grand-average waveform in channel O1 time-locked to the onset of the visual stimuli. (b) and (c) Head topography of the grand-average waveforms considering a visualization window between [100 120]ms: (b) Original signal and (c) Corrected signal.

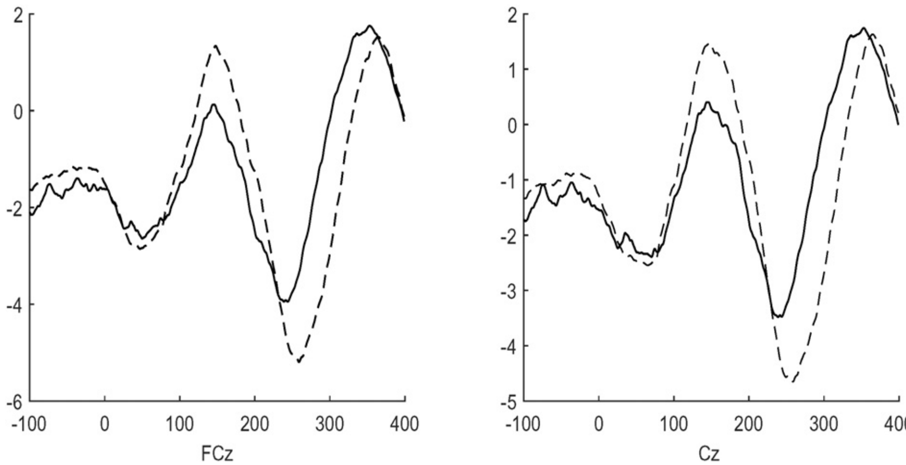

Fig. 6. Grand-average waveforms of individual ERPs for FCz and Cz channels considering two subsets of trials: Wrong responses (*dash line*) and Right responses (*solid line*)

exhibited significantly longer latencies ($M_{lat} = 253$ms) than FC4 ($M_{lat} = 250$ms). Latencies in FC3 and FCz (both $M_{lat} = 250$ms) did not differ significantly from the other locations. Neither the main effect of error condition, $F(1,57) = 1.82$, $p = .182$, partial $\eta^2 = .031$, nor the interaction between error condition and electrode, $F(2.32,132.51) = 1.57$, $p = .209$, partial $\eta^2 = .027$, were significant for the peak latency. Thus, results indicate, as expected, that errors elicited significantly more negative FRN amplitudes than correct responses. Furthermore, there were no significant differences in peak latency.

4. Discussion and conclusions

The present work aimed to demonstrate that it is possible to study event-related potentials during the performance on a non-modified version of the Halstead Category Test (HCT) (DeFilippis, 2002), a neuropsychological instrument clinically used to explore cognitive executive function associated with the ability to formulate and apply abstract principles. The possibility of coupling neuropsychological data with information about brain dynamic activity associated with performance on the test can usefully increase the clinical potential of the test, allowing for a higher specificity in detecting brain dysfunction, an issue that has received some criticism in the past (Allen et al., 2007; Choca et al., 1997). It is important to use a non-modified version of the HCT to be able to use normative data and cut-off scores for clinical diagnosis. Since the clinical version of the test is not a classical paradigm for ERP recordings (being a time unconstrained task, where participants are free to visually explore the stimuli before making a response), we aimed to demonstrate that it is possible to use a SSA based filtering technique (Tomé et al., 2010) to process contaminated signal with high amplitude ocular and movement artifacts. As shown in this paper, this technique allows obtaining a clean signal with minimal loss of relevant information and no phase or latency distortions. More traditional approaches to processing ERP data, where segments containing artifacts are typically eliminated, would implicate the loss of a considerable amount of trials (Luck, 2005). This would result in a very low signal-to-noise ratio, since the clinical version of the test also contains a limited number of trials.

The principle underlying each subtest of the HCT (except the seventh, which is a memory subtest) is the same throughout the entire subtest (but it may change from one subtest to another), and participants receive feedback on their performance after each trial. Thus, based on that feedback, they should opt for maintaining the same rule for the next stimulus, if performance was correct, or they should change their reasoning and try to extract a different abstract principle, if performance was incorrect, until they discover the general principle underlying that particular subtest. Due to the nature and difficulty of the task, there is a high likelihood of committing errors and the degree of uncertainty in the outcome of each trial is considerably high until the moment when the participant is confident that he/she has determined the underlying abstract principle of that subtest. Thus, performance on such a task requires that participants monitor the outcome of their behavior for errors and change their course of action accordingly, in an effort to avoid making another error in the subsequent trial. Hence, we expected that a feedback-related negativity (FRN) wave would be elicited in

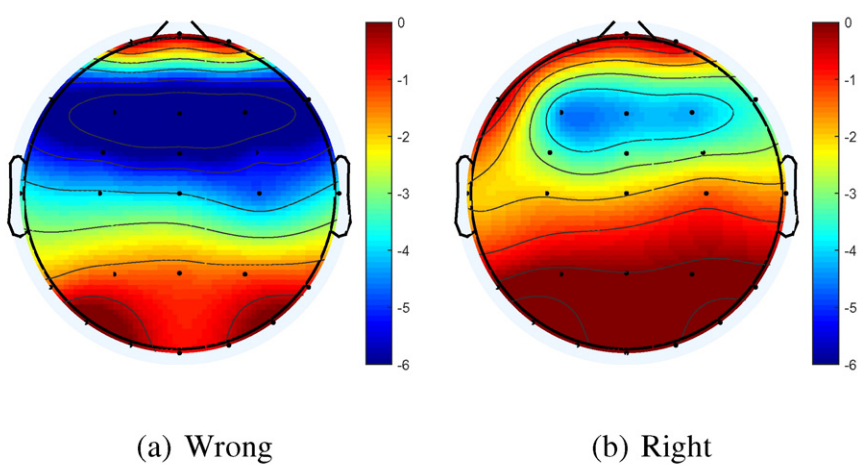

Fig. 7. Head topography of the grand-average waveforms considering a visualization window centered at 250 ms: (a) Wrong trials and (b) Right trials

response to feedback stimuli, which would be more negative for error trials (when feedback was negative) than for correct trials (when feedback was positive). This ERP pattern is consistent with an error monitoring process (Hajcak et al., 2006; Miltner et al., 1997).

After data processing and artifact removal through SSA filtering, a P100 visual component associated with the processing of all visual stimuli (Pratt, 2011) was clearly identified in epochs time-locked with the onset of the stimulus, which confirms the adequacy of the filtering procedure. Importantly, a FRN component was also identified in epochs time-locked with the feedback onset. This component had a fronto-central topographical distribution and peaked around 250 ms after feedback onset. As predicted, error trials elicited significantly more negative amplitudes than correct trials, particularly at midline fronto-central electrode sites. The FRN is normally observed in tasks difficult enough that the subjects do not know the accuracy of their judgments until the feedback occurs, i.e., tasks with a certain degree of uncertainty (e.g., Gehring and Willoughby, 2002), which is consistent with performance on the HCT. Thus, this result supports the idea that performance on the HCT involves error monitoring processes, which allow participants to adjust their response strategy in face of negative feedback, in order to find the correct underlying rule/abstract principle. The present finding is consistent with previous studies using another executive function test, the Wisconsin Card Sorting Test (WCST), which requires partially similar cognitive processes (Kraus and Horowitz-Kraus, 2014; Vilà-Balló et al., 2015). In summary, the present study demonstrated that the application of a nonzero phase filtering technique successfully removed high-amplitude ocular and movement artifacts from EEG signals. In order to guarantee the same gain distortion, the SSA filter coefficients were computed only once by using the corresponding EOG average signal.

The HCT is a well-established neuropsychological measure of nonverbal reasoning, abstract concept formation and cognitive flexibility, which are aspects of executive function. The identification of the FRN component associated with feedback processing during performance on the HCT is consistent with the role of the anterior cingulate cortex (ACC), the most likely neural generator of the FRN (Hauser et al., 2014; Walsh and Anderson, 2012), in error detection and conflict monitoring. The present results are compatible with views of ACC function that link this structure to performance-monitoring processes and predict increased ACC activity and, consequently, larger FRN amplitudes, when performance is poor (Miltner et al., 1997; Holroyd and Coles, 2002). However, alternative views of ACC function, such as the predicted response-outcome model (Alexander and Brown, 2011), propose that the main function of the ACC is to predict the likely outcomes of actions and to signal when an outcome is unexpected, independently of that outcome being good or bad. Thus, according to this theory, an increased FRN should be elicited after unexpected outcomes regardless of their valence. Evidence that similar FRN amplitudes for negative and positive unexpected feedback occur when the unexpectedness of both types of feedback is equivalent has provided support for this theory (Ferdinand et al., 2012). In the present study, negative feedback is less frequent than positive feedback and therefore it is not possible to test these different theories. This was indeed not the aim of this work, as we intended to maintain the original format of the HCT. Future studies with different objectives may introduce an adaptation of the HCT structure in order to balance the frequencies of positive and negative feedback, and thus be able to specifically explore the different predictions made by theories about ACC function regarding the FRN component as an electrophysiological correlate of performance-monitoring.

At the clinical level, the present results may contribute to a better understanding of the HCT as a neuropsychological assessment instrument, as well as to its clinical utility. The possibility of coupling behavioral performance on the test with the recording of ERPs will potentially increase the specific information that can be extracted from the test when applying it to individuals who may suffer from altered feedback processing. This might affect general performance on the HCT, but a global error score might not be sensitive enough to pinpoint the specific cognitive ability that is impaired. Using the FRN as an electrophysiological marker of feedback processing might help to identify specific deficits in those processes (e.g., Vilà-Balló et al. 2015). To this end, future research should explore the possibility of analyzing single-subject ERP correlates of performance on the HCT, since there is evidence that error-related potentials can be measured on a single-trial basis (e.g., Falkenstein et al., 2000; Wiersema et al., 2005). In future studies, it would also be interesting to explore how the FRN and other error-related components vary as the learning of the new rules proceeds during performance on the HCT. It is possible that the FRN is larger at the beginning of the task, when the correct response is yet unknown, whereas other error-related components, such as the ERN, which are initially absent, may become larger when the participant has learned the adequate response (Hoffmann and Falkenstein, 2012). A trial-by-trial analysis would also be critical to understanding these dynamic changes in the correlates of the error-monitoring neural system as learning progresses.

Acknowledgments

This research was supported by Bial Foundation with Grant ref. 136/08 to Isabel M. Santos and by funding from Foundation for Science and Technology (FCT) and the POPH/FSE Program (grant reference SFRH/BPD/101112/2014) to Ana R. Teixeira. The funding agency had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Adams, K.M., Gilman, S., Koeppe, R., Kluin, K., Junck, L., Lohman, M., Johnson-Greene, D., Berent, S., Dede, D., Kroll, P., 1995. Correlation of neuropsychological function with cerebral metabolic rate in subdivisions of frontal lobes of older alcoholic patients measured with [18F] fluorodeoxyglucose and positron emission tomography. Neuropsychology 9, 275–280. http://dx.doi.org/10.1037/0894-4105.9.3.275.

Alexander, W.H., Brown, J.W., 2011. Medial prefrontal cortex as an action-outcome predictor. Nat. Neurosci. 14, 1338–1344. http://dx.doi.org/10.1038/nn.2921.

Alexandrov, T., 2009. A method of trend extraction using singular spectrum analysis. REVSTAT 7, 1–22.

Allen, D.N., Goldstein, G., Mariano, E., 1999. Is the Halstead Category Test a multidimensional instrument? J. Clin. Exp. Neuropsychol. 21, 237–244. http://dx.doi.org/10.1076/jcen.21.2.237.926.

Allen, D.N., Caron, J.E., Duke, L.A., Goldstein, G., 2007. Sensitivity of the Halstead Category Test factor scores to brain damage. Clin. Neuropsychol. 21, 638–652. http://dx.doi.org/10.1080/13854040600744821.

Anderson, S.W., Damasio, H., Jones, R.D., Tranel, D., 1991. Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. J. Clin. Exp. Neuropsychol. 13, 909–922. http://dx.doi.org/10.1080/01688639108405107.

Barceló, F., 2003. The Madrid Card Sorting Test (MCST): a task switching paradigm to study executive attention with event-related potentials. Brain Res. Protocol. 11, 27–37. http://dx.doi.org/10.1016/S1385-299X(03)00013-8.

Barceló, F., Rubia, F.J., 1998. Non-frontal P3b-like activity evoked by the Wisconsin Card Sorting Test. Neuroreport 9, 747–751. http://dx.doi.org/10.1097/00001756-199803090-00034.

Barceló, F., Sanz, M., Molina, V., Rubia, F.J., 1997. The Wisconsin Card Sorting Test and the assessment of frontal function: a validation study with event-related potentials. Neuropsychologia 35, 399–408.

Bellebaum, C., Daum, I., 2008. Learning-related changes in reward expectancy are reflected in the feedback-related negativity. Eur. J. Neurosci. 27, 1823–1835. http://dx.doi.org/10.1111/j.1460-9568.2008.06138.x.

Bernat, E.M., Nelson, L.D., Steele, V.R., Gehring, W.J., Patrick, C.J., 2011. Externalizing psychopathology and gain/loss feedback in a simulated gambling task: dissociable components of brain response revealed by time-frequency analysis. J. Abnorm. Psychol. 120, 352–364. http://dx.doi.org/10.1037/a0022124.

Choca, J.P., Laatsch, L., Wetzel, L., Agresti, A., 1997. The Halstead Category Test: a fifty year perspective. Neuropsychol. Rev. 7, 61–75. http://dx.doi.org/10.1023/b:nerv.0000005944.98635.16.

DeFilippis, N.A., 2002. The Category Test, Computer Version. Psychological Assessment Resources.

DeFilippis, N.A., McCampbell, E., 1997. The Booklet Category Test. Psychological Assessment Resources.

Falkenstein, M., Hoormann, J., Christ, S., Hohnsbein, J., 2000. ERP components on reaction errors and their functional significance: a tutorial. Biol. Psychol. 51, 87–107. http://dx.doi.org/10.1016/S0301-0511(99)00031-9.

Ferdinand, N.K., Mecklinger, A., Kray, J., Gehring, W.J., 2012. The processing of unexpected positive response outcomes in the mediofrontal cortex. J. Neurosci. 32, 12087–12092. http://dx.doi.org/10.1523/JNEUROSCI.1410-12.2012.

Gehring, W.J., Willoughby, A.R., 2002. The medial frontal cortex and the rapid processing of monetary gains and losses. Science 295, 2279–2282. http://dx.doi.org/10.1126/science.1066893.
Golyandina, N., Nekrutkin, V., Zhigljavsky, A., 2001. Analysis of Time Series Structure: SSA and Related Techniques. Chapman & HALL/CRC.
Hajcak, G., Moser, J.S., Holroyd, C.B., Simons, R.F., 2006. The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. Biol. Psychol. 71, 148–154. http://dx.doi.org/10.1016/j.biopsycho.2005.04.001.
Hauser, T.U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., Brem, S., 2014. The feedback-related negativity (FRN) revisited: new insights into the localization, meaning and network organization. NeuroImage 84, 159–168. http://dx.doi.org/10.1016/j.neuroimage.2013.08.028.
Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., Curtiss, G., 1993. Wisconsin Card Sorting Test Manual: Revised and Expanded. Psychological Assessment Resources, Odessa.
Hoffmann, S., Falkenstein, M., 2012. Predictive information processing in the brain: errors and response monitoring. Int. J. Psychophysiol. 83, 208–212. http://dx.doi.org/10.1016/j.ijpsycho.2011.11.015.
Holroyd, C.B., 2004. A note on the oddball N200 and the feedback ERN. Learning and cognitive control laboratory. In: Ullsperge, M., Falkenstein, M. (Eds.), Errors, Conflicts, and the Brain: Current Opinions on Performance Monitoring. Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, pp. 211–218.
Holroyd, C.B., Coles, M.G., 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. 109, 679–709. http://dx.doi.org/10.1037/0033-295X.109.4.679.
Kraus, D., Horowitz-Kraus, T., 2014. The effect of learning on feedback-related potentials in adolescents with dyslexia: an EEG-ERP study. PLoS One 9, e100486. http://dx.doi.org/10.1371/journal.pone.0100486.
Luck, S.J., 2005. An Introduction to the Event-related Potential Technique. MIT Press.
McNally, S., Dsurney, J., McGovern, J., DeFilippis, N., Chan, L., 2015. Concurrent validity of new subscale scores for the Booklet Category Test. Assessment 1–9. http://dx.doi.org/10.1177/1073191115588783.
Miltner, W.H.R., Braun, C.H., Coles, M.G.H., 1997. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a generic neural system for error detection. J. Cogn. Neurosci. 9, 788–798. http://dx.doi.org/10.1162/jocn.1997.9.6.788.
Minassian, A., Perry, W., Carlson, M., Pelham, M., DeFilippis, N., 2003. The category test perseveration, loss of set, and memory scales: three new scales and their relationship to executive functioning measures. Assessment 10, 213–221. http://dx.doi.org/10.1177/1073191103253498.
Pratt, H., 2011. Sensory ERP components. In: Kappenman, E.S., Luck, S.J. (Eds.), The Oxford Handbook of Event-related Potential Components. Oxford University Press, New York, pp. 89–114. http://dx.doi.org/10.1093/oxfordhb/9780195374148.013.0050.
Rousselet, G.A., 2012. Does filtering preclude us from studying ERP time-courses? Front. Psychol. 3. http://dx.doi.org/10.3389/fpsyg.2012.00131.
Teixeira, A.R., Tomé, A.M., Lang, E.W., Gruber, P., da Silva, A.M., 2006. Automatic removal of high-amplitude artifacts from single-channnel electroencephalograms. Comput. Methods Prog. Biomed. 83, 125–138. http://dx.doi.org/10.1016/j.cmpb.2006.06.003.
Tomé, A.M., Teixeira, A.R., Figueiredo, N., Santos, I.M., Georgieva, P., Lang, E.W., 2010. SSA of biomedical signals: a linear invariant systems approach. Stat. Interface 3, 345–355. http://dx.doi.org/10.4310/SII.2010.v3.n3.a.
Vilà-Balló, A., Cunillera, T., Rostan, C., Hdez-Lafuente, P., Fuentemilla, L., Rodríguez-Fornells, A., 2015. Neurophysiological correlates of cognitive flexibility and feedback processing in violent juvenile offenders. Brain Res. 1610, 98–109. http://dx.doi.org/10.1016/j.brainres.2015.03.040.
Walsh, M.M., Anderson, J.R., 2012. Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. Neurosci. Biobehav. Rev. 36, 1870–1884. http://dx.doi.org/10.1016/j.neubiorev.2012.05.008.
Widmann, A., Schröger, E., 2012. Filter effects and filter artifacts in the analysis of electrophysiological data. Front. Psychol. 3. http://dx.doi.org/10.3389/fpsyg.2012.00233.
Wiersema, J.R., van der Meere, J.J., Roeyers, H., 2005. ERP correlates of impaired error monitoring in children with ADHD. J. Neural Transm. 112, 1417–1430. http://dx.doi.org/10.1007/s00702-005-0276-6.