Biological Psychology 93 (2013) 377–385

Contents lists available at SciVerse ScienceDirect

Biological Psychology

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

The ERN is the ERN is the ERN? Convergent validity of error-related brain activity across different tasks

Anja Riesel[a,b,*], Anna Weinberg[b], Tanja Endrass[a], Alexandria Meyer[b], Greg Hajcak[b]

[a] *Institut für Psychologie, Humboldt-Universität zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany*
[b] *Department of Psychology, Stony Brook University, New York, USA*

ARTICLE INFO

Article history:
Received 11 April 2012
Accepted 9 April 2013
Available online 19 April 2013

Keywords:
Construct validity
Discriminant validity
Error-related negativity
Error positivity
ERN
Pe

ABSTRACT

Error-processing is increasingly examined using the error-related negativity (ERN) and error positivity (Pe) – event-related potentials (ERPs) that demonstrate trait-like properties and excellent reliability. The current study focuses on construct validity by applying a multitrait–multimethod approach, treating error-related ERPs (i.e., ERN, Pe and the difference between error minus correct, referred to as ΔERN and ΔPe, respectively) as traits measured across multiple tasks (i.e., Flanker, Stroop, and Go/NoGo). Results suggest convergent validity of these ERPs ranging between .62 and .64 for ΔERN. Values were somewhat smaller for ERN (range .33–.43), Pe (range .37–.49) and ΔPe (range .30–.37). Further, the correlations for ERN and Pe are higher within components across tasks than between different components suggesting discriminant validity. In conclusion, the present study revealed evidence for convergent and discriminant validity of error-related ERPs, further supporting the use of these components as psychophysiological trait markers.

1. Introduction

Adaptive behavior in a changing world requires a flexible system that monitors performance and detects errors. Psychophysiological research on performance monitoring has flourished since error-related event-related potentials (ERPs) were discovered 20 years ago (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993). Studies of performance monitoring have focused in particular on the error-related negativity (ERN; Gehring et al., 1993) or error negativity (Ne; Falkenstein et al., 1991), a sharp negative deflection that appears shortly after the commission of an error over frontocentral electrodes. The anterior cingulate cortex (ACC) has been suggested as the primary generator of the ERN based on studies using both functional neuroimaging (Debener et al., 2005; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004) and source localization techniques (Dehaene, Posner, & Tucker, 1994).

In addition to the ERN, other response-related ERPs are used to examine action monitoring. Several studies report a smaller, but similar looking, negative-going component following correct responses, called the correct response negativity (CRN; Ford, 1999; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000). The ERN is typically followed by the error positivity (Pe; Falkenstein et al., 1991). The Pe has a centroparietal distribution and occurs within 200–500 ms after incorrect responses.

Despite a considerable body of research, the functional significance of these electrophysiological measures of action monitoring is still debated. The Pe has been related to error awareness (Endrass, Reuter, & Kathmann, 2007; Hughes & Yeung, 2011; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001) and some suggest it represents a P3-like response to infrequent error commission (Arbel & Donchin, 2009, 2011; Overbeek, Nieuwenhuis, & Ridderinkhof, 2005; Ridderinkhof, Ramautar, & Wijnen, 2009).

The ERN on the other hand is assumed to signal the need to adjust behavior and to increase cognitive control to improve future performance (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Falkenstein et al., 1991; Gehring et al., 1993; Holroyd & Coles, 2002). More specifically, the ERN has been thought to reflect a neural correlate of conflict monitoring (Botvinick et al., 2001; Yeung, Botvinick, & Cohen, 2004), reinforcement learning (Holroyd & Coles, 2002) or error-likelihood (Brown & Braver, 2005). In addition, the ERN is related to motivational and individual difference variables and is thought to be a trait marker that reflects individual differences in the subjective value of errors based on context, personality, and learning history (Hajcak, 2012; Olvet & Hajcak, 2008; Weinberg, Riesel, & Hajcak, 2012).

Despite this increasing interest in the ERN as a potential trait-like biomarker, the psychometric properties of the ERN have not been thoroughly investigated. There is evidence that the internal consistency and temporal stability of error-related ERPs (ERN, CRN,

---

[*] Corresponding author at: Humboldt-Universität zu Berlin, Institut für Psychologie/Klinische Psychologie, Rudower Chaussee 18, 12489 Berlin, Germany. Tel.: +49 30 2093 4918; fax: +49 30 2093 4859.
E-mail address: anja.riesel@hu-berlin.de (A. Riesel).

0301-0511/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.biopsycho.2013.04.007

Pe) as well as the difference between the ERN and CRN (referred to as ΔERN) are excellent (Olvet & Hajcak, 2009b; Segalowitz et al., 2010; Weinberg & Hajcak, 2011). Specifically, estimates of test–retest reliability range between .40 and .82 over a period of 2–6 weeks (Olvet & Hajcak, 2009b; Segalowitz et al., 2010) and are similar in size (ranging from .56 to .67) after as long as 2 years (Weinberg & Hajcak, 2011).

However, few studies investigated the degree to which the ERN is comparable *across tasks* within the same individuals. This is particularly important given that the ERN is measured in a variety of tasks, including the Flanker (e.g., Gehring et al., 1993), Go/NoGo (e.g., Bates, Kiehl, Laurens, & Liddle, 2002) and Stroop tasks (e.g., Hajcak & Simons, 2002). Initial evidence suggests that the ERN elicited in a Flanker and a Go/NoGo task is highly correlated within the same individuals (Segalowitz et al., 2010). Further, the ERN is often scored at FCz where it is maximal (e.g., Bates et al., 2002; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Gentsch, Ullsperger, & Ullsperger, 2009; Luu, Tucker, Derryberry, Reed, & Poulsen, 2003; Ridderinkhof et al., 2002), but it is unknown where the ERN has the highest convergence across tasks.

Nonetheless, there is some evidence for task-dependent modulations of the ERN (Grundler, Cavanagh, Figueroa, Frank, & Allen, 2009; Mathews, Perez, Delucchi, & Mathalon, 2012; Nieuwenhuis, Nielen, Mol, Hajcak, & Veltman, 2005; Olvet & Hajcak, 2009a). For example, research suggests that enhanced ERN amplitudes in obsessive-compulsive disorder may only be found in response-conflict tasks and not in tasks with probabilistic stimulus–response mappings (Grundler et al., 2009; Mathews et al., 2012; Nieuwenhuis et al., 2005). It has even been proposed that different, albeit overlapping, neural systems can underlie the ERN, depending on the specific task (i.e., execution of an incorrect motor response vs. suboptimal choice; Cavanagh, Grundler, Frank, & Allen, 2010; Grundler et al., 2009). In short, it is not yet clear whether the ERN measured across common tasks reflects a unitary phenomenon or has task-dependent characteristics. Thus, there is a need to directly compare indices of error-related brain activity across different tasks to examine the construct validity of the ERN.

If correlations between ERNs measured across tasks were low, this would suggest that the ERN is not a singular entity, and that correlations with individual difference measures may depend heavily on the task used to elicit the ERN. In this case, it would be useful to specify results in terms of the task employed (e.g. Stroop-ERN). Given that the ERN is discussed as a promising biomarker, which could be useful for diagnostic or prognostic purposes, this question is of central importance in guiding such research efforts.

To this end, error-related ERPs (i.e., ERN, Pe) were assessed in the current study using three commonly employed speeded response tasks (i.e., Flanker, Stroop, Go/NoGo). An adjusted multitrait–multimethod matrix (MMTM; Campbell & Fiske, 1959) was applied to examine whether indices of error-related brain activity (i.e., traits) converge across different tasks (i.e., methods). This was done in order to examine how comparable (convergent validity) and distinguishable (discriminant validity) error-related components are across tasks.

2. Methods

2.1. Participants

Forty-seven undergraduate students (20 female) from Stony Brook University participated in this study. Two participants committed fewer than six errors and were therefore excluded from further analysis, since evidence suggests that between 6 and 8 error trials are needed to reliably quantify the ERN and Pe (Olvet & Hajcak, 2009c; Pontifex et al., 2010). Data from two subjects were excluded due to excessive EEG artifacts. The final sample consisted of 43 participants (19 female). All participants had normal or corrected-to-normal vision and reported no history of head trauma or neurological disease. The mean age was 19.14 years (*SD* = 1.42). 38.6% of the sample was Caucasian/European, 45.5% was Asian-American, 6.8% was Hispanic, 2.3% was African-American and 6.8% identified as "other." All participants received verbal and written information about the aims and procedure of the study and written consent was obtained. All participants received course credit for their participation.

2.2. Task and procedure

The experiment consisted of three tasks: a modified Flanker task, a Go/NoGo task, and a Stroop task. The order of tasks was counterbalanced across participants. All tasks were administered using Presentation software (Neurobehavioral Systems, Inc., Albany, CA, USA). Prior to each task, the participants performed a practice block containing 20 trials. All three tasks consisted of 420 trials presented in 7 blocks of 60 trials. All stimuli were presented for 200 ms. An intertrial interval (ITI) that varied randomly from 600 to 1000 ms followed the response. Throughout all tasks, participants were encouraged to be both fast and accurate in their performance. To encourage both fast and accurate behavior, performance-based feedback was presented at the end of each block. If performance accuracy was below 75%, a message appeared instructing participants to respond more accurately. When performance was above 90%, participants were instructed to respond faster. Error rates between 10 and 25% were followed by the feedback "You're doing a great job." For the Go/NoGo task, the performance feedback was given with regard to error rates that were calculated for NoGo trials only (i.e., errors of commission). The total duration of the three tasks combined was approximately 60 min.

Flanker task: On each trial of the Flanker task (Eriksen & Eriksen, 1974; Kopp, Rist, & Mattler, 1996), five horizontally aligned white arrowheads were presented and participants were instructed to respond with the left or right mouse button in accordance with the direction of the central arrowhead. Half the trials were compatible (e.g. flanker arrows and target point in the same direction) and half were incompatible (e.g. flanker arrows and target point in opposite directions). The trials were displayed in a pseudorandomized order. At a viewing distance of approximately 65 cm, the set of arrows filled 2° of visual angle vertically and 10° horizontally.

Stroop task: On each trial, one of three color words ('red', 'green', 'blue') was shown, and was presented in either red or green font. Subjects were instructed to press the left mouse button if the color word was presented in red, and press right button if the color word was presented in green. Thus, 1/3 of trials were compatible (e.g. color word and font color require the same response, 'red' in red font, 'green' in green font), 1/3 were incompatible (e.g. color word and font color require different responses, 'red' in green font, 'green' in red font), and 1/3 were neutral (e.g. the color word, 'blue' in red or green font). At a viewing distance of approximately 65 cm, each word occupied between 2° and 3° of visual angle.

Go/NoGo task: In the Go/NoGo task, a green triangle was presented on each trial. Participants were instructed to press the right mouse button in response to an upright triangle, which occurred on 80% of the trials. Additionally, participants were told to withhold responses to slightly tilted triangles (10°), which occurred on 20% of the trials. At a viewing distance of approximately 65 cm, each triangle occupied $3^\circ \times 3^\circ$ of the visual angle.

2.3. Psychophysiological recording, data reduction and analysis

The continuous EEG was recorded using an elastic cap and the ActiveTwo BioSemi system (BioSemi, Amsterdam, Netherlands). Sixty-four electrode sites were used, based on the 10/20 system, as well as two electrodes on the right and left mastoids. All electrodes were sintered Ag–AgCl electrodes. The Electrooculogram (EOG) was recorded using four additional facial electrodes: two electrodes placed approximately 1 cm outside of the right and left eyes and two electrodes mounted approximately 1 cm above and below the right eye. To improve the signal-to-noise ratio, the EEG signal was pre-amplified at the electrode with a gain of 1× by a BioSemi ActiveTwo system (BioSemi, Amsterdam, Netherlands). The EEG was digitized with a sampling rate of 512 Hz using a low-pass fifth order sinc filter with a half-power cutoff of 102.4 Hz. A common mode sense (CMS) active electrode producing a monopolar (non-differential) channel was used as recording reference. The EEG was analyzed using Brain Vision Analyzer (Brain Products, Gilching, Germany).

2.4. ERP analysis

Offline, the data were referenced to the average of the left and right mastoids, and band-pass filtered with low and high cutoffs of 0.1 and 30 Hz, respectively. Eye movement artifacts were corrected using the algorithm developed by Gratton, Coles, and Donchin (1983). Response-locked epochs with a duration of 1200 ms, including a 400 ms prestimulus interval, were extracted. A semi-automatic procedure was used to detect and reject artifacts. Epochs containing a voltage step of more than 50 μV between sample points, a voltage difference of 300 μV within a segment, and a maximum voltage difference of less than 0.50 μV within 100 ms intervals were rejected. In addition, visual inspection of the data was conducted to detect and reject any remaining artifacts. Response-locked ERPs were averaged separately for each participant, each task, and for incorrect and correct responses. For all tasks, trials with response times below 100 ms and above 700 ms were excluded from averaging. Because the ERN can begin prior to the completion of the motor response, we used the 400–200 ms pre-response interval as the baseline in order to avoid subtracting out activity of interest (Weinberg, Olvet, & Hajcak, 2010). To quantify the ERN and