Neuroscience Letters 444 (2008) 64–68

Contents lists available at ScienceDirect

Neuroscience Letters

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

Brain potentials reveal the role of conflict in human errorful and errorless learning

Marcus Heldmann[a,*], Ursula Markgraf[b], Antoni Rodríguez-Fornells[c,d], Thomas F. Münte[b,e]

[a] Department of Neurology II, Otto von Guericke University, Leipziger Strasse 44, 39120 Magdeburg, Germany
[b] Department of Neuropsychology, Otto von Guericke University, 39106 Magdeburg, Germany
[c] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[d] Faculty of Psychology, University of Barcelona, 08035 Barcelona, Spain
[e] Center for Behavioral Brain Sciences, Magdeburg, Germany

ARTICLE INFO

Article history:
Received 17 April 2008
Received in revised form 15 July 2008
Accepted 17 July 2008

Keywords:
Errorless learning
Memory
ERN
Conflict model
Errorlikelihood model

ABSTRACT

The avoidance of errors during learning, so-called errorless learning, results in increased memory performance. In the present study subjects had to learn items in an errorful and an errorless manner. After each learning session learned items were presented again, but now intermixed with items not learned before. Response-locked event-related potentials were used to investigate the neural underpinnings of cognitive control mechanisms during recognition of items learned under errorless and errorful conditions. Irrespective of the response's correctness a typical error-related negativity (ERN) was observed for items classified as learned before. In contrast to the apparent difference in memory performance between learning modes, ERN amplitudes to hits and false alarms were not different. The present pattern of results can be explained neither in terms of error detection nor the conflict monitoring account. Instead we are trying to argue that the findings add support to a theoretical proposal which posits that variations of the ERN amplitude can be best explained by the subjects' perceived likelihood of making an error.

Errors in memory retrieval may occur because remembering is often a reconstructive process [9,13,15]. One possible cause for problems with recall or recognition are errors made during the encoding process of the critical memory traces (e.g. when exposed in parallel to multiple new persons/faces that have to be associated with the appropriate names). Indeed, preventing people or animals [21] from making errors during encoding, henceforth "errorless learning" (EL), results in increased memory performance compared to "errorful" (EF) trial-and-error learning [2]. The benefit of errorless learning has been explained by the reduction of competition from false memories [2]. To deal with such competition during memory retrieval, cognitive control mechanisms are thought to supervise the results of retrieval operations [16].

One well-known correlate of cognitive control processes is the error-related negativity (ERN), an event-related potential (ERP) occurring about 70–100 ms after an erroneous response at frontocentral electrode sites [8,10]. The major part of this component is generated in the anterior cingulate cortex (ACC) [7]. Contrasting EL and EF learning modes with ERPs recorded during memory retrieval, Rodriguez-Fornells et al. [19] reported the ERN to be sensitive to conflicting memory traces during retrieval. These authors used a variant of the word fragment completion task during learning to induce multiple competing memory traces for a given fragment in the EF condition, while only one item was introduced in the EL condition. During the retrieval phase participants had to decide by button press whether the presented word had been learned before. For the errorful learning condition sizable ERN amplitudes were reported for items that were rated as having been previously learned irrespective of whether this was in fact the case (hits) or not (false alarms). This pattern of results was interpreted in terms of the conflict monitoring theory of the ERN [3,6], which holds that the ERN reflects competing response tendencies. In this particular case, it might reflect the evaluation of the output of the retrieval process. Interestingly, for items which were not recognized although learned in the study phase (miss trials) or for the items which were correctly rejected, no ERN was observed. This finding may be taken to suggest that the post-retrieval checking mechanism is engaged only if retrieval leads to a familiarity and/or recollection (episodic) experience (as can be assumed for hits or false alarms).

However, the pattern for EL items in this experiment was less well explained. Although presumably no conflicting memory traces were active during recollection, an ERN was reported for correctly recognized items (hits), which was, however, much smaller than

---

[*] Corresponding author. Fax: +49 391 6117540.
E-mail address: marcus.heldmann@med.ovgu.de (M. Heldmann).

0304-3940/$ – see front matter © 2008 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.neulet.2008.07.042

the one obtained for EL false alarm responses. It was speculated that this result could be caused by the fact that during the study phase EL and EF items were presented randomly intermixed. As a consequence, a spillover effect coming from the non-target activations of the EF trials might have caused the ERN to the EL hits. If this interpretation is correct, the ERN to EL hits should disappear or at least be greatly diminished in amplitude, if EL and EF learning conditions were presented in separate blocks. An alternative explanation is that the larger ERN observed in the EL condition for false alarms compared to hits is related to the larger salience of the error (these items were in fact new words which had never been presented before in the study phase).

The present study thus seeks to further investigate the interesting distinction between errorful and errorless learning (see [1]) and to replicate the previous findings in relation to the ERN component [19]. In the present design we are trying to address three critical aspects of the Rodriguez-Fornells et al. [19] study: (i) to avoid the spillover effect possibly caused by the intermixed presentation of EL und EF items we used a blocked design with separate sessions for the two conditions; (ii) to get a more distinct control condition new words unrelated to the learned items were introduced in the recognition phase for both conditions; (iii) to make the depth of word processing in the EL condition more comparable to the EF condition a sentence production task was introduced.

Twenty-six right-handed neurological healthy volunteers participated in this study after giving written informed consent. Three subjects were excluded because of technical artifacts and 12 because of too low false alarm rates. Hence, data of 11 participants (eight women, mean age 23.1 years) are reported. The study protocol was approved by the Ethics Committee of Magdeburg University.

Subjects participated in one EF and one EL learning session spaced 3–7 days apart with the order of learning conditions counterbalanced across subjects. Each session comprised five runs each of which featured a learning phase and a subsequent recognition phase. During the learning phase subjects had to perform a word fragment completion task for 30 word fragments each (nearly identical to [19]). In the EF condition the first three letters of a word were given and the subjects had to come up with three alternatives to complete this fragment to a word. After guessing the experimenter revealed which word was the target and asked the subject to learn it. If subjects failed to guess the intended target word, it was introduced by the experimenter and learned by the subject. In the EL learning condition the first three letters of the word were introduced followed by the full to-be-remembered word ("I am thinking about a word starting with A–N–Z, the word is 'Anzeige'"). To ensure a deeper processing of words in the errorless condition participants had to produce a sentence containing the to-be-remembered word.

For each of the presented word fragments at least two German words exist with a high and comparable guessing probability (word fragment completion norms were kindly provided by A. Richardson-Klavehn and E. Düzel). For each fragment one word was used during the learning phase as target word, while the other was used as distracter during the recognition phase. During the learning phase of the EF condition it was made sure that the distracter words were produced by the participants during the guessing process and to that effect a competing memory trace was created. During each recognition phase subjects were presented with 30 targets, 30 distracters and 30 additional new words in a randomized order. The participants' task was to indicate by button press (left/right index finger, target buttons were counterbalanced across subjects), whether or not a given word was a target word or not. The words were presented in white letters on a black background in the middle of a computer screen. The stimuli subtended 0.57° of visual angle in height and between 1.7° and 4.9° in width. The SOA was 1800–2500 ms, the stimulus duration 300 ms.

The EEG was recorded using tin electrodes (impedances <5 kΩ) mounted in an elastic cap, located at 29 standard positions (Fp1/2, F3/4, C3/4, P3/4, O1/O2, F7/8, P7/8, T3/4, Fc5/6, Cp5/6, Pg1/2, Cp1/2, Fpz, Fz, Cz, Pz), and referenced to the left mastoid process. Offline, scalp electrodes were re-referenced to the mean activity at the left and right mastoid process. To enable the offline rejection of eye movement artifacts, horizontal and vertical electrooculograms (EOGs) were recorded using bipolar montages. All channels were amplified (bandpass 0.05–30 Hz) and digitized with 4 ms resolution. Using individualized amplitude criteria on the eye channels, trials with eye movement artifacts were excluded from the analysis. Response-locked ERPs were averaged for epochs of 900 ms length, starting 300 ms prior to each response and 1–8 Hz bandpass filtered. For each condition and each subject the mean amplitude at electrodes Fz, Cz and Pz in the time window 0–100 ms (baseline −300 to 0) after response was calculated and entered into analyses of variance (ANOVA). The combination of learning condition (EF/EL), stimulus type (target, non-target and new word) and response (correct/incorrect) results in 12 different categories. Because the frequency of false alarms for new words was near to zero, this category was neglected in the analysis. Note that in the Rodriguez-Fornells study no really new words were used, instead they labeled the correct rejections (CRs) of the EL condition as new words, because these distracter words had not been produced by or shown to the participants during the learning phase. For hypothesis testing we calculated repeated measurement ANOVAs. Prior to this analysis data were subjected to a vector-normalization procedure ([14], for a critical discussion see [22]) to remove overall amplitude differences of ERN. The Huynh–Feldt epsilon coefficient was applied to correct ANOVAs for non-sphericity. The original degrees of freedom but corrected *p*-values will be reported. Following the ANOVAs the original, non-vector normalized mean amplitudes were used to calculate planned comparisons testing for differences between hits, false alarms and new corrections within each learning condition as well as for response category differences between learning conditions.

The signal detection measure d' revealed that the accuracy was significantly better for errorless ($d' = 2.89$) compared to errorful trials ($d' = 1.25$; $t = 10.5$, $p < 0.001$, d.f. = 10). The memory accuracy was thus much higher than in Rodriguez-Fornells et al. [19], who had used intermixed presentation of EL and EF words during learning (EL = 1.04, EF = 0.45). Repeated measures ANOVA for RTs with the factor learning mode (EF) and response type (hit/false alarm/correct rejection/new words correct rejection) indicated significantly faster response for EL learned items (mean RT EL: 781 ms, EF: 939 ms; $F = 5.59$, $p = 0.04$, d.f. = 1,10). The main effect response type ($F = 2.69$, $p = 0.13$, d.f. = 4,40) as well as the interaction learning mode × response type ($F = 1.12$, $p = 0.31$, d.f. = 4,40) failed to reach significance level.

Response-locked averages (Fig. 1) show a negativity for hits and false alarms compared to correctly rejected new words in both learning conditions. The corresponding topographical maps (Fig. 3) show a fronto-central maximum typical for the ERN with some spread of activity to more posterior sites. A repeated measures ANOVA with the factors learning mode (EL/EF), response type (hit/false alarm/new correct rejection) and electrode site (Fz/Cz/Pz) revealed significant response type ($F = 11.42$, $p < 0.001$) and electrode site ($F = 3.63$, $p < 0.05$, both d.f. = 2,20) main effects. Neither the learning mode main effect ($F = 0.16$, d.f. = 1,10) nor any of the resulting interactions were significant (learning mode × response type, $F = 1.68$; learning mode × electrode site, $F = 0.8$; response type × electrode site, $F = 1.91$; all d.f. = 2,20; learning mode × response type × electrode site, $F = 1.81$, d.f. = 4,40).

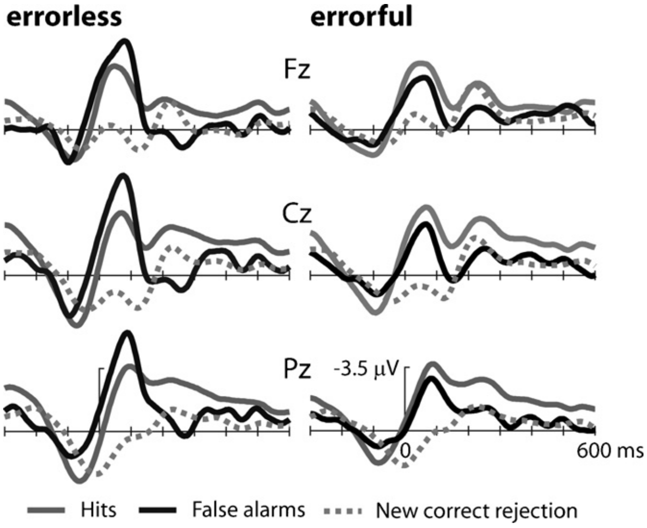

Fig. 1. Response-locked ERPs, filtered with bandpass 1–8 Hz, baseline 300 ms. Hits and false alarms are associated with a negativity peaking approximately 100 ms after the button press, thus having the typical characteristics of an ERN.

Planned comparisons within learning modes did not show differences of ERN amplitudes between hits and false alarms (EL: $F = 3.09$, $p = 0.11$; EF: $F = 4.31$, $p = 0.064$, both d.f. = 10), but comparisons of hits vs. new CRs (EL: $F = 56.00$; EF: $F = 32.27$; both $p < 0.001$) and false alarms vs. new CRs (EL: $F = 22.51$, $p < 0.001$; EF: $F = 10.56$, $p = 0.008$; all d.f. = 10) did.

Comparing the ERN to EL/hits and EF/hits or to EL/new CRs and EF/new CRs did not show a significant difference (all $F < 2.13$, all $p > 0.17$, all d.f. = 10), whereas the comparison between EL/FA and EF/FA did ($F = 5.43$, $p = 0.04$, d.f. = 10).

Fig. 2A illustrates waveforms of the further event types (misses, CR to distracter words) at Cz together with hits and CRs to new words. Although misses were target words erroneously classified as not seen before and CR to distracter words and CR were correct responses, no ERN-like waveform can be shown. The corresponding repeated measures ANOVA with the factors learning type, response type (misses/CR/new CR) and electrode site resulted in a significant main effect for electrode site and two significant interactions (electrode site × response type ($F = 3.77$, $p < 0.05$, d.f. = 4,40) and electrode site × learning mode ($F = 3.89$, $p < 0.05$, d.f. = 2,20)). The triple interaction learning mode × response type × electrode site ($F = 1.68$, d.f. = 4,40) as well as the interaction learning mode × response type ($F = 0.07$, d.f. = 2,20) failed to become significant.

In the present study errorless learning (compared to errorful learning) had a profound effect on the quality in memory performance of healthy subjects. In contrast, no impact of learning mode on the ERN amplitude was observed, but a significant influence of the subjects' judgment regarding whether or not an item had to be learned during the study phase. Surprisingly, the correctness of the subjects' recognition had no significant influence on ERN amplitude variations.

The memory performance was increased quite dramatically compared to the Rodriguez-Fornells et al. [19] study by using a blocked presentation of learning conditions and by requiring participants to produce a sentence containing the target word. Interestingly, not only items learned in the EL condition but also those learned in the EF mode benefited from blocked presentation. Importantly, the *difference* between EL and EF was enhanced by blocked presentation. The behavioral results of this study underscore the potential of EL learning in particular in patients with memory disturbance [12,18] but probably also in normal persons (e.g. school children learning orthography). With regard to the effects of errorless learning in amnesic patients, Baddeley and Wil-

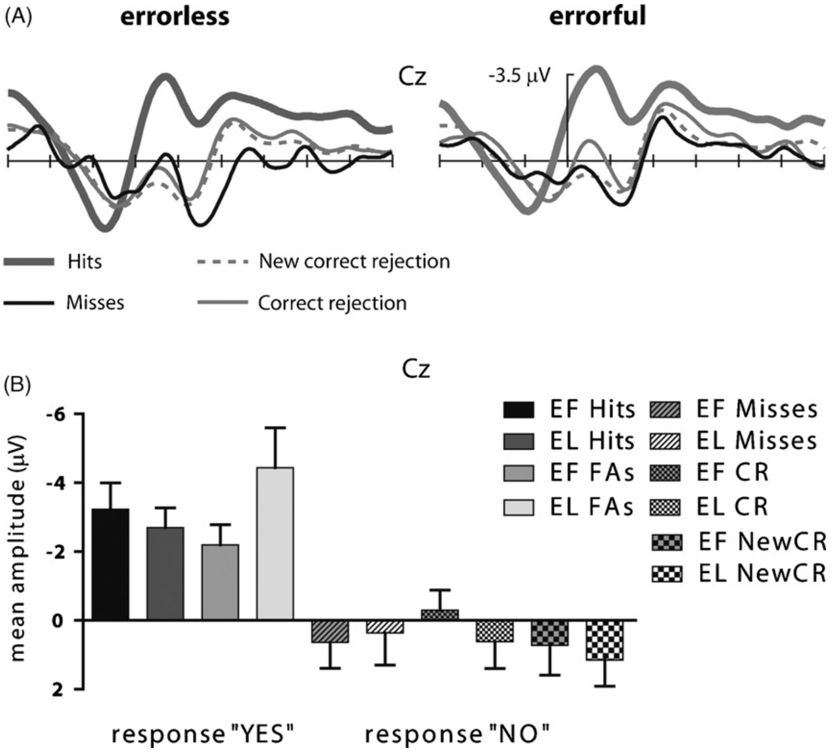

Fig. 2. (A) Response-locked ERPs at Cz, filtered with bandpass 1–8 Hz, baseline 300 ms, all "NO" responses contrasted against correct "YES" responses to target words (hits). (B) Mean amplitudes at electrode site Cz (0–100 ms after response).

son [2] suggested that the EF condition might lead to more memory errors because the trial-and-error learning procedure enhances the activation level of incorrect guesses. This in turn will lead to conflict at the time of retrieval, as both (correct) target words and (incorrect) distracter words compete. By contrast, the EL condition is thought to prevent such competition, because only the correct response gets activated during learning. One potential argument against the advantage of errorless learning in the present study could be the sole use of a sentence production task in the EL condition. However in the Rodriguez-Fornells et al. [19] study the EL conditions' hit rate was relatively small indicating a disadvantage for errorless learned items. In fact, in this previous study, we argued that the errorful task required an activity by the participant (self-generation of possible candidate words) whereas the errorless condition did not, leading to a level of processing effect. To equate hit rates between errorful and errorless conditions we therefore decided to introduce an active task (i.e. sentence generation) for the errorless condition. Thus, the additional task is not meant to yield an advantage for errorless learned items, but rather intends to reduce a disadvantage with regard to the depth of word processing.

Similar to the previous study [19], we observed ERN components of different amplitude to EF–hits, EF–FA, EL–hits and EL–FA, but not to miss trials, or correct rejections of either previously presented words or new words. This is compatible with the idea that the ERN in this memory paradigm is related to post-retrieval checking processes which are triggered when an item leads to familiarity or recollection experience. As such experiences are not present for misses or correct rejections, such trials are not associated with an ERN. An interesting question is why the monitoring system is triggered only in the case of the existence of a possible memory trace. Further studies are needed in order to understand this issue.

In our previous study we interpreted the difference between the false alarms and hits in the EL condition with regard to the response conflict model of the ERN [3,6]. This model holds that the degree to which a "conflict monitor" is activated depends on the product of the current activations for each of two concurrently available response alternatives. However, in the face of the remarkable and even greater difference [19] in memory performance in the present study we expected to see these reflected by response-locked ERPs. In particular, we expected a smaller ERN amplitude for hits in the EL condition compared with EF. This was not the case, however. Instead nearly identical ERN amplitudes were observed for FAs and hits in both learning conditions. While this effect for EF learned items replicates the findings by Rodriguez-Fornells et al. [19] and could be explained in terms of the conflict monitoring theory, the indistinguishable amplitudes of EL items do not fit within this framework.

In addition to the conflict monitoring model of the ERN, there are two alternative theoretical approaches that can contribute to our understanding of the present results: the error detection approach [8] and its extension, the reinforcement learning model (RLM) of the ERN [11], and the errorlikelihood model (ELM) ([4,5], but see also [17]). The RLM states that an action error can be understood as a negative reinforcement signal, which is processed within the mesencephalic dopaminergic system. These error-related changes in dopaminergic activity enable the organism to adapt ongoing behavior and to avoid errors in the future. Here, we will concentrate on the ELM, as it seems to be able to explain the missing ERN differences for hit responses. The ELM postulates that it is not the perceived conflict or the detection of an error *per se* which is causing an activation of the anterior cingulate gyrus (observable electrophysiologically as an ERN response [4]), but the perceived probability of making an error when responding to a given task [5]. In the present study subjects had to respond with "yes", if the word was recognized as a target word learned in the preceding learning phase and with "no" for new and distracter items. In case of perfect responding the ratio of yes:no responses would be 1:2. In terms of the ELM the chance of making an error for yes responses, then is twice as high as for no responses. According to the ELM one should therefore observe increased ACC activity for all yes responses, regardless of learning condition and of response type (hit or FA). As is illustrated in Fig. 2B, this is borne out by the data quite well: there is a major difference of ERN amplitude between yes and no responses. Accordingly, the ELM has the potential to explain these major overall differences between response types. The smaller but significant differences between the ERN amplitudes to the FA trials in the EL and EF conditions, however, are not easily explained by the ELM. This difference can be

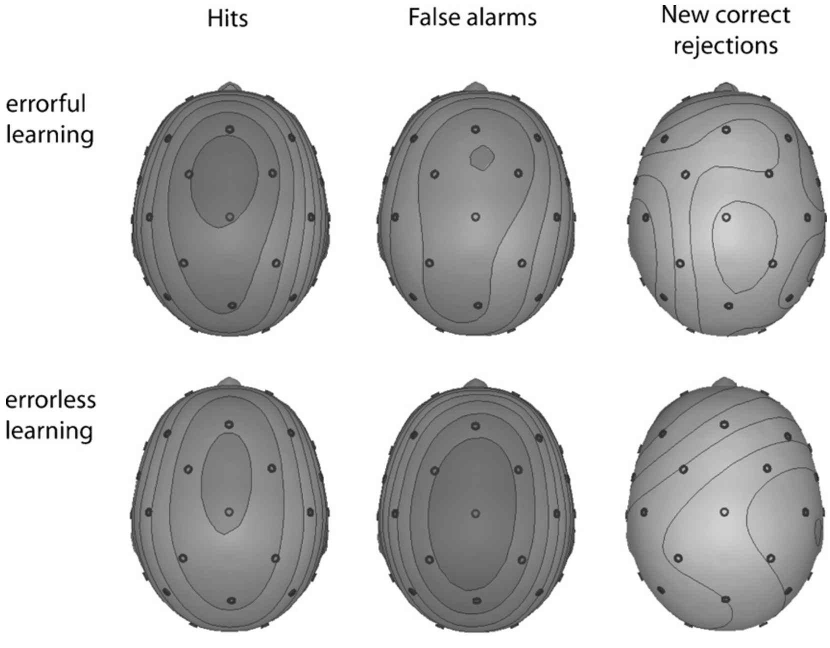

Fig. 3. Spline-interpolated isovoltage maps at 70 ms after response. Each step represents 0.4 μV, darkest color is most negative.

explained by the error detection model (more salient error in case of EL false alarms).

Both peak latency and topographical maximum between Fz and Cz qualify the present ERP response as an instance of the ERN. As evidenced by the isovoltage maps in Fig. 3 the distribution of this ERN-type effect extends towards midparietal scalp in the present study, i.e. more posteriorly than in typical action monitoring experiments [20]. In our previous experiment on errorless learning, we showed a shift from a fronto-central (between 50 and 100 ms) to a more posterior maximum (between 300 and 350 ms, [19, Fig. 7]). A similar topography has also been found by Nessler and Mecklinger [16] in a false memory task. These authors have speculated that the anterior portion of the effect is related to action monitoring/error detection processes, whereas the posterior portion of the effect was attributed to evaluative processes. While the present data are certainly compatible with such a view, further research using a larger sample size is needed to disentangle these two putative portions of the ERN component in memory tasks.

Acknowledgements

Supported by grants from the DFG (MU1311/11-3, KFO 163) to TFM. Also supported by the BMBF (contract 01GO0202). ARF receives support from the Spanish Government (SEJ2005-06067/PSIC and the Ramon y Cajal research program).

References

[1] N.D. Anderson, F.I. Craik, The mnemonic mechanisms of errorless learning, Neuropsychologia 44 (2006) 2806–2813.
[2] A. Baddeley, B.A. Wilson, When implicit learning fails: amnesia and the problem of error elimination, Neuropsychologia 32 (1994) 53–68.
[3] M.M. Botvinick, T.S. Braver, D.M. Barch, C.S. Carter, J.D. Cohen, Conflict monitoring and cognitive control, Psychol Rev 108 (2001) 624–652.
[4] J.W. Brown, T.S. Braver, Learned predictions of error likelihood in the anterior cingulate cortex, Science 307 (2005) 1118–1121.
[5] J.W. Brown, J.R. Reynolds, T.S. Braver, A computational model of fractionated conflict-control mechanisms in task-switching, Cogn Psychol 55 (2007) 37–85.
[6] C.S. Carter, T.S. Braver, D.M. Barch, M.M. Botvinick, D. Noll, J.D. Cohen, Anterior cingulate cortex, error detection, and the online monitoring of performance, Science 280 (1998) 747–749.
[7] S. Dehaene, M.I. Posner, D.M. Tucker, Localization of a neural system for error detection and compensation, Psychol Sci 5 (1994) 303–305.
[8] M. Falkenstein, J. Hohnsbein, J. Hoormann, L. Blanke, Effects of cross-modal divided attention on late ERP components. II. Error processing in choice reaction tasks, Electroencephalogr Clin Neurophysiol 78 (1991) 447–455.
[9] D.A. Gallo, False Memory Research in DRM and Related Tasks, Psychology Press, New York, 2006.
[10] W.J. Gehring, B. Goss, M.G. Coles, D.E. Meyer, E. Donchin, A neural system for error detection and compensation, Psychol Sci 4 (1993) 385–390.
[11] C.B. Holroyd, M.G. Coles, The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity, Psychol Rev 109 (2002) 679–709.
[12] R.P. Kessels, S.T. Boekhorst, A. Postma, The contribution of implicit and explicit memory to the effects of errorless learning: a comparison between young and older adults, J Int Neuropsychol Soc 11 (2005) 144–151.
[13] A. Koriat, M. Goldsmith, A. Pansky, Toward a psychology of memory accuracy, Annu Rev Psychol 51 (2000) 481–537.
[14] G. McCarthy, C.C. Wood, Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models, Electroencephalogr Clin Neurophysiol 62 (1985) 203–208.
[15] M. Moscovitch, Confabulation, in: D.L. Schacter (Ed.), Memory Distortion. How Minds, Brains, and Societies Reconstruct the Past, Harvard University Press, Cambridge, 1995, pp. 226–251.
[16] D. Nessler, A. Mecklinger, ERP correlates of true and false recognition after different retention delays: stimulus- and response-related processes, Psychophysiology 40 (2003) 146–159.
[17] S. Nieuwenhuis, T.S. Schweizer, R.B. Mars, M.M. Botvinick, G. Hajcak, Error-likelihood prediction in the medial frontal cortex: a critical evaluation, Cereb Cortex 17 (2007) 1570–1581.
[18] A.L. Pitel, H. Beaunieux, N. Lebaron, F. Joyeux, B. Desgranges, F. Eustache, Two case studies in the application of errorless learning techniques in memory impaired patients with additional executive deficits, Brain Inj 20 (2006) 1099–1110.
[19] A. Rodriguez-Fornells, C. Kofidis, T.F. Munte, An electrophysiological study of errorless learning, Cogn Brain Res 19 (2004) 160–173.
[20] A. Rodriguez-Fornells, A.R. Kurzbuch, T.F. Munte, Time course of error detection and correction in humans: neurophysiological evidence, J Neurosci 22 (2002) 9990–9996.
[21] H.S. Terrace, Discrimination learning with and without “errors”, J Exp Anal Behav 6 (1963) 1–27.
[22] T.P. Urbach, M. Kutas, The intractability of scaling scalp distributions to infer neuroelectric sources, Psychophysiology 39 (2002) 791–808.