# Structured Credit Assignment in Mice

Kevin J. Miller[1,2], Laurence Freeman[2], Yu Jin Oh[2], Matthew M. Botvinick[1,2] and Kenneth D. Harris[2]

1. Google DeepMind 2. University College London

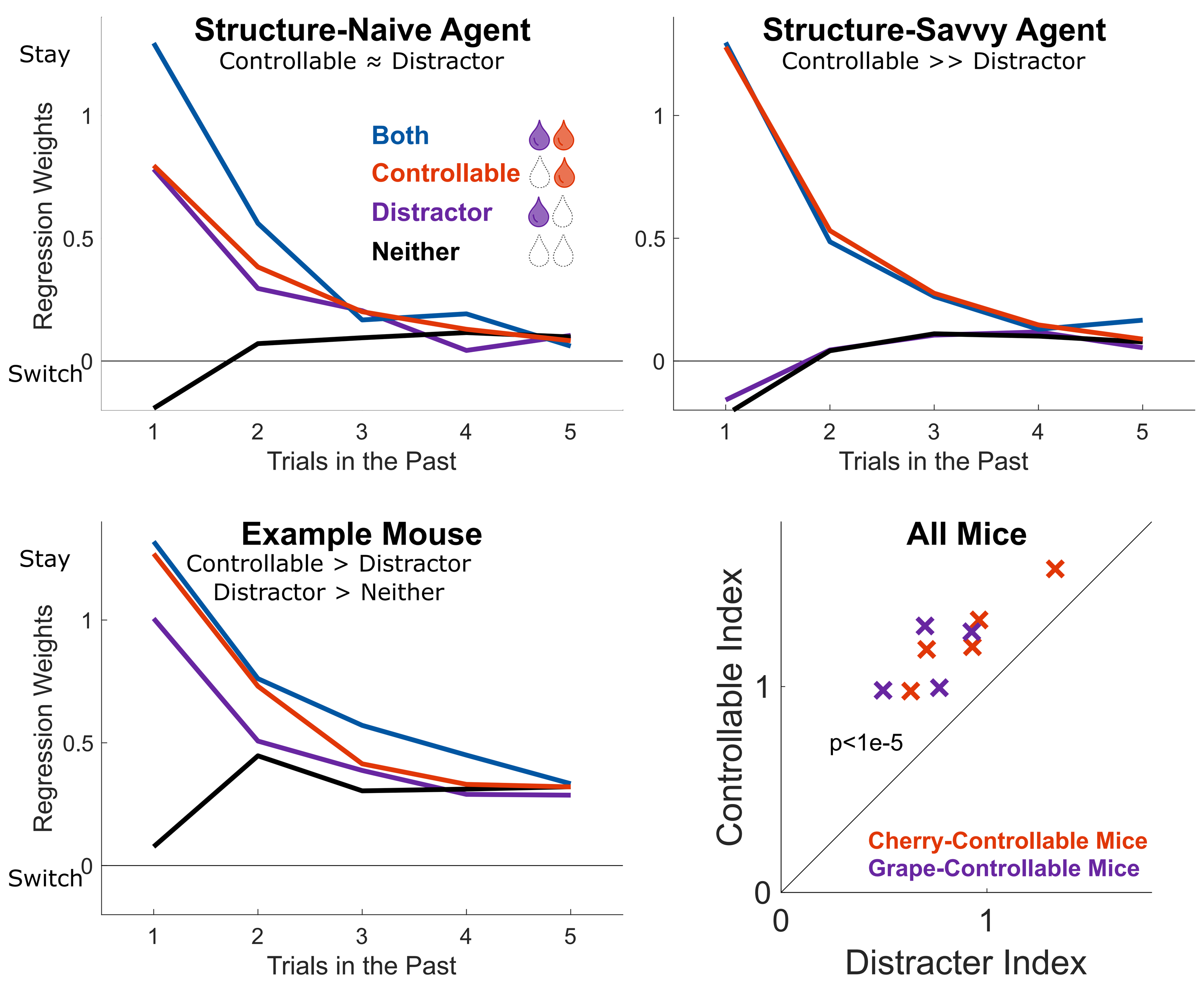
## Introduction

- Reinforcement learning requires associating rewards with one or more previous states or actions. The “credit assignment problem” asks which of many possible associations to form.
- Good solutions to the credit assignment problem result in rapid learning, a major goal of AI research. **"Structured credit assignment"** strategies (Harutyunyan et al., 2019; Mesnard et al., 2021; Raposo et al., 2021) involve use of known task structure to guide credit assignment.
- Neuroscience research (Gershman et al., 2009; Jocham et al., 2016; Moran et al, 2019) suggests that humans also use **"noncontingent credit assignment"**, associating reward broadly to prior actions.
- Studies of credit assignment in the brain typically focus on humans or primates, limiting the experimental toolkit available.

## Mice Weight Controllable More than Distractor Rewards

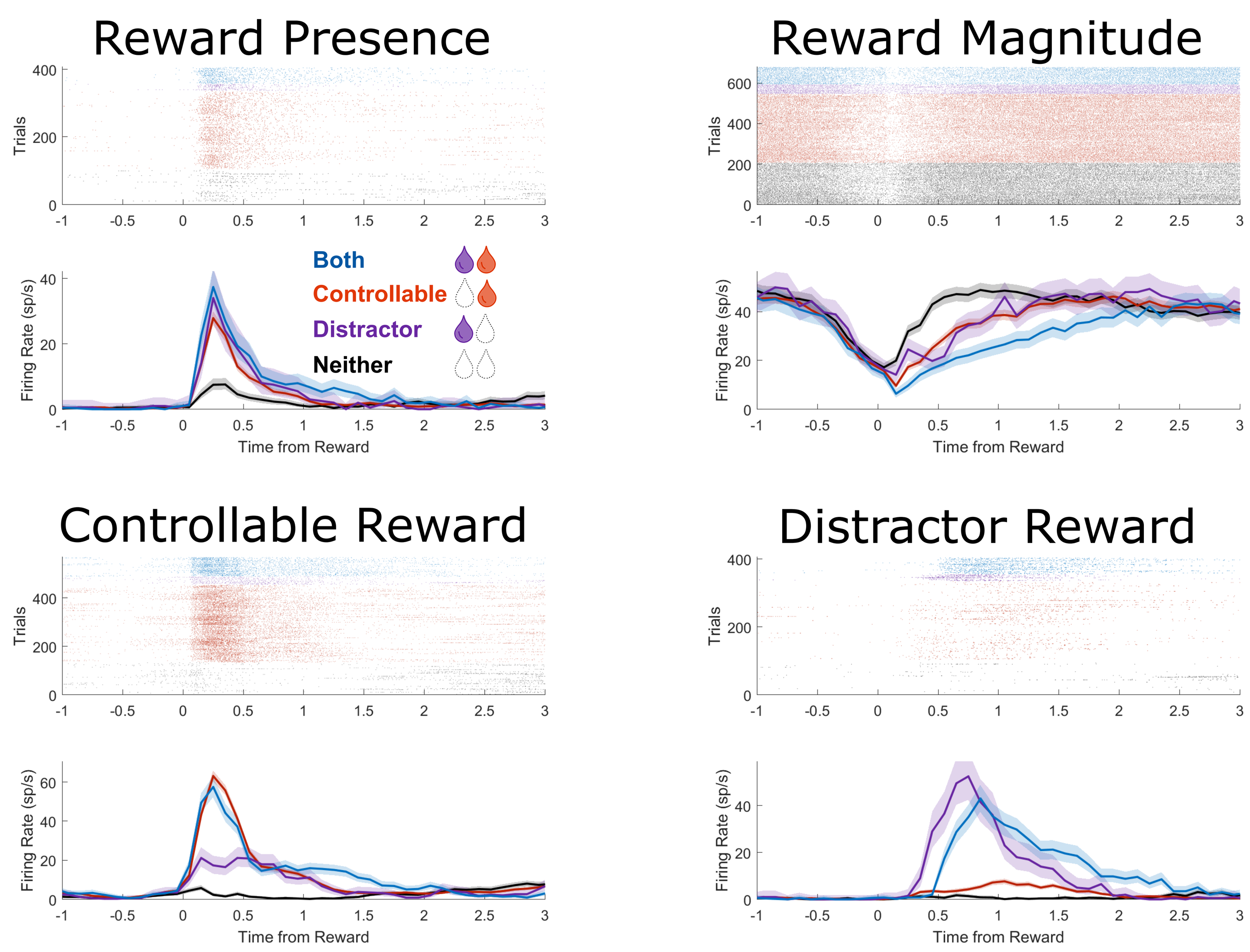
Structured Learning: Controllable rewards affect choice more than distractor rewards

Noncontingent Learning: Distractor rewards still affect choice



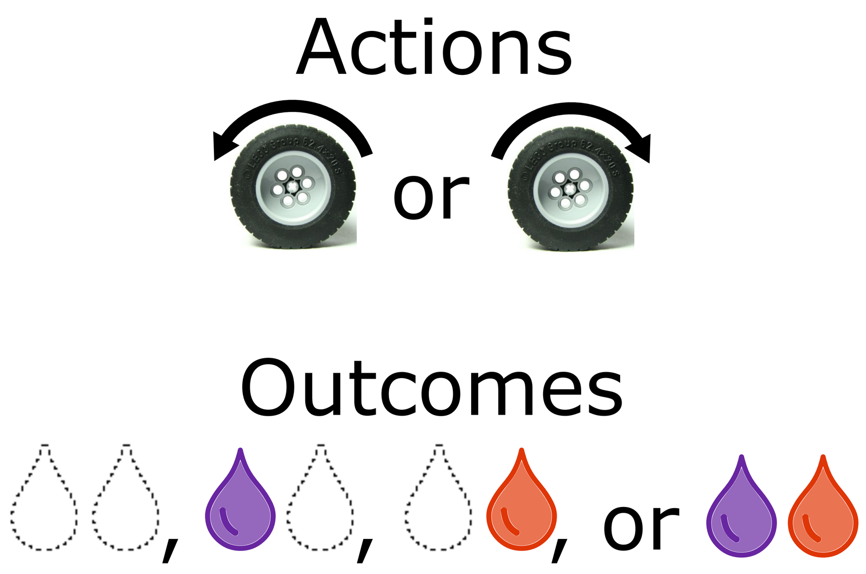

## Neural Responses to Reward

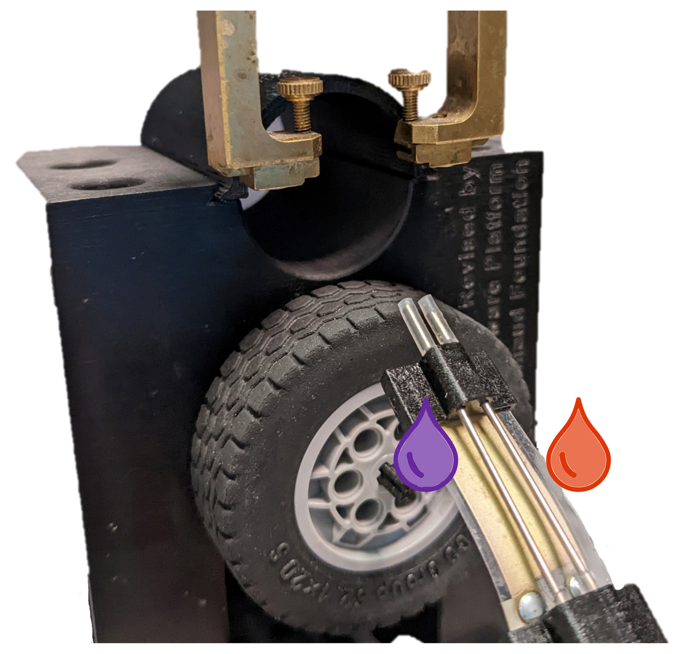
Example Units: Diversity of Responses to Reward



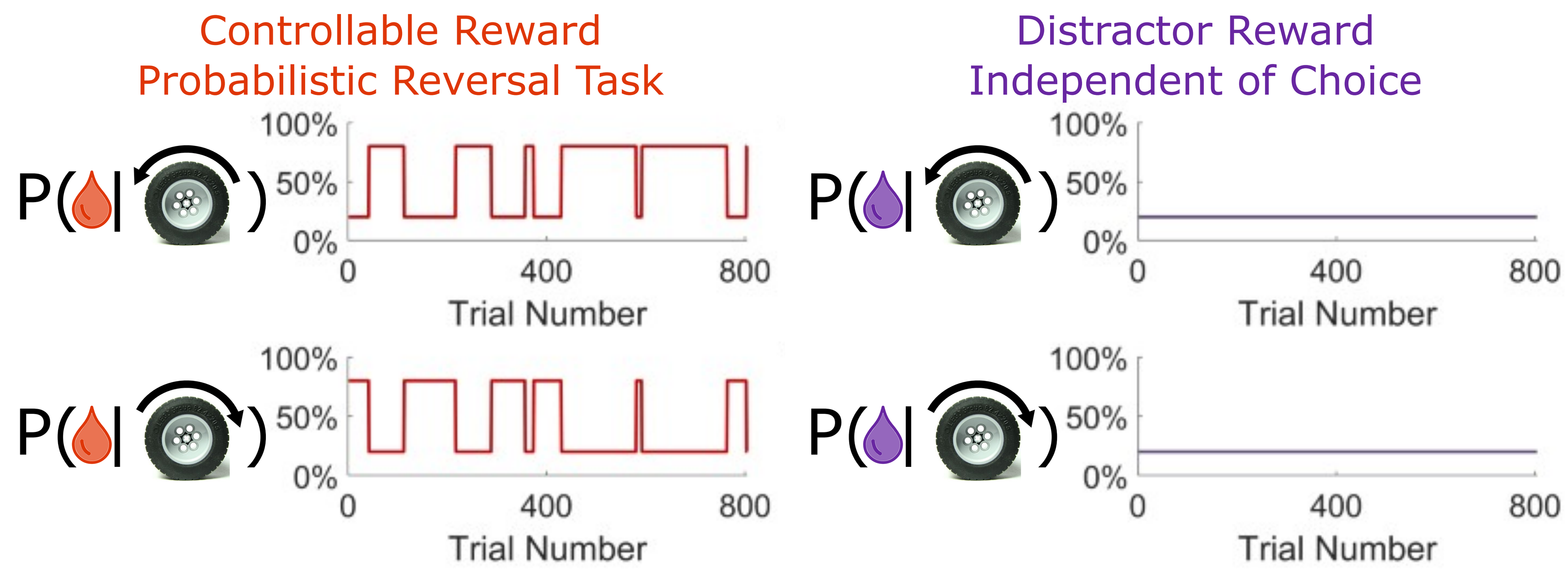

## Distractor Rewards Task for Mice

Rewards of one flavor ("controllable") are causally dependent on mouse's choice. Rewards of another ("distractor") are independent of choice (Jocham et al., 2016).



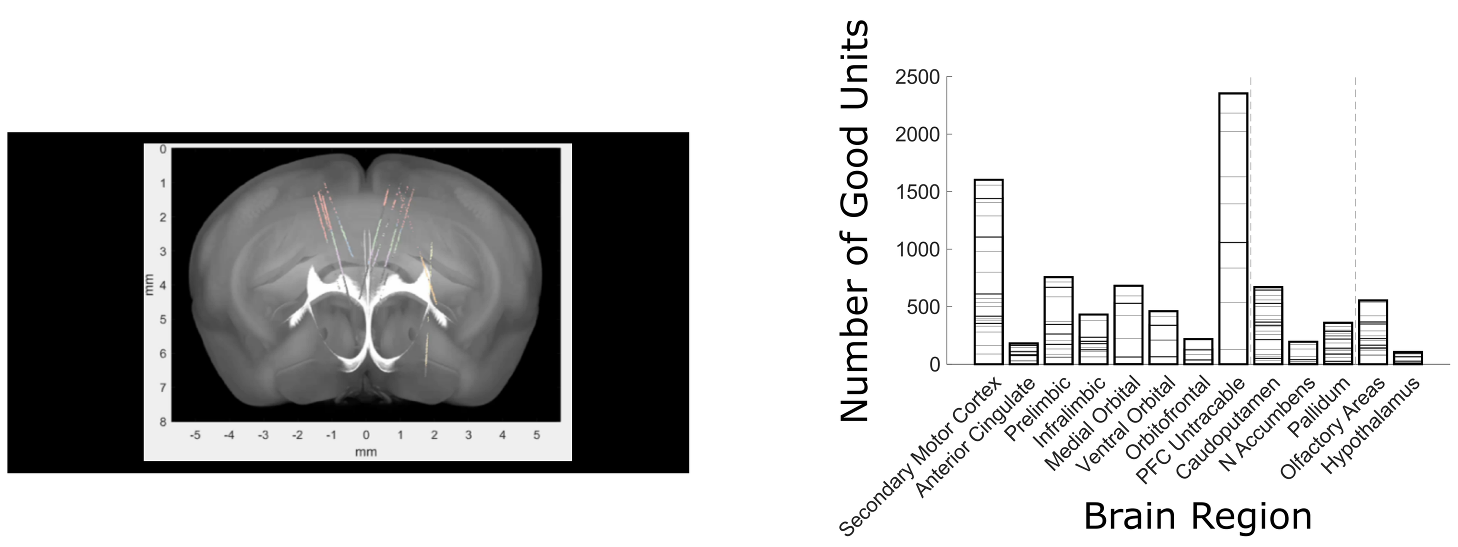



### Example Session




## Neuropixel Recording Sites

Large dataset covering prefrontal cortex and striatum




## Conclusions

- Mice show behavioral evidence of both structured and noncontingent learning
- Neurons in many brain regions respond to rewards of both types

## Directions

- Do neural responses weight rewards differently?
  - Weight both equally: Reward as sensory/motor/hedonic
  - Weight controllable: Reward as signal to learn
- Do neural responses integrate past rewards differently?
  - Weight both equally: Memory of past rewards
  - Weight controllable: Expectation of future reward