When is it worth working: Brain circuits related to motivation and reward expectancy.

Barry J. Richmond,
Laboratory of Neuropsychology
National Institute of Mental Health National Institutes of Health,
Bethesda, MD, USA

Most, if not all animals, base a large part of their behavioral repertoire on associations between context and predicted outcome. If, for example, we are watching television and an advertisement for food appears, we might realize that we are hungry, and we might seek food (the advertiser hopes that we will choose to seek their product). How does the incentive to act arise, and how do the associations leading to this behavior arise?

Using simple behaviors in which visual cues or other contexts indicate outcome contingencies such as reward size or whether a reward will be delivered or not, we can induce systematic changes in motivation in monkeys. We assume the behavioral performance is proportional to the value of the outcome predicted by the visual cues, or other contextual information, thereby providing a simple measure of motivation. The behavioral adjustments arise within a few trials (2-4) after the cues are first presented, and the value implied by the cue seems to be learned without effort, even when the behavioral response seems counter-productive. The behavioral adjustments are long-lasting, lasting seemingly forever. This apparently simple behavioral approach allows us to examine many aspects of value calculation and perception.

We have found differentiable signals related to predicted outcomes in many brain regions and/or neuronal classes (12 or 13 different signals). Using classical selective ablation studies we have learned that the rhinal cortex in the anterior medial temporal lobe, the orbitofrontal cortex and the lateral prefrontal cortex play different roles in learning and predicting the outcomes.

We also have been looking at where in the brain the generalization that was implied in the first paragraph (recognizing a whole class of stimuli) might arise. Surprisingly our monkeys do rapid visual stimulus generalization (20-40 exemplars of a class are enough) after complete lateral prefrontal cortex ablations, and even after inferior temporal area TE ablations.

When there are schedules of trials, we can arrange circumstances in which value seems to be as expected, but performance is severely suboptimal (in terms of maximizing reward over time). Thus, the monkeys never adapt to improve their performance as classical reinforcement learning (implemented as a temporal difference model) predicts. In this task performance remains 'suboptimal' forever, at least in terms of obtaining reward. We also see that in these schedules of trials, the monkeys are sensitive to the number of trials already performed, again in contrast to what simple temporal difference learning would predict. This latter suggests that there is a framing or 'sunk cost' effect as is seen in human investment behavior.

References

Behavior and theory:


and Upcoming Reward in Monkey Orbitofrontal Cortex, *Cerebral Cortex*:18, 93-103.


*Behavior with intervention:*


