Good vibrations: Rhythms and plasticity in neural correlates of value in rat orbitofrontal cortex

van Wingerden, E.J.M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
GLOSSARY

Acquisition: see Reversal Learning

Actor/Critic Model: Actor/Critic models of reinforcement learning are a class of Temporal Difference (TD) Reinforcement Learning models that follow a division of labor in associative learning. Actor-critic models have a separate structure that maps states onto actions (i.e. if X, do Y). This structure is called the actor. A second structure, called the critic learns and stores the state value function (the outcome). Whenever the actor performs an action, the new state will be criticized by the critic through the computation of the TD error, which is the actual minus the expected outcome. Both actor and critic learning is based on this error.

Afferent: anatomical term for an incoming connection, i.e. axon terminals from a distant structure. See also Efferent

Agranular (cortex): anatomical term for layered cortex that lack a clearly identifiable layer 4, which usually holds granulated cell bodies. Layer 4 is a relatively recent evolutionary addition, thereby categorizing neural areas with agranular cortex as old(er) in evolutionary terms.

Coherence: coherence among neuronal groups is the pattern of phase-locking among oscillations in the communicating neuronal groups. Specifically, it indexes the consistency of phase alignment (the phase offset need not be zero) between two oscillatory processes over a time period.

After the idea for a Glossary emerged, I ran across Lucantonio et al. (2012), featuring a “Glossary of Terms” with substantial overlap to the terms on my list. For the sake of definition convergence, I have adopted their definitions (marked with a *) for terms that are also explained in this Glossary.
Common currency: a common currency is a representation of value on an abstract scale. When one wants to compare apples and oranges, one does so by converting their subjective properties in a judgment of value on a common scale: the common currency. It is thought that neural activity (BOLD activity in human fMRI experiments) in ventral striatum reflects the value of stimuli on a common currency scale.

Credit: One of the problems an animal encounters in solving a learning task through associative learning is to find out which stimulus, as part of the configuration of stimuli constituting the current environment or context holds the most predictive power, i.e. is most consistently aligned with the outcome and therefore should be credited (instantiated by increased synaptic strength between cue and outcome representations) for the outcome.

Critic: See Actor/Critic Model

Devalued; Reinforcer Devaluation*: A procedure in which the conditioned response to a cue is tested after the unconditional stimulus (for example, food) is separately devalued by motivational (for example, inducing satiation with prefeeding) or associative (for example, pairing the food reward with Lithium-Chloride-induced illness) manipulations. The normal consequences of such manipulations are a decrease in the conditioned responding to the cue previously paired with the reward.

Efferent: anatomical term for an outgoing connection, i.e. an axon projecting to a distant structure. See also Afferent

Exemplars: when associating cues with outcomes, multiple cues could be linked to the same outcome. In this case, it is interesting whether neurons discriminate between the different exemplars, i.e. the different cues that predict the same outcome. See for example Schoenbaum et al. (1999)
Chapter 7. GLOSSARY

Forward modeling: Arguably one of the major advances of evolved organisms over more primitive forms is the capacity to detach momentarily from the sensory input to play an internal scenario, a forward model, that predicts the course of events to happen, based on previous, learned experience. This capacity is essential for example to override instantaneous impulses when the long-term negative consequences of an action outweigh its initial positive outlook.

Gamma: a range of frequencies, usually defined between 30-100 Hz, found in brain oscillations and thought to reflect local interactions between interneurons and principal cells.

Go/NoGo discrimination learning task: A learning task based on Signal Detection Theory.

Goal-directed behavior*: Instrumental responding that is sensitive both to the contingencies between responses and outcomes and to the current desirability of these outcomes. As such, the hallmark of goal-directed behavior is its immediate flexibility in the face of changes in outcome availability or current value. As goal-directed responding is expected to adapt to new information immediately even if this information is acquired in a different context, it has been suggested that such behavior must rely on model-based representations and on flexible, online, mental simulation–based evaluations of actions. The term goal-directed behavior describes behavior at the phenomenological level, in contrast to model-based learning, which describes the computational processes that may explain goal-directed behavior.

Hits, Misses, False Alarms & Correct Rejections: See Signal Detection Theory.

Incentive: Pavlovian incentive learning reflects the acquisition of motivational properties by conditioned stimuli (CSs) through their association with appetitive and aversive reinforcers (Dickinson and Balleine, 2002).
Lobotomy: The surgical procedure employed to damage large parts of the frontal lobe. The procedure as used by Walter Freeman employed a switchblade that was used to cut white-matter axonal tracts, essentially disconnecting the frontal cortex from the rest of the brain. The procedure usually affected social behavior, though with high variance in the actual effects.

Model-free learning*: A set of reinforcement learning methods that use prediction errors to estimate and store scalar cue or action values from experience. These stored (‘cached’) values indicate the predicted total future reward if an action or cue are pursued, and they are used to bias choices of actions in order to gain maximal rewards. This decision strategy is simple but inflexible: the values are simply scalar numbers, separated from the identity of the expected future outcomes themselves or the specific events that will ensue en route to obtaining the outcomes.

Model-based learning*: A set of reinforcement learning methods in which an internal model of the environment is learned and used to evaluate available actions or cues on the basis of their potential outcomes. In these methods, values are not learned incrementally through prediction errors and stored for future use, but rather are computed ‘on the fly’ when needed, by mental simulation of sequences of events and outcomes using the internal world model. All that is required is a world model that includes predictions about the immediate consequences of each action or state in a sequence. This approach to action selection is computationally expensive owing to the need to mentally simulate and ‘test’ alternative courses of action, but it allows one to flexibly adapt behavior in a changing environment, and specifically, to adapt immediately to changes in the current value of the outcome. As goal-directed behavior is defined as a behavior that changes flexibly according to new knowledge about outcomes and event sequences, it has been suggested that this type of behavior must rely on model-based learning and decision mechanisms.
Chapter 7. GLOSSARY

Motivation: Motivation reflects a desire for a (basic) commodity like food, water, warmth or sex. The prospect of acquiring such a commodity constitutes an incentive. Motivation is assumed to drive internally generated action, an idea that can be traced back to Descartes, who wrote that “the passions dispose the soul to desire those things which nature tells us are of use, and to persist in this desire, and also bring about that same agitation of spirits which customarily causes them to dispose the body to movement which serves for the carrying into effect of these things”

N-methyl-D-aspartate or NMDA-receptor: A class of glutamate activated opn channel found in the brain that requires co-occurrence of glutamate binding and postsynaptic depolarization to open. In the open state, it allows influx of sodium, and, importantly, calcium. It is implicated in short and long-term modifications of synaptic strength

Pavlovian over-expectation*: A Pavlovian conditioning procedure in which two conditioned stimuli are first separately paired with the same unconditioned stimulus (for example, food) and then the cues are combined (in a compound cue) and paired with the same unconditioned stimulus. After such training, the subjects are given a test session in which cues are presented alone, without the unconditioned stimulus; subjects typically show a reduction in conditioned responding to the previously compounded cues, as soon as the first test trial. This decline in responding is thought to reflect a decrease in the cues’ associative strength during the compound phase, owing to the fact that during this stage the two cues together overpredict the unconditional stimulus, resulting in a negative prediction error that drives learning and hence reduction in cues’ associative strength
Phase-amplitude coupling (PAC): PAC reflects the tendency of one oscillatory rhythm to influence another rhythm. Biophysical constraints allow only for slower-to-faster interactions, where the phase of the slower rhythm influences the amplitude of the faster process. It is defined as the phase distribution (on the phase-providing, slow oscillation) of the faster rhythm’s amplitude envelope (instantaneous energy of the oscillation). If the amplitude distribution across phases is more consistent (i.e. large amplitudes at peaks, low amplitudes at troughs, for example), PAC takes higher values for magnitude. The phase of PAC reflects the preferred phase of coupling; this aspect is not investigated in this thesis.

Primary reinforcer: An object or stimulus that holds natural value (like food, water, warmth or sex) and can be used to strengthen a response. Primary reinforcers can fuel an internal motivation, prompting an animal to action to obtain it.

Resonance: Sub-threshold resonance of a neuron reflects a preference for inputs in a certain frequency regime, where ‘preference’ means that the input following the preferred frequency is proportionally maximal effective in eliciting output. The specific resonance tuning of a neuron’s membrane potential depends on the interplay of delayed or slow membrane currents that affect excitability by moving the potential towards or away from the threshold for spike generation in the wake of the initial refractory period after a spike has occurred.

Reversal learning*: Any task in which subjects first learn to discriminate among different cues associated with different probabilities of reward and punishment (acquisition phase), and then the cue-outcome (reward) associations are reversed (reversal phase).

Reward or outcome prediction: According to reinforcement learning theory, a cue can come to elicit an outcome prediction through associative learning. Through many pairings, the cue is associated with the outcome so that it can act as a substitute for this outcome to drive behavior, for example to approach or avoid a food well.
Signal Detection Theory: Signal Detection Theory describes whether a stimulus is correctly decoded by an observer. Suppose there are two stimuli, one instructing to go (A), and one to withhold a response (B). If A is transmitted, and the observer acts, this constitutes a Hit. If she fails to act, it is a Miss. If, on the other hand, B is transmitted, and the observer acts, this constitutes a False Alarm (an unwarranted response), where the correct course of action would be to withhold: a Correct Rejection response (Green and Swets, 1966; Macmillan, 2002)

Spike-Timing Dependent Plasticity or STDP: STDP is found in mammalian synapses and relates to the finding that the exact timing of pre- and postsynaptic spikes determines whether the synapse is strengthened or weakened. Specifically, when postsynaptic firing follows presynaptic firing, strengthening occurs, in correspondence with the principle of causal association, already put forward by Aristotle

Theta: a range of frequencies, usually defined between 4-12 Hz, found in brain oscillations.

Value (cached)*: A concept linked to the term ‘Utility’, used in economics. A goal’s value designates the amount of subjective pleasure associated with attaining this goal and can be treated as a mathematical entity. It is assumed that free choice between options “reveals preferences” for those options, and that these revealed preferences reflect the ratio of utility derived from gaining them.