Article

Neuron

Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex

Highlights

- Monkeys transition between exploratory and exploitative goal states
- During exploration, prefrontal choice-predictive activity is virtually absent
- Prefrontal dynamics are disrupted during exploration both within and across trials
- Exploration enhances reward-dependent learning in brain and behavior

Authors

R. Becket Ebitz, Eddy Albarran, Tirin Moore

Correspondence

rebitz@gmail.com

In Brief

Exploratory choices permit the discovery of new rewarding options. Ebitz et al. report that spatially selective, choice-predictive neurons in the prefrontal cortex do not predict choice before exploratory decisions. Reduced prefrontal control may underlie flexible decision-making and trial-and-error discovery.

Ebitz et al., 2018, Neuron *97*, 1–12
January 17, 2018
https://doi.org/10.1016/j.neuron.2017.12.007

CellPress

Please cite this article in press as: Ebitz et al., Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex, Neuron (2017), https://doi.org/10.1016/j.neuron.2017.12.007

Neuron
Article

CellPress

Exploration Disrupts Choice-Predictive Signals and Alters Dynamics in Prefrontal Cortex

R. Becket Ebitz,[1,3,4,*] Eddy Albarran,[1] and Tirin Moore[1,2]
[1]Department of Neurobiology, Stanford University School of Medicine, Stanford, CA 94305, USA
[2]Howard Hughes Medical Institute
[3]Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA
[4]Lead Contact
*Correspondence: rebitz@gmail.com
https://doi.org/10.1016/j.neuron.2017.12.007

SUMMARY

In uncertain environments, decision-makers must balance two goals: they must "exploit" rewarding options but also "explore" in order to discover rewarding alternatives. Exploring and exploiting necessarily change how the brain responds to identical stimuli, but little is known about how these states, and transitions between them, change how the brain transforms sensory information into action. To address this question, we recorded neural activity in a prefrontal sensorimotor area while monkeys naturally switched between exploring and exploiting rewarding options. We found that exploration profoundly reduced spatially selective, choice-predictive activity in single neurons and delayed choice-predictive population dynamics. At the same time, reward learning was increased in brain and behavior. These results indicate that exploration is related to sudden disruptions in prefrontal sensorimotor control and rapid, reward-dependent reorganization of control dynamics. This may facilitate discovery through trial and error.

INTRODUCTION

In complex environments, reward contingencies are seldom fully known. In these circumstances, there is a limit to the effectiveness of an "exploitative" strategy. Trying to maximize immediate reward by repeatedly choosing known-value options risks missed opportunities to discover better alternatives. Thus, decision-makers occasionally deviate from exploiting in order to "explore"—they sample alternative actions, gather information about the environment, and thereby increase the potential for future reward (Kaelbling et al., 1996; Sutton and Barto, 1998). Designing a system flexible enough to both exploit and explore is a classic problem in reinforcement learning (RL) (Sutton and Barto, 1998), and its solution is a prerequisite for intelligent, adaptive behavior in natural decision-makers (Rushworth and Behrens, 2008). However, only a few studies have examined how exploration is implemented in the brain (e.g., Daw et al., 2006; Quilodran et al., 2008; Pearson et al., 2009; Kawaguchi et al., 2015), and it remains unclear how the mapping of sensory input onto motor output is adjusted in order to pursue these different strategies in an otherwise identical environment.

The brain is biased toward representing and selecting rewarding options. For example, neurons in oculomotor regions such as the frontal eye field (FEF) (Leon and Shadlen, 1999; Roesch and Olson, 2003, 2007; Ding and Hikosaka, 2006; Glaser et al., 2016) and the lateral intraparietal area (LIP) (Platt and Glimcher, 1999; Sugrue et al., 2004) signal high-value gaze targets more robustly than low-value targets. At the behavioral level, high-value targets cause rapid, vigorous orienting responses (Takikawa et al., 2002; Reppert et al., 2015), and previously rewarded options continue to capture gaze and bias attention even when explicitly devalued (Takikawa et al., 2002; Anderson et al., 2011; Hickey and van Zoest, 2012). This bias improves the detection of goal-relevant targets and would help during exploitation. However, it interferes with the goal of exploring alternative options.

How can the brain efficiently overcome its reward-seeking bias in order to discover better options? One way might be to choose more randomly during exploration, perhaps by adding noise or indeterminacy to neural computations involved in choice and attention. This is an efficient way to produce exploration in artificial agents (Sutton and Barto, 1998), and humans also seem to explore largely randomly (Wilson et al., 2014). However, random selection in behavior need not imply an indeterminate selection process in the brain, and there is no empirical evidence for indeterminate selection. Alternatively, the representations of chosen options could be enhanced during exploration, perhaps due to some bias toward uncertain options (Rushworth and Behrens, 2008; Schultz et al., 2008). This latter hypothesis might have cognitive consequences. For example, in regions involved in directing attention, increasing choice-selective representations could increase reward learning, because attention facilitates learning (Pearce and Hall, 1980; Swan and Pearce, 1988; Pearce and Bouton, 2001; Niv et al., 2015). Such an observation could provide a mechanistic basis for normative accounts that predict that learning should increase during exploration (Kaelbling et al., 1996; Sutton and Barto, 1998; Yu and Dayan, 2005; Daw et al., 2006; Cohen et al., 2007; O'Reilly, 2013). Of course, learning could increase during exploration via other mechanisms. It remains unclear whether learning is increased

