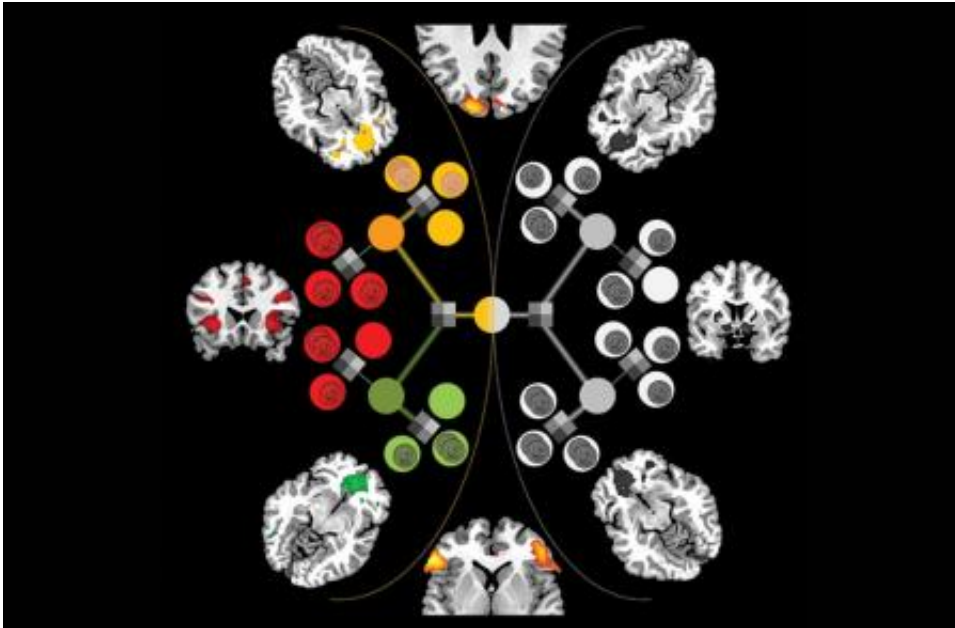# A tour of neural reinforcement learning

February 9 2015, by Douglas Smith

Several areas of the brain are activated in the process of making value judgements, as these fMRI scans show. Credit: John O'Doherty/Caltech

It is often said that people who do not learn from history are doomed to repeat it. Not being one of those people requires a network of different brain regions to work in concert.

John P. O'Doherty, professor of psychology and director of the Caltech Brain Imaging Center, discusses our current understanding of how we learn from experience.

Q: What do you do?

A: I study how we learn from experience. Humans and other animals have to make decisions all the time to maximize their benefits and minimize danger. These decisions range from what I should have for dinner or should I cross the road—which could have life-changing consequences if I'm wrong—to the selection of a life partner. I don't claim that "Who should I marry?" is equivalent to "Carrots or Brussels sprouts?" but we do think that many decisions share certain commonalities. So we look at very simple tasks that give us a window into how the brain solves problems to maximize future rewards.

We study brain activity by putting your head in an fMRI scanner. "MRI" stands for magnetic resonance imaging, and you've probably had one if you've had a sports injury. The "f" stands for "functional," and an fMRI scan detects changes in the oxygenation levels in the blood. If a certain part of the brain is active, its oxygen supply increases. We map those increases onto the brain's anatomy in 3-D while our volunteers perform some task that involves learning.

A task might be playing virtual slot machines. You have a choice of three machines, and we tell you one machine pays better than the others. So you choose one, press the button, and get instant feedback—you win or you lose. As you try to work out which machine is better, we monitor the patterns of activity in various parts of your brain. One of our goals is to find the part of the brain that represents the experienced value of the things we meet in the world—how good it feels to win, or how bad to lose.

We're also interested in how the brain changes its expectations. As you play the machines, you're constantly revising your estimate of which machine is better. We have computational models that we think represent how the brain internalizes feedback, and we're trying to find

[brain areas](link) where the activity matches those models.

We think that understanding the neural circuits and computations that underpin our decision-making capacity may shed some light on certain psychiatric disorders, such as obsessive-compulsive disorder, depression, and addiction. On some level, all of these can be seen as decision-making gone wrong. Addiction, for example, involves a choice—voluntary or otherwise—to engage in a certain pattern of behavior.

Q: Setting aside clinical disorders, why do people make garden-variety bad decisions? What leads us to cross a busy road and almost not make it?

A: First, it's important to emphasize that humans are collectively pretty good at making decisions. That's why we've been so successful as a species. But there could be all sorts of reasons why an individual might make a poor decision. For example, you might underestimate how fast the traffic is moving.

My lab is particularly interested in how two distinct decision-making mechanisms may interact to produce bad outcomes. One mechanism is "goal-directed," in which you evaluate the consequences of your action in light of the goal you're pursuing. This requires a lot of mental energy. In contrast, "habit-controlled" decision-making is basically stimulus-response—you react to some cue from the environment. Habits can be very beneficial, because you can execute them quickly without thinking deeply. Once you learn to ride a bicycle, for example, you don't have to concentrate on keeping your balance. It becomes routine, and you can focus your mental energy on other things. Poor decisions can result when the habit system drives your behavior when you really should be solving things in a goal-directed manner. This may be how addiction becomes

compulsive. The goal-directed system says, "I don't want to take this drug any more," but the habitual system overrides it.

Q: How did you get into this line of work?

A: Even as a kid I was interested in science and its unsolved mysteries. I was actually keen on astronomy as a teenager and really considered going in that direction. Then I started getting interested in how computers work, which led me to start wondering about how the most complex computer that we know of works, namely our brain. So I basically had a career choice between studying the universe or studying the brain, which are probably the world's two greatest outstanding mysteries. I decided to take my chances on the brain.

At the time, the field of cognitive neuroscience was based on the paradigm that the brain is like a digital computer, and brain processes were modeled in essentially in the same way. There were lots of studies of memory, such as recalling lists of words, but very little was known about how the brain assigns a greater value to some things than others. But it's a really fundamental question, because the ability to work out whether something is good or bad—and to maximize behaviors that lead to good things and avoid bad things—is critical for survival. Digital computers typically don't make value judgments of that sort unless they are programmed to do so. So that's what excited me, trying to unlock how it is that the brain assigns value to things in the world.

Provided by California Institute of Technology

study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.