Department of Clinical Neuroscience

Cognitive Control in Reinforcement Learning - Brain Structure and Function

AKADEMISK AVHANDLING

som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i Hillarp, Retzius väg 8.

Fredagen den 15 December, 2017, kl 09.00

av

Björn C. Schiffler

Huvudhandledare:
Dr. Sara L. Bengtsson
Karolinska Institutet
Department of Clinical Neuroscience

Bihandledare:
Prof. Martin Ingvar
Karolinska Institutet
Department of Clinical Neuroscience

Prof. Peter Fransson
Karolinska Institutet
Department of Clinical Neuroscience

Fakultetsopponent:
Dr. Claudia Danielmeier
University of Nottingham
School of Psychology

Betygsnämnd:
Prof. Martin Lövdén
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society

Prof. Peter Juslin
Uppsala Universitet
Department of Psychology

Dr. Marc Guitart-Masip
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society

Stockholm 2017

Abstract

Adapting behaviour according to internal or external feedback is a fundamental property of cognitive control. For example, humans tend to slow down when they make mistakes, a process called post-error slowing (PES), which has previously received extensive attention in research on response inhibition. However, whether PES is actually an adaptive process which helps avoid future mistakes or a maladaptive one which siphons cognitive resources is still not clear.

The overall aim of the work in this thesis was to investigate how post-error slowing contributes to the stabilization of performance after errors and which brain areas are involved in this response inhibition process. We did this by combining behavioural experiments, computational modelling and neuroimaging techniques to provide a comprehensive analysis of latent decision processes and their neural correlates.

Specifically, in **Study I**, we analyzed data from a probabilistic reinforcement learning task in combination with functional Magnetic Resonance Imaging to explore which brain regions signalled enhanced future post-error slowing when receiving negative feedback. On a behavioural level, we studied whether PES was associated with how well participants learned, as assessed in a later test phase. We showed that post-error slowing was associated with brain activity in a central cognitive control region, the right inferior frontal gyrus (rIFG) as well as brain regions in occipital cortex which overlapped with the representation of absolute prediction errors, a measure reflecting deviance from expectations, i.e., surprise at feedback.

In **Study II**, we found that cortical thickness in rIFG as a measure of grey matter integrity was related to inter-individual differences in post-error slowing, both for direct next trials and trials further apart in time. This analysis was supported by a drift diffusion model of the underlying decision processes, which demonstrated that an increased decision boundary after an error, indicating enhanced response caution, was related to cortical thickness variability, particularly in anterior parts of the rIFG.

Finally, in **Study III** we used drift diffusion modelling on a large-scale behavioural dataset during a visual search task to illuminate decision processes of up to five trials after an error and how post-error adaptation benefits accuracy recovery several trials after the error. Post-error slowing was marked by both adaptive and non-adaptive decision processes which changed dynamically over several trials after an error. While adaptive increases in decision threshold were sustained for several trials after an error, reductions in evidence accumulation only transiently affected performance on the next trial after the error. Further, post-error increases in response caution and evidence accumulation were also associated with better performance on future trials.

These studies illustrate that there is valuable information to be gained about response inhibition processes beyond looking at the simple relation of post-error slowing and accuracy. Computational modelling allowed us to compartmentalize various decision processes and relate adaptations in these processes directly to brain anatomy. We hope the results from the studies presented in this thesis can provide a framework for future work on how the brain learns from mistakes and adapts to a continually changing environment.