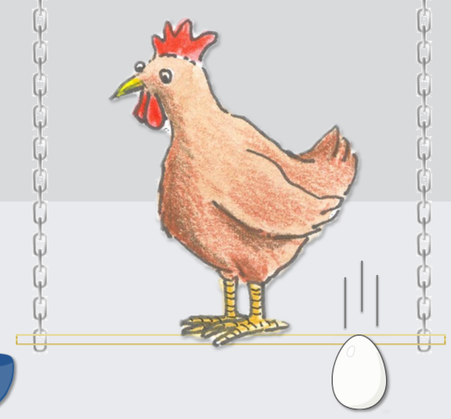
Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

Of chickens, eggs, and yolk: The electrophysiology of breaking a rule

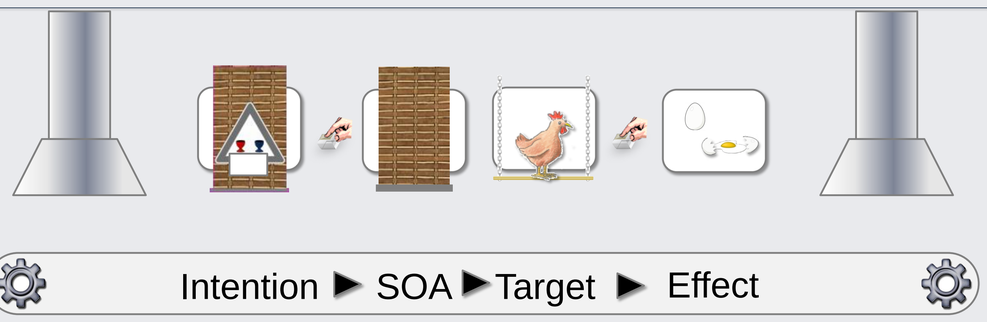


Roland Pfister & Wilfried Kunde

BACKGROUND

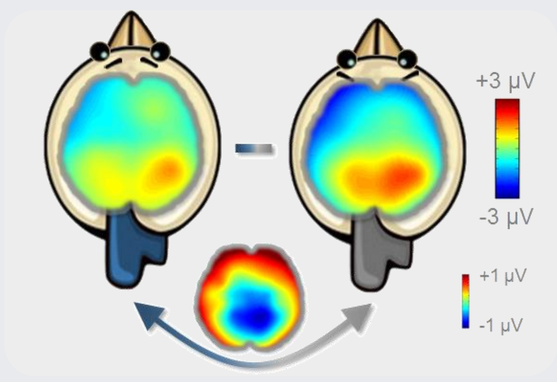
Not all rules can be obeyed at all times, and failures to obey a rule can be either unintended or intended (Reason, 1990). Interestingly, whereas the electrophysiology of unintended errors has been studied extensively in the last decades, virtually nothing is known about the electrophysiological signature of intended rule violations. The present study targeted two ERP components to address this question: The error-related negativity (ERN / N_E) and the P300.

Because the **ERN** is assumed to reflect monitoring of unexpected events (e.g., Holroyd & Coles, 2002) we did not expect any ERN-like waveforms for intended rule violations. In contrast, we expected violation-specific effects on the **P300** component, which has been linked to the (direct) mapping from stimuli to canonical responses (Verleger et al., 2005). Because rule violations are the opposite of such canonical responses, we expected attenuated and/or delayed P300 responses for rule violations as compared to rule-based behaviour.



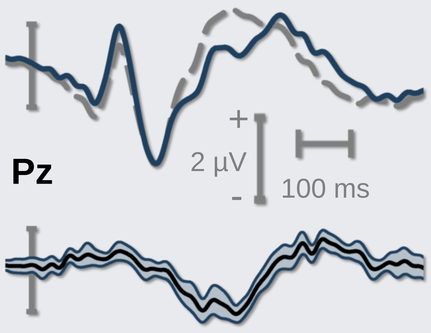
The task was framed as a game in which participants controlled an 'egg factory' (N = 16). In each trial, to ensure smooth operation of the factory, they had to place an egg cup under the rear of a left- or right-looking chicken by pressing a left or right response key.

Crucially, at the beginning of each trial, participants indicated whether they wanted to conform to factory rules (and **perform correctly**) or whether they wanted to **violate the factory rules** and commit an error by intention.



The **stimulus-locked analysis** of the ERPs elicited by the target stimulus is shown to the left, and a clearly attenuated and delayed **P300** response emerged for rule violations as compared to normal, rule-based responses (time to peak amplitude at Pz: 448 ms vs. 404 ms; p = .008, d = 0.77).

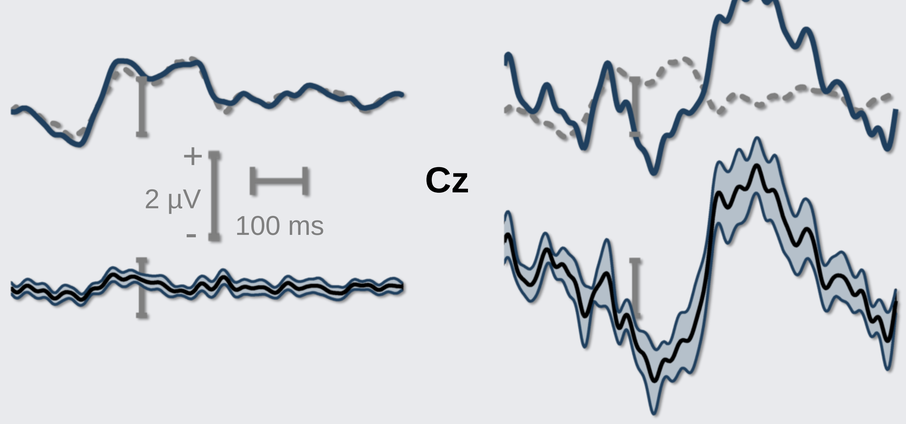
Upper left: Mean voltage distributions, 300-450 ms post-stimulus. The blue mock head represents rule violations, the grey mock head represents normal, rule-based responses. The plot in between both heads shows the difference between the conditions.



Lower left: Exemplar ERP data for the candidate electrode Pz (upper plot) and the difference wave (lower plot). The blue line of the ERP represents rule violations, the dashed, grey line represents normal, rule-based responses. The coloured area around the difference wave indicates ± 1 standard error of paired differences (Pfister & Janczyk, 2013), computed separately for each time point.

RESULTS

The **response-locked analysis** of the ERPs did not show any specific effects for the comparison of rule violations to correct responses, confirming the prediction of an absent ERN in this case (**lower left plot**; the blue line represents rule violations). By contrast, a clear ERN was present for the comparison of unintended errors and correct responses (**lower right plot**; the blue line represents errors). The coloured areas indicate ± 1 standard error of paired differences (Pfister & Janczyk, 2013), computed separately for each time point.



SUMMARY & CONCLUSIONS

The present results are clear-cut: First, intended rule violations do not give rise to ERN-like components in the ERP as unintended errors do. Second, the P300 response to the stimulus prompting a violation was attenuated and delayed compared to normal, rule-based responses.

These findings indicate that rule violations are a complex type of behaviour, distinct from normal, rule-based responding. Human agents thus seem to be unable to simply reverse a to-be-violated rule; rather they continue to be influenced by the original rule. Rule breaking therefore reminds of **ironic effects** of facilitating behaviour that an agent intends to suppress (Wegner, 2009). The relation of rule violations to such ironic effects on the one hand and the exact processes that underlie the observed results on the other hand clearly await further investigation.

REFERENCES

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679-709.

Pfister, R., & Janczyk, M. (2013). Confidence intervals for two sample means: Calculation, interpretation, and a few simple rules. *Advances in Cognitive Psychology*, *9*(2), 74-80.

Reason J. (1990). *Human error*. New York: Cambridge University Press.

Verleger, R., Jaśkowski, P., & Wascher, E. (2005). Evidence for an integrative role of P3b in linking reaction to perception. *Journal of Psychophysiology*, *19*(3), 165-181.

Wegner, D. M. (2009). How to think, say, or do precisely the worst thing for any occasion. *Science*, *325*, 48-50.

Contact: roland.pfister@uni-wuerzburg.de