Computational mechanisms of stimulus-stimulus and stimulus-reward associative learning

Sandra Iglesias[1,*] · Christoph Mathys[1,3] · Kay H. Brodersen[1,2] · Lars Kasper[1,3] · Marco Piccirelli[1] · Klaas E. Stephan[1,4]

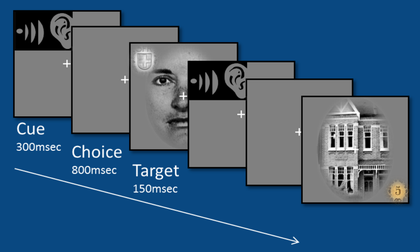
[1] Laboratory for Social and Neural Systems Research, Institute for Empirical Research in Economics, University of Zurich, Switzerland [2] Department of Computer Science, ETH Zurich, Switzerland [3] Institute for Biomedical Engineering, ETH Zurich
[4] Wellcome Trust Centre for Neuroimaging, University College London, UK
[*] Correspondence: sandra.iglesias@econ.uzh.ch

1 Summary

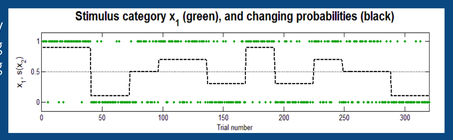
- Synaptic plasticity is critical for reconfiguration of neuronal circuits during normal learning and development [1] but also for pathological learning and disorders. Although numerous studies in various species have investigated different forms of learning, we still lack a precise understanding of the differences in physiological and computational mechanisms involved and particularly of the relative importance of different modulatory transmitters.
- Learning is driven by prediction errors. Dopamine has been found to play a crucial rule in prediction errors related to reward[2], whereas ACh might play a crucial role for synaptic plasticity during perceptual and reward learning.
- Using computational modeling and two different learning tasks, a stimulus-stimulus association (SSA) learning task and a stimulus-reward association (SRA) learning task, we conducted a behavioral and an fMRI-study to investigate relevant neural circuits and potential mechanisms of neuromodulation.
- Here we present preliminary behavioral and fMRI results. The main activation for precision-weighted prediction errors was found to be represented in prefrontal areas. Prediction errors were mainly represented in/ near the basal forebrain.
- Further analyses will include models of effective connectivity and genetic analyses to characterize putative dopaminergic and cholinergic mechanisms involved in these learning forms.

2 Experimental paradigm

Participants (behavioral study: 47; fMRI study: 34) engaged in both studies in two learning tasks; in a stimulus-stimulus association (SSA) and in a stimulus-reward association (SRA) task. Both tasks required continuous learning of probabilistic, time-varying associations between cues and targets. The perceptual input was identical in both tasks, but the task was different. In the SSA learning task, participants had to predict the visual stimulus (face or house), given the auditory stimulus (high or low tone). In the SRA learning task, participants had to predict the monetary value (indicated by superimposed coins, 5 francs or 0.05 francs) that appeared randomly in one of the 4 edges, given the auditory stimulus (high or low tones).



Probabilities were only transiently stable (range: 24 to 40 trials), changing pseudorandomly between strong (p=0.9, 0.1), moderate (p=0.7, 0.3) and no association (p=0.5).

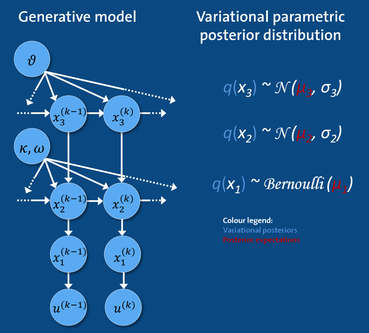


Literature

[1] Gu, Q. (2002). Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity. Neuroscience, 111(4), 815-835.
[2] Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. Behav Brain Funct, 6, 24.
[3] Matthys, C., Daunizeau, J., Friston, K., & Stephan, K.E. A Bayesian foundation for individual learning under uncertainty. Front. Hum. Neurosci. 5:39.
[4] Fletcher, P. C., Anderson, J. M., Shanks, D. R., Honey, R., Carpenter, T. A., Donovan, T., et al. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. [Research Support, Non-U.S. Gov't]. Nature neuroscience, 4(10), 1043-1048.

3 Modelling

Bayesian learning models prescribe an optimal way how agents learn under uncertainty, and they provide trial-wise prediction error estimates that can inform models of synaptic plasticity. However, they are computationally too complex for real-time learning and therefore biologically unrealistic.
We have developed an extremely efficient variational approximation to ideal Bayesian learning; this allows for inference on an agent's belief about causal relations between stimuli in a changing world.



The participant is taken to receive a sequence of inputs u(1), u(2),... (the associated pairs of cues and targets). It uses these to make inferences on a hierarchy of hidden states x1, x2,... of its environment. While x1 is binary, all higher states are continuous. Continuous states change by performing Gaussian random walks that are hierarchically coupled (one state's step size is determined by the next higher state).

In this model with two coupled random walks, the nature of learning is determined by three parameters ϑ, κ, and ω. These parameters can be estimated from behavioral data.

ϑ is the step size of the random walk in x3. Reducing it leads to little learning in x3 due to agent's small uncertainty about x3's true value. ω regulates the step size in x2. Reducing it leads to little learning primarily in x2 and secondarily in x3 since this, representing the log-volatility in x2, cannot change much if x2 remains stable. κ determines how strongly x2 and x3 are coupled. Small κ diminishes learning in x3 despite great uncertainty while learning in x2 remains largely unaffected.

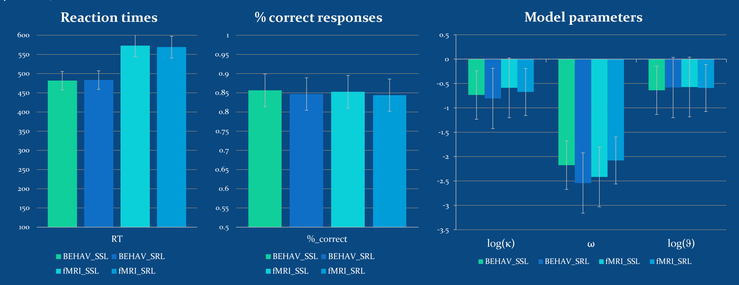
Furthermore, we derive trial-by-trial update equations, which describe the agents update of its beliefs about the environment, and which contain i.a. parameters that, analogous to reinforcement learning models, represent prediction errors and precision-weighted prediction errors[3].

Prediction error $= \mu_1^{(k)} - s\left(\mu_2^{(k-1)}\right)$; $\quad p(\mu_2^{(k)}) \sim N(\mu_2^{(k-1)}, \exp(\kappa x_3 + \omega))$

Precision-weighted prediction error $= \sigma_2^{(k)}\left(\mu_1^{(k)} - s\left(\mu_2^{(k-1)}\right)\right)$

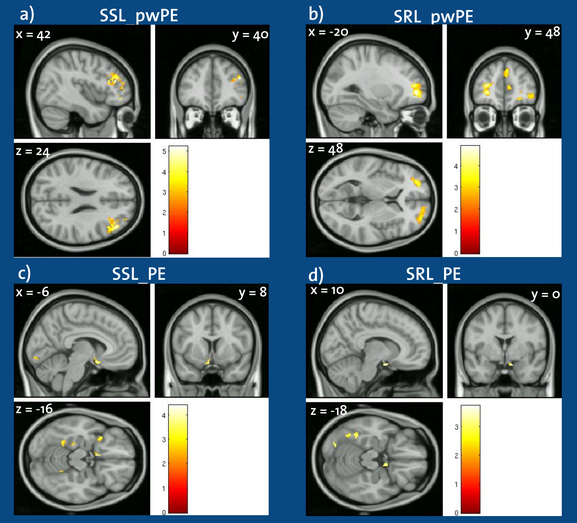
4 Behavioral Results

In both studies, reaction times (RT) and performance (% correct responses) did not show a significant difference between both learning tasks (behavioral: RT: $t(46) = -0.16$, $p>0.05$; %CR: $t(46) = 1.21$, $p>0.05$; fMRI: RT: $t(36) = 0.36$, $p>0.05$; %CR: $t(36) = 1.26$, $p>0.05$). Also, no significant differences between both learning tasks were found for the three model parameters ϑ (behavioral: $t(46) = -0.55$, $p>0.05$; fMRI: $t(36) = 0.21$, $p>0.05$), κ (behavioral: $t(46) = 0.83$, $p>0.05$; fMRI: $t(36) = 0.93$, $p>0.05$), and ω (behavioral: $t(46) = 1.34$, $p>0.05$; fMRI: $t(36) = -1.29$, $p>0.05$). Furthermore, the parameters did not significantly differ, when comparing them between the behavioural and the fMRI study (ϑ: $t(82) = -0.62$, $p>0.05$; κ: $t(82) = 1.36$, $p>0.05$; ω: $t(82) = -1.09$, $p>0.05$).



5 Functional Results

- For the brain activity analyses, we examined the representation of trial-wise precision-weighted prediction errors (pwPE) and of prediction errors (PE). Trial-by-trial estimates of pwPE and PE were generated for each subject by the hierarchical Bayesian learning model.
- For the SSA learning task, the statistical parametric maps of the pwPE showed significant activation (p<0.05 cluster-level corrected across the whole brain, voxel-level threshold $p < 0.01$) in the right medial and inferior prefrontal cortex. (a). For the SRA learning task the significant activation (p<0.05 cluster-level corrected across the whole brain, voxel-level threshold $p < 0.01$) was located in bilateral superior prefrontal cortex, right medial prefrontal cortex and ACC, and in the right orbitofrontal cortex (b).
- Prediction errors in the SSA learning task were found to be represented in/near the basal forebrain (c; uncorrected, $p < 0.01$) and visual areas, such as fusiform gyrus. A similar activation pattern was found for prediction errors during the SRA learning task (d; uncorrected, $p < 0.01$).



6 Discussion

Precision-weighted prediction errors, i.e. the value of prediction errors, were found to be mainly represented in pre- and orbitofrontal areas, both for SSA and SRA learning. Responses in prefrontal cortex have been previously related to unpredictable and surprising stimuli[4], i.e. when uncertainty plays a critical role. In particular, medial prefrontal cortex (MFC) has been suggested to be important in cognitive control and performance adjustments. Therefore, one may speculate that MFC alters processing in neural circuits implementing perception and behavior according to the precision of prediction errors. This will be examined in future analyses using DCM. In both tasks, prediction errors elicited activation near/in the basal forebrain. This region, which contains cholinergic nuclei and is close to dopaminergically innervated regions (i.e., Nucl. accumbens), has previously been implicated in reward processing.[2]

These preliminary results, that we obtained by a Bayesian learning model that predicted the agent's trial-by-trial beliefs about the environment, will be completed by additional subjects (measured, but not yet analyzed), genetic analyses of SNPs related to DA and ACh, and connectivity analyses using DCM. In particular, we aim to quantify the "hidden" neurophysiological mechanisms that affect synaptic plasticity and shape different learning forms. By combining computational models, models of effective connectivity, and genetic information, we hope to better characterize the role of different neurotransmitters, such as dopamine and acetylcholine, for different forms of learning.