

Are all errors created equal?

Sven Lesche¹

¹ Ruprecht-Karls-University Heidelberg

Author Note

This is a research proposal completed for the seminar ‘FOV Elektrophysiologie’. It was completed using R-markdown, all code needed to replicate this work can be found at <https://github.com/SLesche/eeg-seminar>.

Correspondence concerning this article should be addressed to Sven Lesche, Im Neuenheimer Feld 695. E-mail: sven.lesche@stud.uni-heidelberg.de

Are all errors created equal?

Introduction

Post-error slowing (PES) describes the increase in response time (RT) on trials following an error compared to trials following a correct response. PES has been observed in a variety of experimental settings and is often accompanied by a post-error change in accuracy (PEA). Some studies observed decreased accuracy following errors while others find accuracy improvements. This has lead researchers to investigate differences in cognitive processes underlying post-error effects.

Damaso et al. (2020) have argued that different types of errors can lead to different types of post-error changes in behavior. They differentiate between *response speed* and *evidence quality* errors.

Damaso et al. (2020) used median splits in RT to show that behavioral differences within-task exist between the two types. They observed increased PES for the faster bin of errors and post-error speeding for the bin containing mainly *evidence quality* errors. Only measuring RT and accuracy changes provides a limited view of post-error effects and fails to capture differences in cognitive processes following the different types of errors. This RT approach is also highly susceptible to regression to the mean. We will aim to improve on the research design of Damaso et al. (2020) by adding two tasks each eliciting either *response speed* or *evidence quality* errors and using neurocognitive measures to measure post-error changes in cognitive processes.

We will deploy three tasks designed to elicit either mainly *response speed* errors, *evidence quality* errors or both. A modified Flanker task will be used to elicit both error types. A modified version of the Behavioral Adaptation Task (Hester et al., 2007) will be used to elicit mainly *response speed* errors and a difficult color discrimination task (Buzzell et al., 2017) will be used to elicit mainly *evidence quality* errors.

Differences in post-error adjustments to cognitive processes will be studied in several measurements obtained in an electroencephalogram (EEG). This includes error-related

components in event-related potentials (ERP), the *error-related negativity* (ERN) and *error-positivity* (Pe). ERN is a fronto-central negative deflection in the response-locked ERP that is more negative following error responses. It peaks around 100 ms after the response and is often associated with error monitoring or response conflict (Botvinick et al., 2001). Pe on the other hand is a centro-parietal positive deflection around 200 - 400 ms following the response. The functional interpretation of this component is less clear. It seems to be a later stage of error processing relevant for error awareness (Di Gregorio et al., 2018).

Of interest in this study are differences in post-error adjustments between conditions containing either mainly *response speed* or *evidence quality* errors. We will obtain EEG measures and investigate post-error changes using t-tests. We predict post-error changes in behavior of all types to be more pronounced following *response speed* errors. Errors due to poor evidence quality should not induce large changes in cognitive processes, as the underlying cognitive processes were not at fault when committing the error.

Hypotheses

H1: Post-error slowing is more pronounced in conditions containing mainly *response speed* errors.

H2: ERN is more negative in conditions containing mainly *response speed* errors.

H3: Pe is more positive in conditions containing mainly *response speed* errors.

Design

Participants

To detect a medium sized difference ($d > 0.50$) between post-error behavior depending on error type (evidence quality vs. response speed) in a t-test for dependent measures with $\alpha = 0.05$ and $\beta = 0.80$ we require a minimum of 27 participants (Faul et al., 2009). To account for potential removal of participants due to bad data quality, a minimum of 30 participants should be collected.

Materials

Flanker Task. Participants have to decide whether the centrally presented letter is a “U” or a “V”. This central letter is surrounded by flanking letters that either show the same letter (congruent) or the respective other letter (incongruent). This task elicits both *evidence quality* and *response speed* errors. We will determine the type of error via the RT of the error in a manner similar to Damaso et al. (2020). Participants will complete 5 blocks of 100 trials each.

Behavioral Adaptation Task. The stimuli “X” and “Y” are presented in a mostly alternating fashion. When the current stimulus is different from the previously presented stimulus, participants have to press the button corresponding to the stimulus. When the stimulus is repeated, participants should give no response at all. This leads to a large number of *response speed* errors. Participants will complete 6 blocks of 250 trials each.

Color Discrimination Task. A stimulus consistent of a smaller circle surrounded by a bigger circle is presented each trial. The participant has to decide whether the inner circle is of the same color as the outer ring. The true difference in colors in incongruent stimuli will be adapted to each participant to ensure error rates of 20%. Errors in this task are mainly due to participants being unsure of whether there is a true difference and thus compromise largely *evidence quality* errors. Participants will complete 8 blocks of 80 trials each.

Procedure

We will conduct all tasks on one measurement occasion. We will counterbalance the order of the three tasks across participants. All tasks are two-choice tasks, appropriate button responses will be counterbalanced across participants. We will set the response-stimulus interval (RSI) to randomly vary between 800 ms and 1000 ms on each trial of all tasks. This aims to reduce anticipation effects and obtain clearer ERPs. All stimuli will be presented for 250ms. All trials have a response deadline of 2000ms. A fixation cross will be presented during the RSI. No feedback will be given outside of practice sessions.

EEG

As we will not conduct any spatial analysis and are mainly looking at ERPs, the quality and placement system of the electrodes does not need to be ideal. An example setup is given here, a lower number of electrodes or a different cap-style will not diminish the results obtained.

EEG will be recorded using 64 in-cap Ag/AgCl electrodes placed in the extended 10-20 system. Electrooculogram (EOG) measures will be taken bipolarly by two electrode placed above and below the left eye to correct for ocular artifacts. All impedances will be kept below 5 $k\Omega$. EEG signal will be recorded with a sampling rate of 1000 Hz (band-pass 0.1 Hz - 100 Hz) and referenced to an in-cap reference located between electrodes Cz and CPz and an in-cap ground just anterior to Fz.

Following data acquisition, we will apply a low-pass filter of 16 Hz and a high-pass filter of 0.1 Hz. The raw data will be down-sampled to 250 Hz. To remove artifacts we will conduct an ICA on the dataset down-sampled to 100 Hz and passed through an additional high-pass filter of 1 Hz. Raw data will be cleaned by removing channels with unusually long flatlines, artifact-rates or line-noise. Channels removed will be interpolated following this procedure. EEG-data will be re-referenced to the average across electrodes and epoched to 200ms prior to and 1000ms following all stimulus and response markers.

Peak-amplitude of the response-locked ERN will be determined at the electrode Cz between 20 - 150 ms following the response. Response-locked peak amplitude of Pe will be measured at the Pz between 200 - 500 ms following the response.

Analysis

The effect of error type on behavioral measures, ERN and Pe will be compared within the Flanker task and between the BAT and color discrimination task. We will define the trials surrounding the 25% fastest errors as belonging to the *response speed* category and the trials surrounding the 75% slowest errors as belonging to the *evidence quality* category in the flanker task. In the BAT, the 85% fastest errors will be termed *response speed*. The opposite

is true for the color discrimination task, here the 85% slowest errors will be termed *evidence quality*. We will then obtain behavioral and electrophysiological measures for those categories and use t-tests to investigate differences in dependent measures between the two error types.

Limitations

This study aims to falsify the concept of a unitary PES as a dependent measure often used in clinical research. We want to show that post-error adjustments differ depending on the type of error committed. We have identified several limitations.

Firstly, the comparison across tasks does not allow differences in post-error behavioral and neurocognitive to be mapped onto differences in error types. The inclusion of the Flanker task aims to combat this. Nonetheless, the overarching aim of this study is to falsify the assumption of a unitary PES-mechanism. Differences across tasks in neurocognitive behavior following errors also achieve this goal.

A possible limitation is the definition of errors used here. The fastest 25% of errors in the flanker task may not exclusively be *response speed* errors. In order to combat this, we introduced two tasks that elicit mainly one of the error types. We will also pre-test the tasks and question participants regarding the reasons for error commission to ensure that our assumptions about the types of errors committed reflect reality.

A further limitation haunting PES research is the number of trials. Even with a minimum of 500 trials in each task, some participants may commit fewer than 10 errors. Higher trial counts and extensive pre-testing will be required to find an appropriate amount of trials per task. This is especially relevant when aiming to investigate ERP components on an individual level. Luck (2005) recommends a minimum of 30 trials even for large components. This may prove to be an issue here.

Despite these limitations, our study will allow us to investigate the presence and effect of two different error types. This will further add to recent findings illuminating the diverse nature of post-error adjustments in behavior and warn future researchers not to consider all post-error adjustments as equal.

References

- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624.
- Buzzell, G. A., Beatty, P. J., Paquette, N. A., Roberts, D. M., & McDonald, C. G. (2017). Error-induced blindness: Error detection leads to impaired sensory processing and lower accuracy at short response–stimulus intervals. *Journal of Neuroscience*, *37*(11), 2895–2903.
- Damaso, K., Williams, P., & Heathcote, A. (2020). Evidence for different types of errors being associated with different types of post-error changes. *Psychonomic Bulletin & Review*, *27*, 435–440.
- Di Gregorio, F., Maier, M. E., & Steinhauser, M. (2018). Errors can elicit an error positivity in the absence of an error negativity: Evidence for independent systems of human error monitoring. *Neuroimage*, *172*, 427–436.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using g* power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160.
- Hester, R., Simoes-Franklin, C., & Garavan, H. (2007). Post-error behavior in active cocaine users: Poor awareness of errors in the presence of intact performance adjustments. *Neuropsychopharmacology*, *32*(9), 1974–1984.
- Luck, S. J. (2005). Ten simple rules for designing and interpreting ERP experiments. *Event-Related Potentials: A Methods Handbook*, *4*.