DEEP REINFORCEMENT LEARNING AND ITS NEUROSCIENTIFIC IMPLICATIONS

BY: MATTHEW BOTVINICK, JANE X. WANG, WILL DABNEY, KEVIN J. MILLER, AND ZEB KURTH-NELSON

INTRODUCTION

- Deep learning has been studied to model numerous systems:
  - Vision, audition, motor control, navigation, and cognitive control

- Neuroscience applications of DL can be traced back to the 1980s

CONTENT

- Conceptual and historical introduction of deep RL
- Highlight studies that have explored relationship between deep RL and brain
- Discuss broad set of topics of interest and caveats

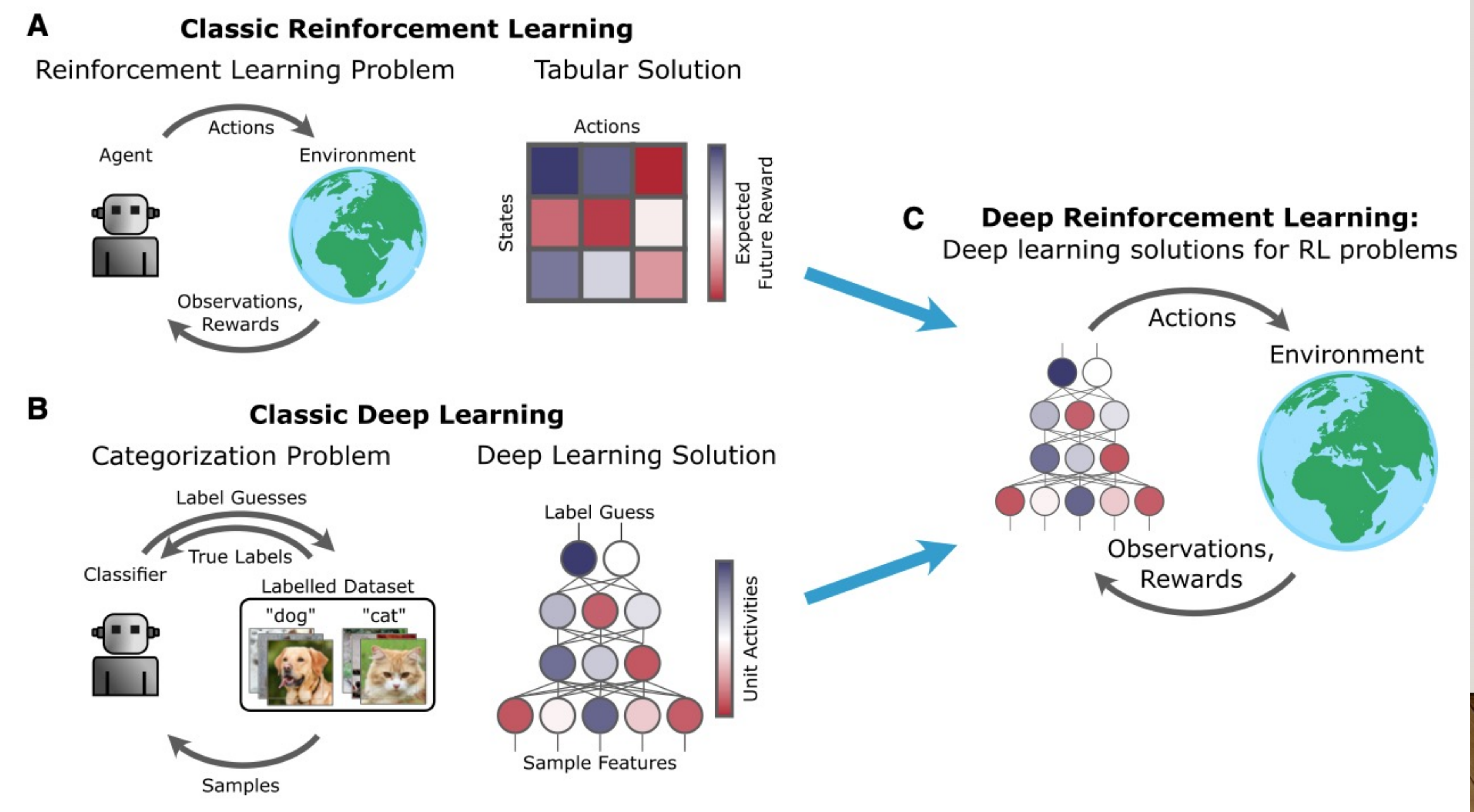
INTRODUCTION TO DEEP RL

- In contrast to supervised learning, the agent does not receive explicit feedback whether it is performing correct actions

- Early RL work used tabular representations -> inefficient

- Generalization across states: function approximation -> DL

INTRODUCTION TO DEEP RL



FUNDAMENTAL DIFFERENCES FOR RL

- Unlike (un)supervised learning requires exploration (trade-off exploration vs exploitation)

- Representations are shaped by exploration and actions

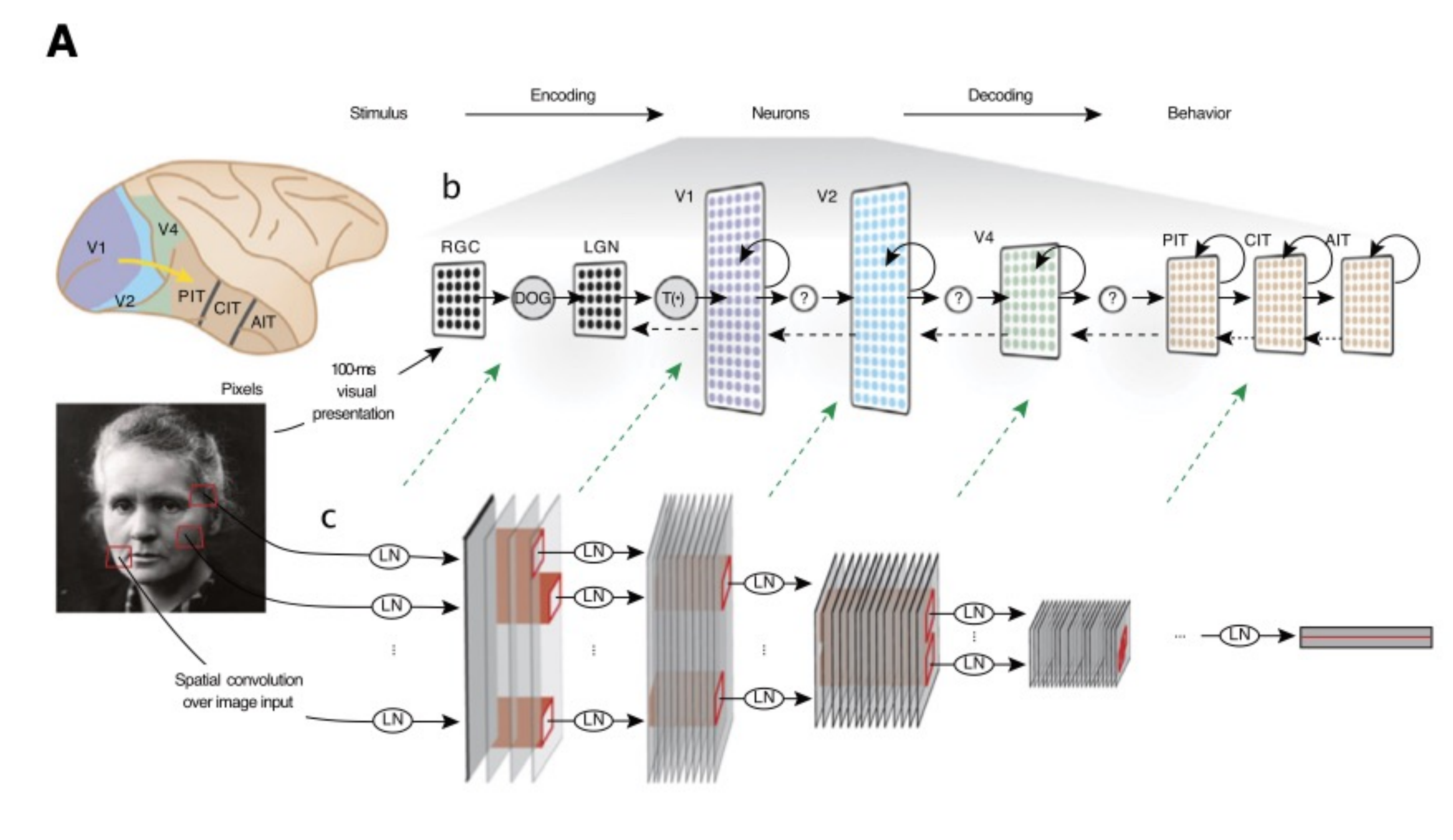
INSTABILITY RL

- Earlier work of RL had issues with collapsing (TD-Gammon)
- DQN -> much more stable
- Example: "experience replay" (past transitions are seen randomly during training)

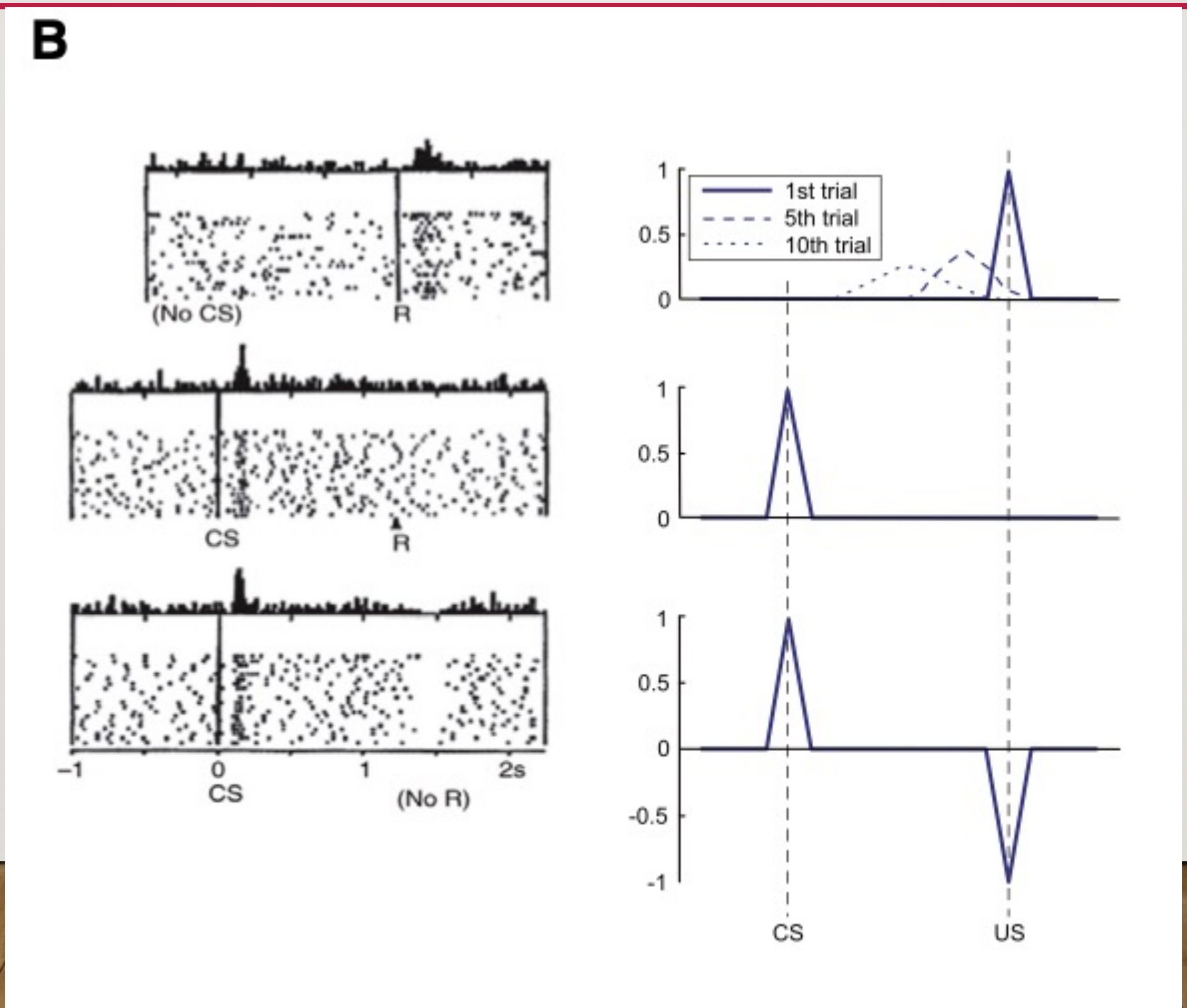
NEURAL REPRESENTATIONS

- DL has been used to study neural representations
- This does not help us understand goal-directed behavior within a sensory-motor loop
- RL already has provided theory wrt mechanisms of learning and decision making

EXAMPLES

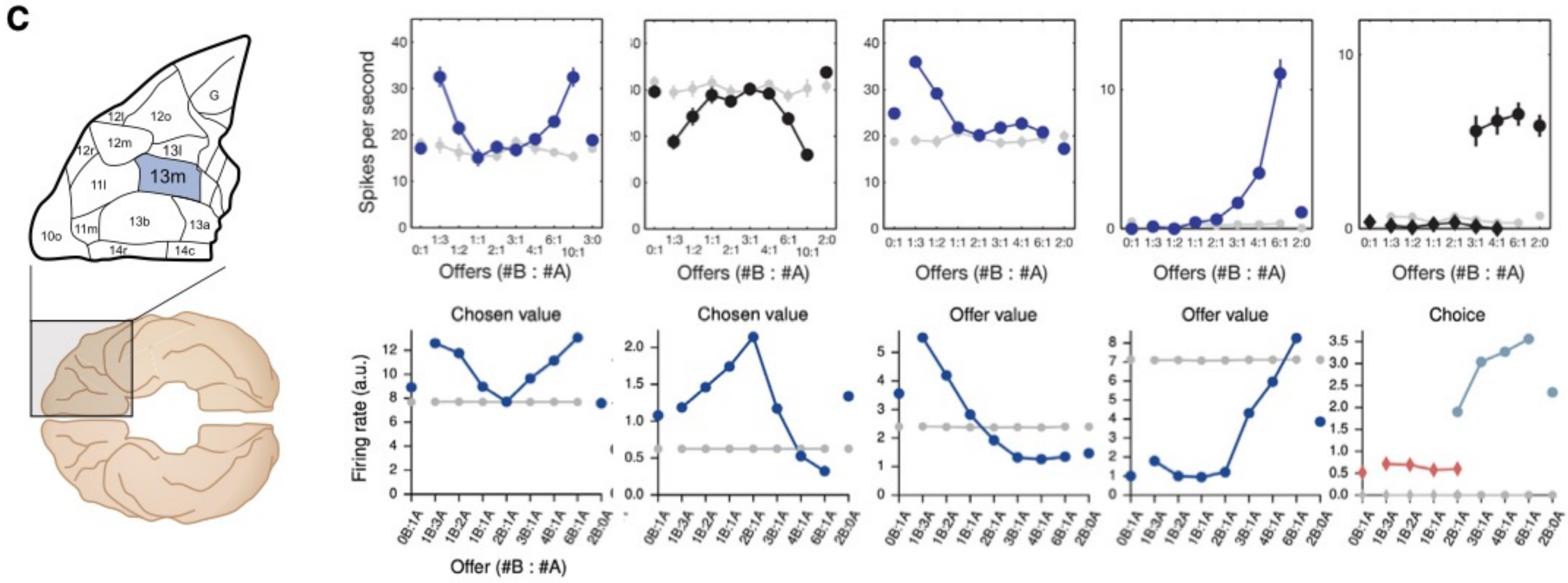

A
Stimulus
Encoding
Neurons
Decoding
Behavior
b
V1
V4
V2
PIT
CIT
AIT
RGC
DOG
LGN
T(·)
V1
V2
V4
PIT
CIT
AIT
?
Pixels
100-ms visual presentation
c
LN
Spatial convolution over image input

EXAMPLES



B
(No CS)
R
CS
R
-1
0
CS
1
2s
(No R)
1st trial
5th trial
10th trial
1
0.5
0
-0.5
-1
CS
US

EXAMPLES



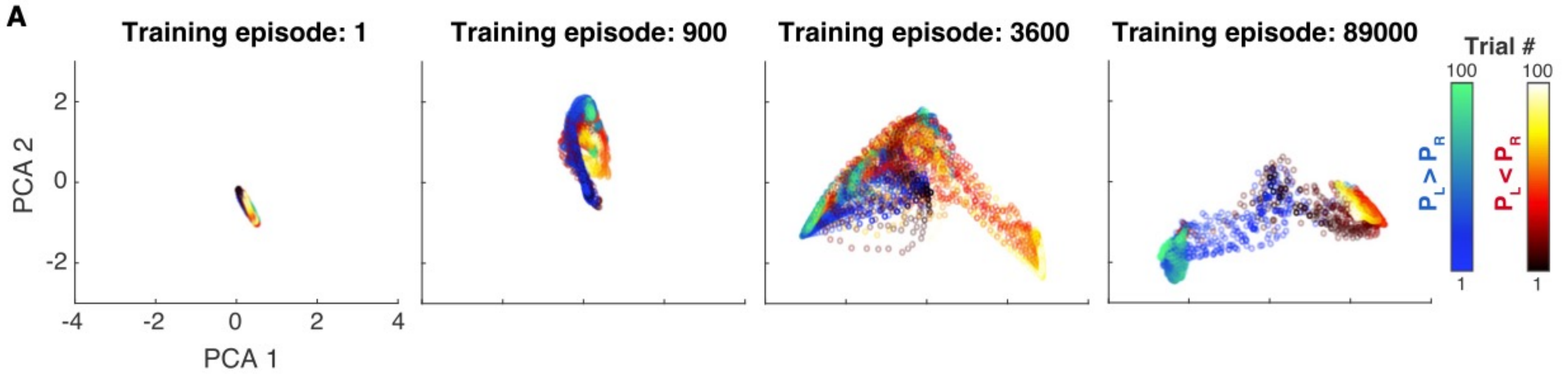
c
G
12l
12o
12m
12r
13l
13m
11l
13b
13a
10o
11m
14r
14c
Spikes per second
Offers (#B : #A)
Chosen value
Chosen value
Offer value
Offer value
Choice
Firing rate (a.u.)
Offer (#B : #A)

EMERGENCE

- New phenomena emerge with RL: processed by which representations support and are shaped by reward-driven learning and decision making

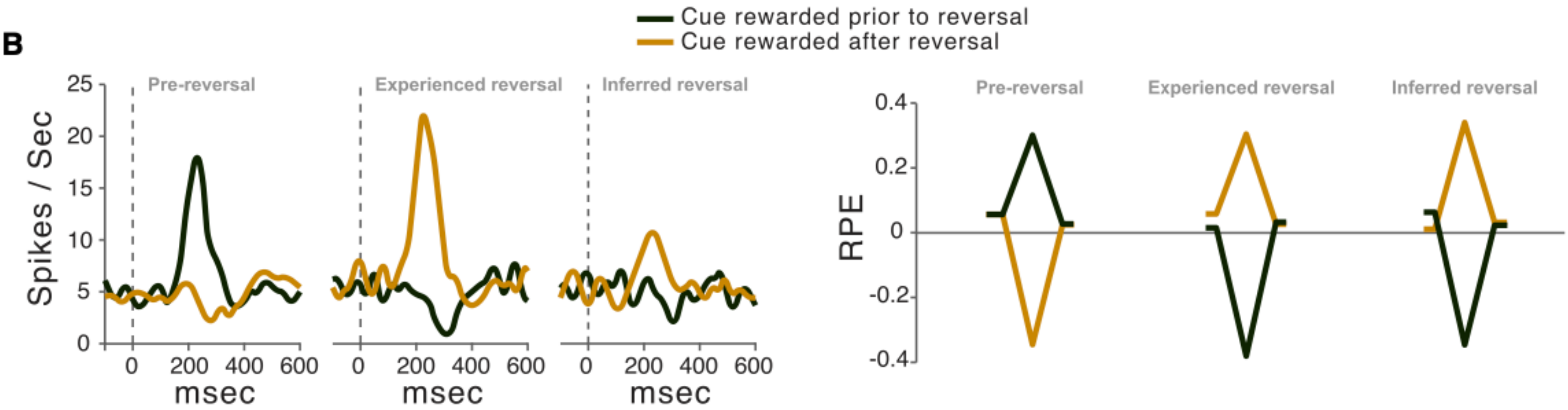
- Deep RL is unique in that it's different from DL and RL alone

EMERGENCE EXAMPLES



Trained on a series of interrelated tasks: deep RL networks will develop the ability to adapt to new tasks of the same kind without weight changes

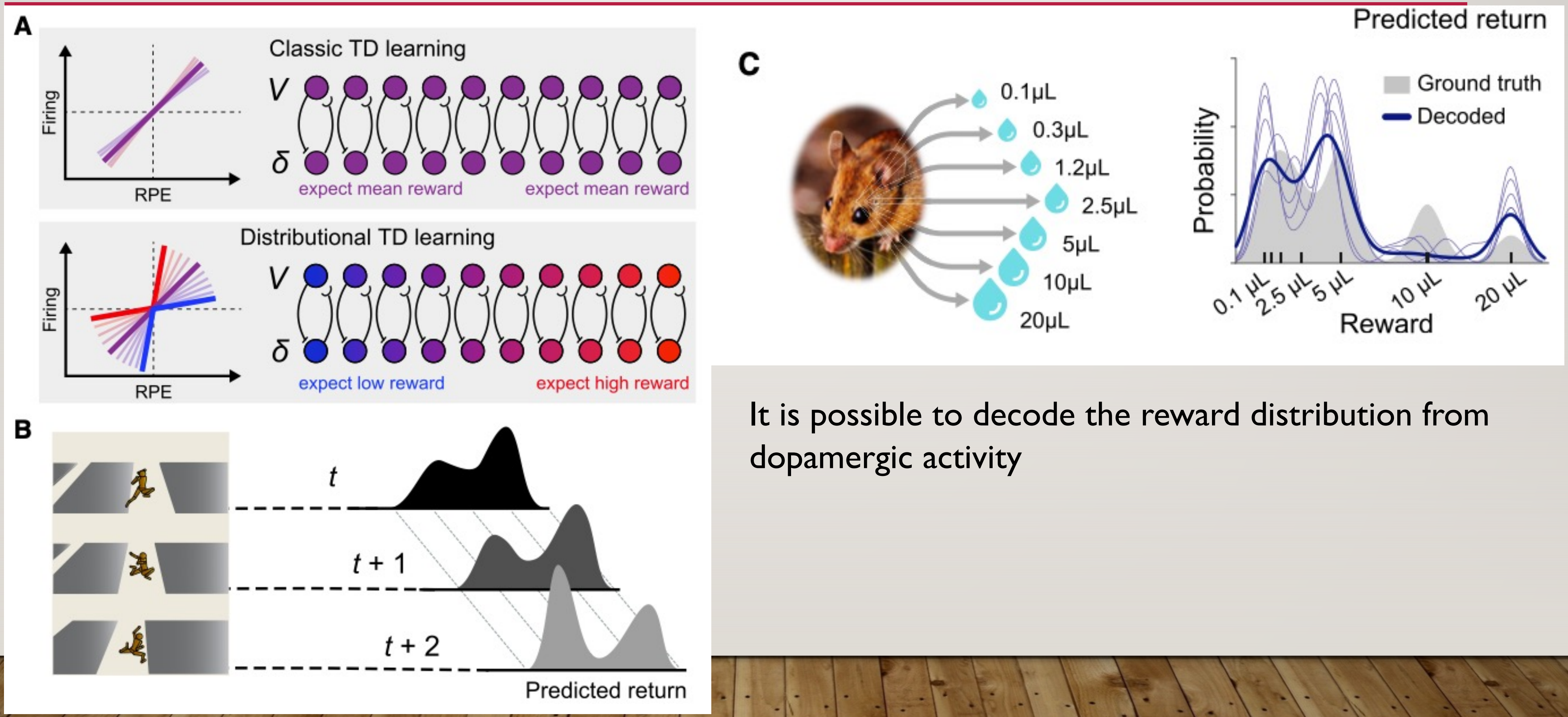
EMERGENCE EXAMPLES



Dopaminergic activity in response to cues ahead of a reversal

Reversal: of what target yielded juice

TD LEARNING



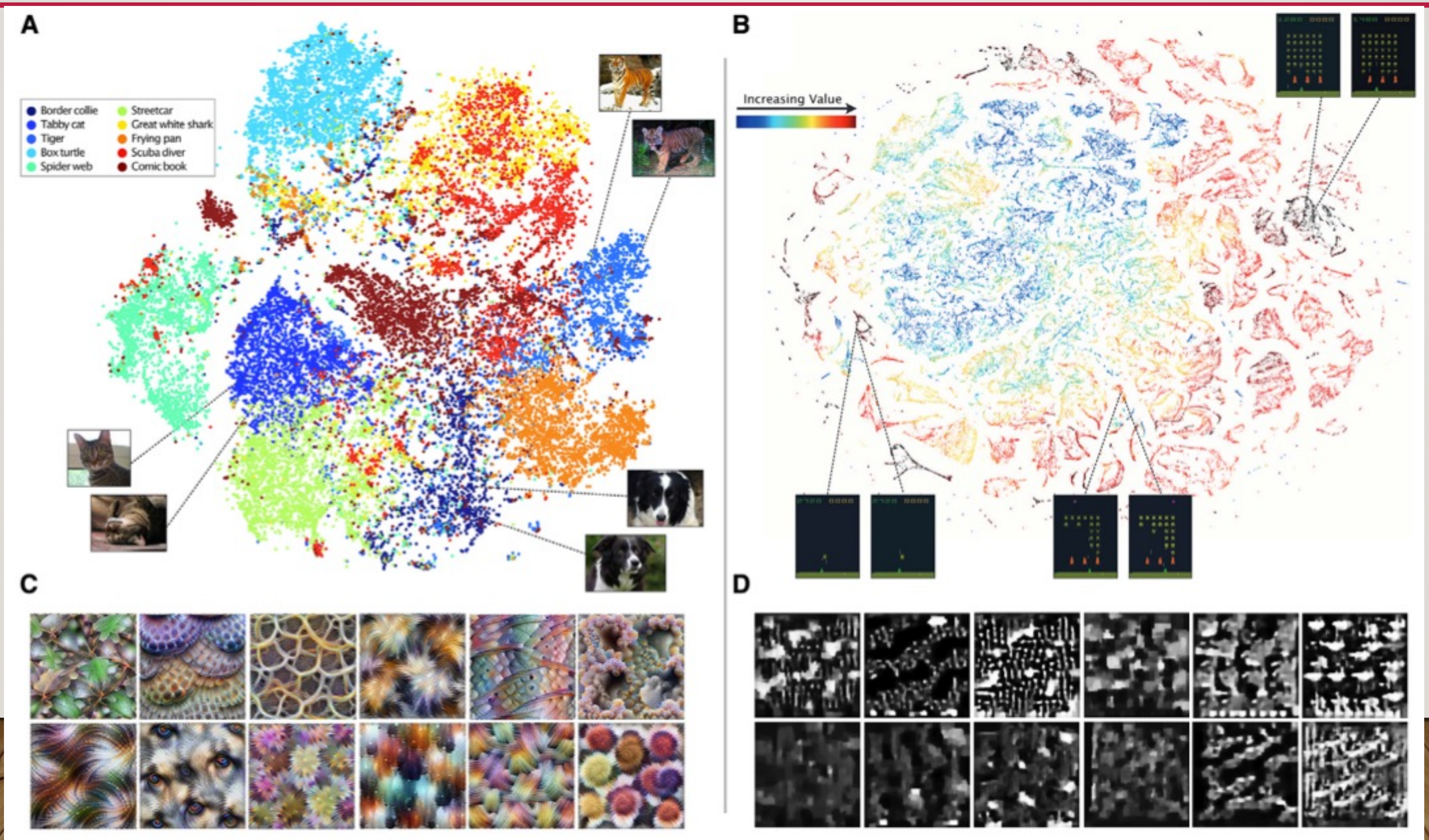
It is possible to decode the reward distribution from dopamergic activity

AREAS WHERE DEEP RL MAY BE INTERESTING

- Representation learning (shaped by action and reward)

- Supervised learning: similar reps to similar labels, deep RL associates images to images with similar functional implications

DEEP RL REPRESENTATIONS



MEMORY

- Experience replay
- The mechanisms in LSTMs have been proposed to be used in the brain

EXPLORATION

- Basing exploration not on novelty but on uncertainty
- Or hypothesis-driven experimentation, which make them better to investigate the neural basis for strategic exploration in animals

COGNITIVE CONTROL AND ACTION HIERARCHIES

- Low-level systems that operate autonomously, and the higher level system only intervenes at a cost

- Deep RL systems have been configured to operate at different timescales at different levels (slower at high level, faster at low level)

SOCIAL COGNITION

- Deep RL with multiple agents can be computational leverage

CHALLENGES AND CAVEATS

- Deep RL's demand for large amounts of data
  - Deep RL systems have not been proven to be capable of matching humans in terms of flexibility

- Long term credit assignment, updating behavior on the basis of rewards that may not accrue until substantial time after actions that were responsible for generating them

CHALLENGES AND CAVEATS

- Most deep RL is still done in an engineering context and not neuroscientific context to model the brain

QUESTIONS