Reward predictability and oscillatory brain processes

Alfredo Vellido[1], Wael El-Deredy[1], Thomas Gruber[2],
David H. Zald[3], Francis P. McGlone[4], and Mathias M. Müller[2]

[1]School of Computing and Mathematical Sciences, Liverpool John Moores University, Liverpool L3 3AF, UK
[2]Centre for Cognitive Neuroscience, University of Liverpool, Liverpool L69 7ZA, UK
[3]Department of Psychology, Vanderbilt University, Nashville, TN, USA
[4]Centre for Cognitive Neuroscience, University of Wales, Bangor LL57 2DG, UK

In spite of the increasing neurophysiological, neuroimaging and neurocomputational interest in the mechanisms of processing reward and reward predictability, the dynamics and manifestation of reward predictability in human neuroelectric activity remains unknown. Human and animal studies have shown that predictability modulates the brain response to reward, and that predictability activates the reward pathways. Theoretical models of dopamine release suggest that this neurotransmitter may be acting as a reward predictor, and that uncertainty may be critical for learning, and hence the guiding of future behaviour. Furthermore, it is increasingly being acknowledged that in order for the brain to predict future stimuli rapidly and efficiently, it must make use of top-down resources by building assemblies of synchronised neural activity. In this study we investigate synchronised EEG activity in response to predictable and unpredictable rewards. A card selection task was used to generate predictable or unpredictable monetary rewards (or non-rewarding controls). We report initial results that demonstrate a complex pattern of synchronised EEG gamma band in different brain regions that differentiates between rewarding, non-rewarding, predictable and unpredictable stimuli. We argue that this synchronised activity provides a framework that integrates imaging data, single cell recording and theoretical models and a means of elucidating the dynamics of reinforcement learning.

Experimental task

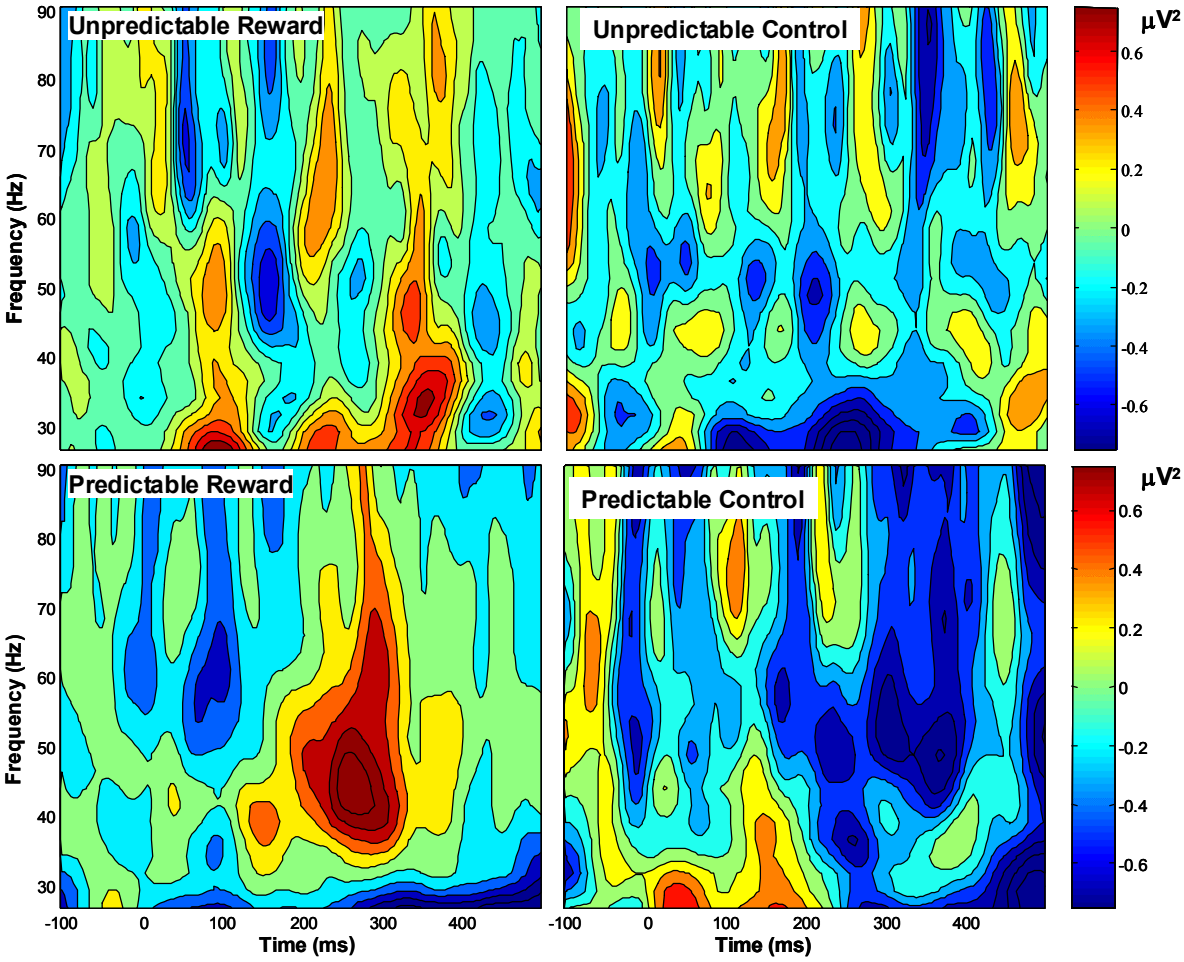
Twelve healthy right-handed subjects completed a card selection task with four experimental conditions: Unpredictable Reward (UR), Predictable reward (PR), Unpredictable Control (UC) and Predictable Control (PC). Each condition was run separately, in two recording blocks. On each trial, subjects selected one of four cards displayed on a computer screen positioned 1 meter in front of them. Subjects responded pressing one of 4 keys marked on a computer keyboard. On rewarded trials, a card with a **£** symbol appeared for 1000 ms. Each **£** sign was worth £0.50, and subjects were given a running total of their earnings after each trial. In the UR condition, subjects were unaware of when they would receive a reward or which card was rewarding. They were informed, however, that the same card would not provide a reward in two consecutive trials. Rewards were given in randomly, with probability 25%. In PR condition, subjects were informed that every fourth card choice would result in a **£** sign regardless of which card they picked. In the remaining 75% of the time a card with ***Φ*** appeared on the screen.

Two control conditions were used to counter balance the two rewarding conditions. Subjected selected the cards in the same way as above. In the UC condition, a card with # sign appeared with probability 25% while in the PC condition, every fourth card selection resulted in a # sign regardless of the card is selected. Subjects were informed that no reward was associated with the control conditions. To ensure that subjects maintained attention during the control conditions, they were asked to count the # signs, and were asked to report the count at three randomly spread occasions during each of the control blocks. In the remaining 75% of the time a card with **%** appeared on the screen.

Recording and data analysis

All subjects completed written informed consent approved by the local ethics committee. EEG was recorded continuously with an EGI (Electrical Geodesics Inc.) 129-electrode array with the vertex (position Cz) as a reference. As suggested by the EGI high input impedance amplifier, impedances were kept below 50 kΩ. Sampling rate was 500 Hz, and all channels were processed on-line by means of a 0.1 to 200 Hz band-pass filter. Eye movements were monitored with a subset of the 128 electrodes.

Oscillatory brain activity in the gamma-band is extracted using Complex Morlet wavelets and visualized on time-frequency power plots.



Grand mean (12 subjects), baseline-corrected Time-Frequency in the gamma band of the 4 conditions, averaged across frontal electrode sites corresponding to 10-20 positions Fp1, Fp2, AF3, AFz, AF4, and Fz. Reward conditions are clearly differentiated from the control. Other electrode sites displayed specific differentiation between rewards types (left parietal) and between predictability conditions (central parietal).