Modeling managment of access to working memory as a self-evalution process for intrinsically motivated prediction

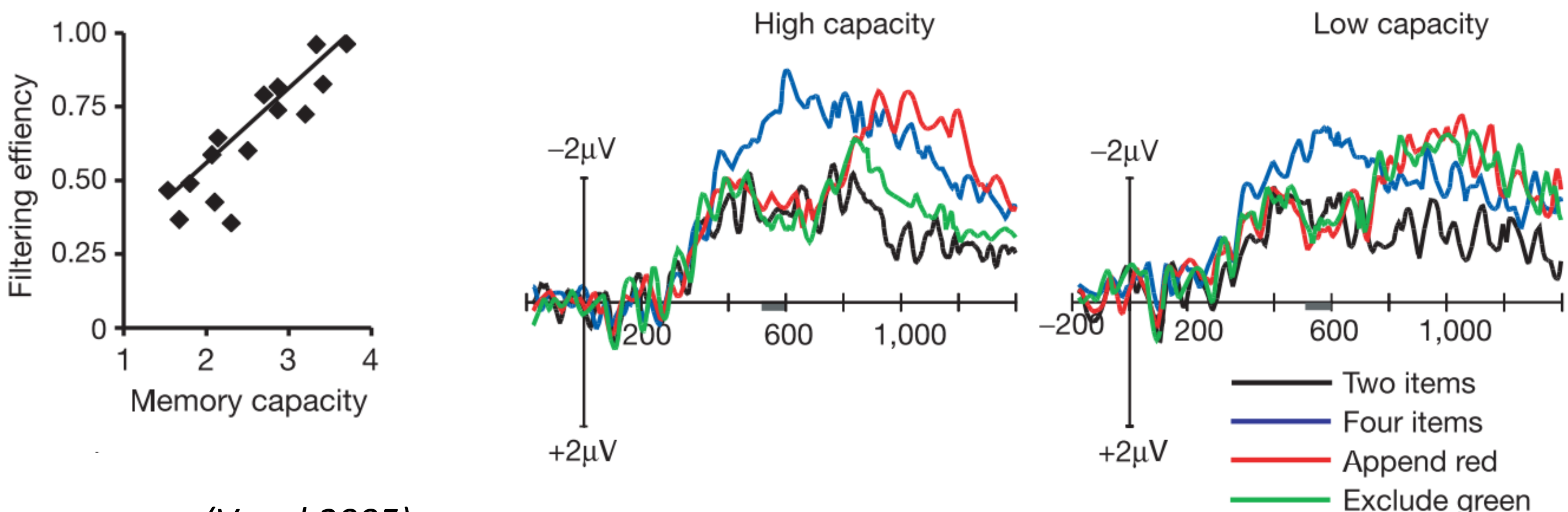
Catherine Wacongne

ICON 2014

- Working memory has a limited capacity *(Vogel 2004)*

➔ Necessity to manage working memory content

- linked to the capactity to filter out irrelevant items



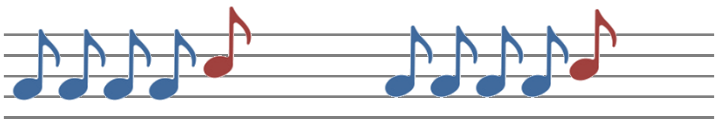
(Vogel 2005)

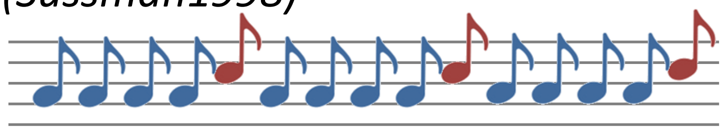
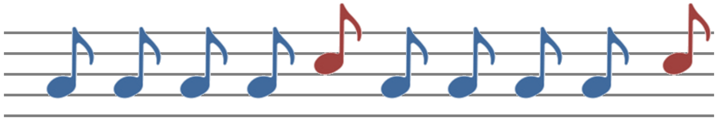
- Implication of the basal ganglia and valuation system (*Mc Nab 2008, O'Reilly & Frank 2006*)

Main question

What is working memory used for in the absence of an external task ?

Auditory processing :

(Bekinschtein 2009)

MMN to 

(Sussman1998)

no MMN

MMN to 

Serial reaction time tasks :

- 1st order transitions = independent of attention.
- higher order transitions is affected by attention demanding secondary tasks *(Curran and Keele 1993)*

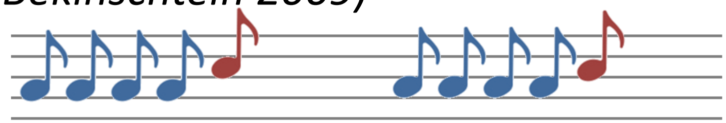
Evidence for predictive systems in the brain

Preattentive processing is strongly affected by temporal gaps

Attentive processing and working memory

Auditory processing :

(Bekinschtein 2009)



P300 to

The attention-related P300 is not affected by variable or long time intervals

Serial reaction time tasks :

Dependencies at distance up to 6 items can be learned (in several thousands of trials) if the subject is **attentive**, although the learning remains **implicit** *(Remillard, 2008, 2010)*

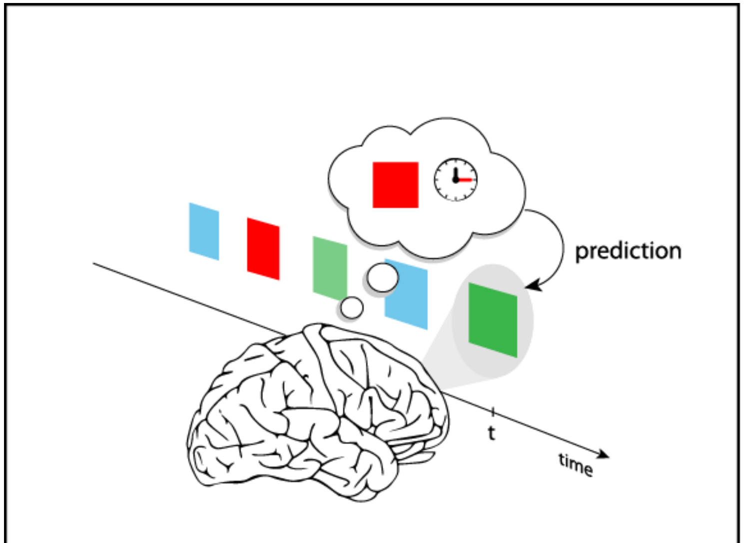
- ➔ Some responses to violation depend on attention and working memory
- ➔ Working memory can bridge temporal gaps
- ➔ Even in the absence of a task

What is working memory used for in the absence of an external task ?

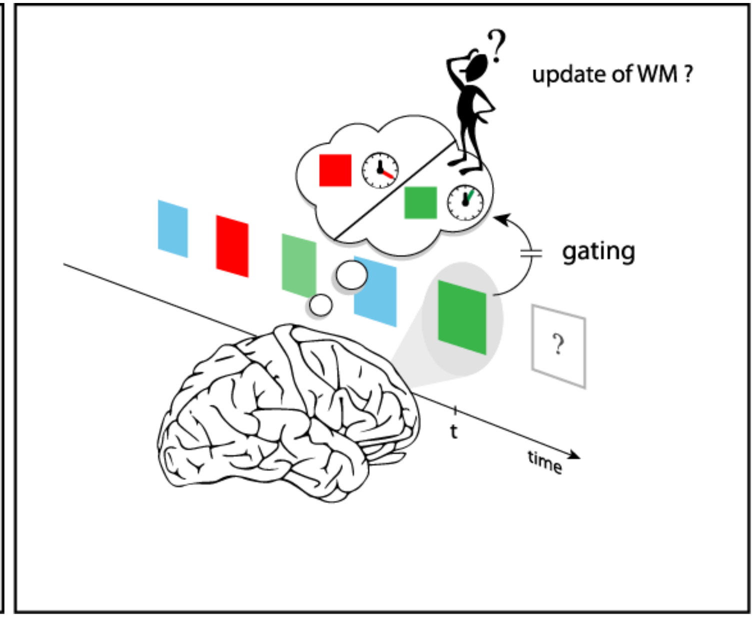
Predicting future events

How can we manage working memory content to predict future events ?

Computational Problem



Goal : predict the upcoming stimulus



"Rules":

- working memory can hold **one** item for arbitrary long time
- Past items that are not kept in working memory cannot be retrieved
- working memory slot encodes identity and number of stimuli since the current content was stored
- Probability estimate of the upcoming item is computed based on the learned P(stimulus|**Slot content**)

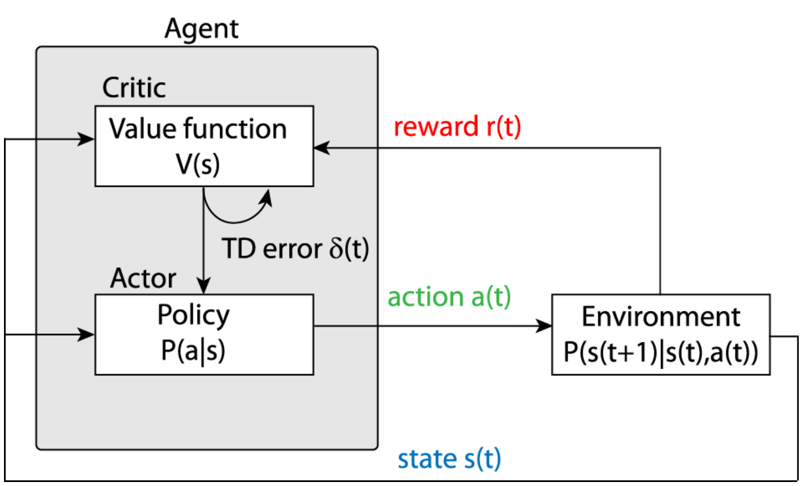
Hacking the reinforcement learning algorithm

Updating can be seen as an internal decision

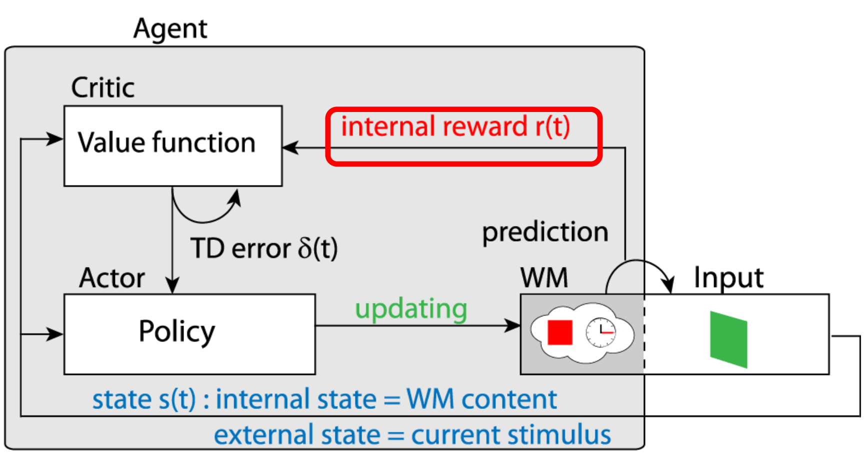
The implication of the value system implies the existence of a reward

Monkeys prefer informative options even when they do not change the objective reward => consistent with the idea of a rewarding value of predictive information *(Bromberg-Martin & Hikosaka, 2009)*

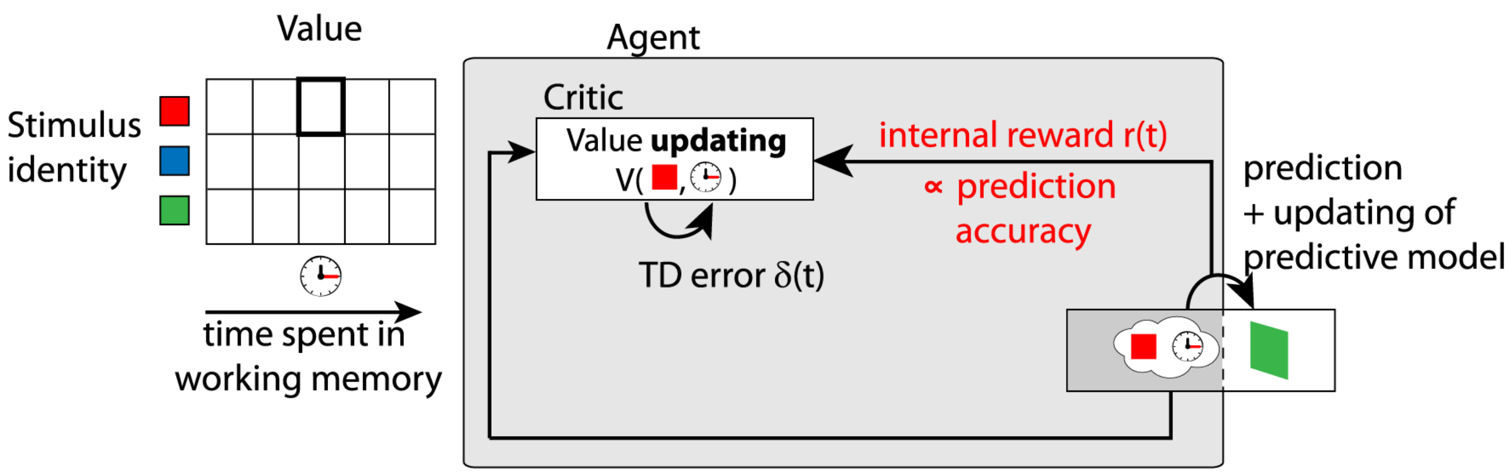
Classical reinforcement learning



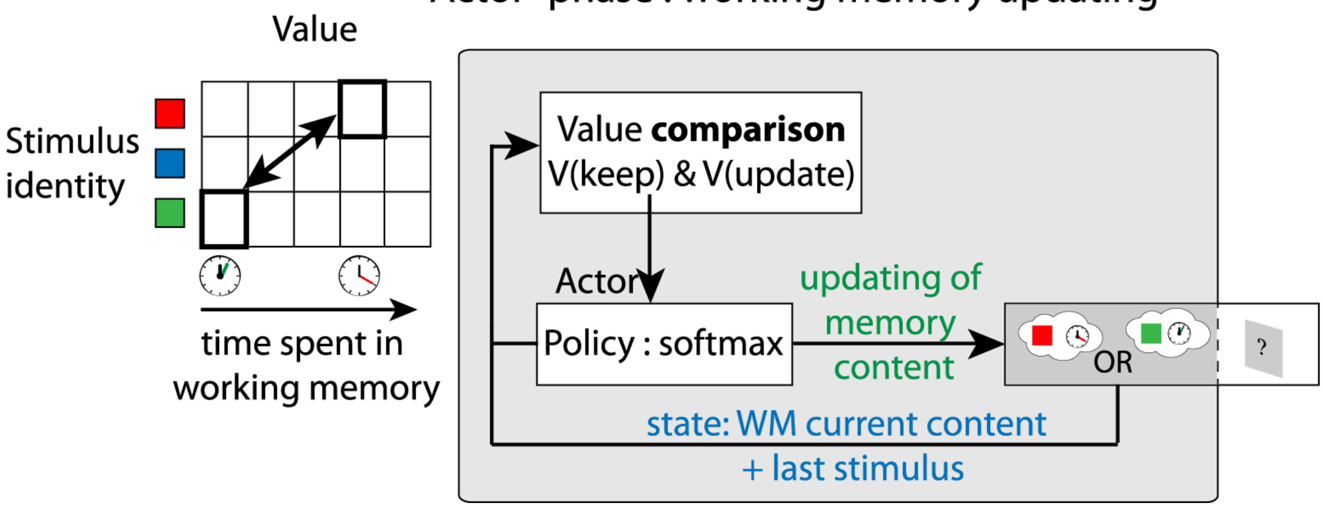
Working memory updating model



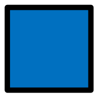
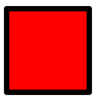
"Critic" phase : value updating



"Actor" phase : working memory updating

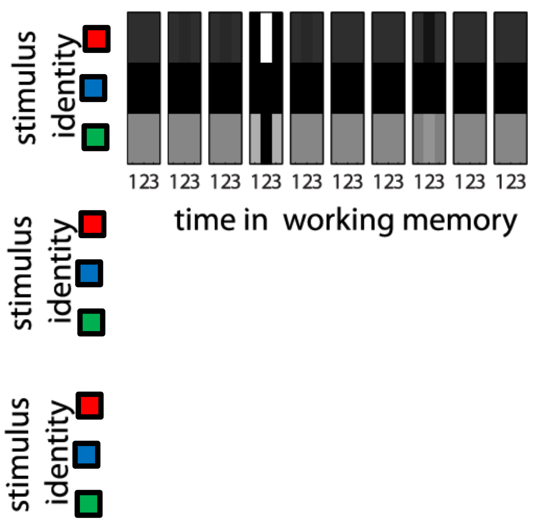


Discovery of long distance predictive relations

- 3 possible stimuli
-  predicts  four time steps later.

temporal distance 4

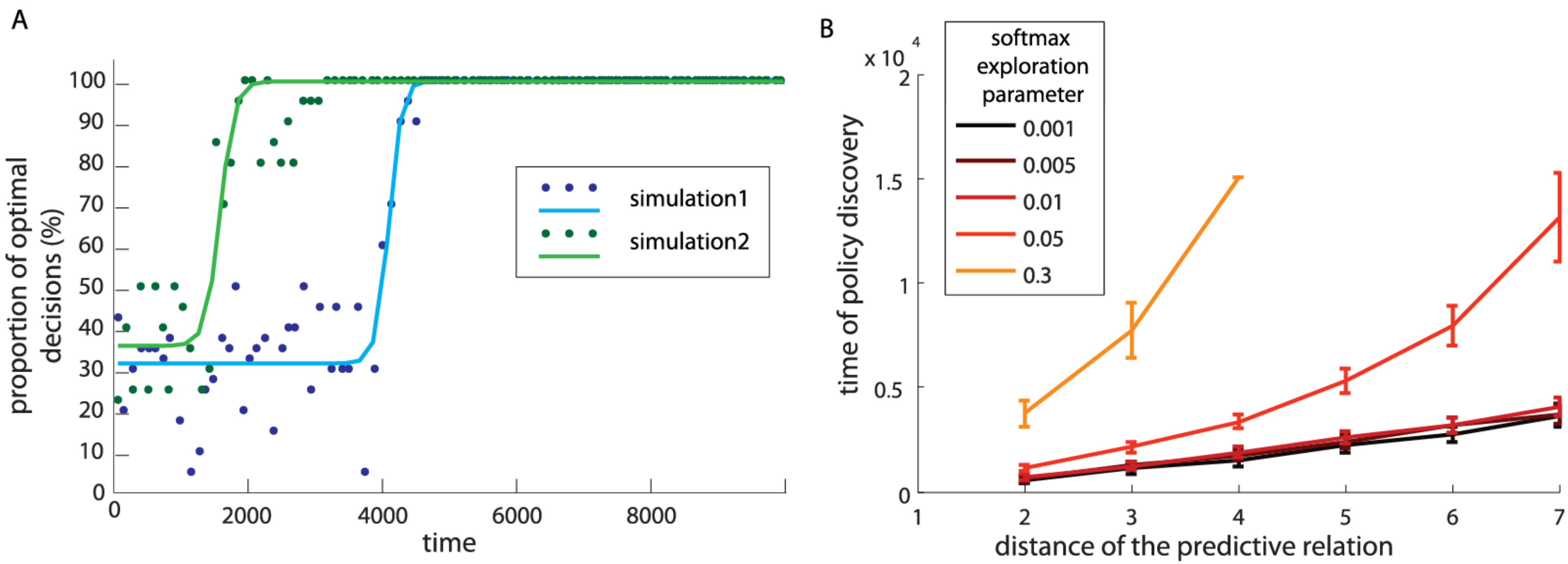
Real transition probability matrix



Discovery of long distance dependencies
Working memory dependent but not necessarily explicit

Strategy discovery

Optimal strategy: store the predictive item until it is not predictive any more

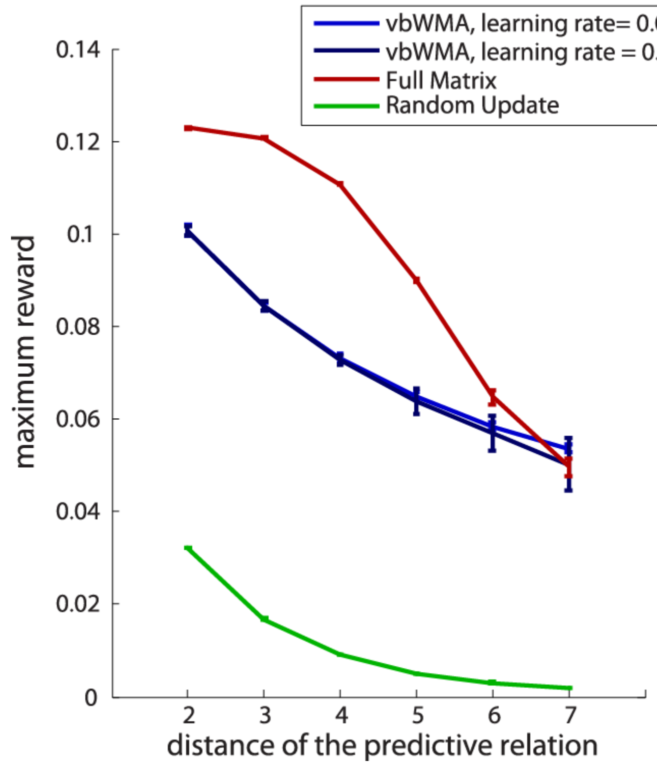


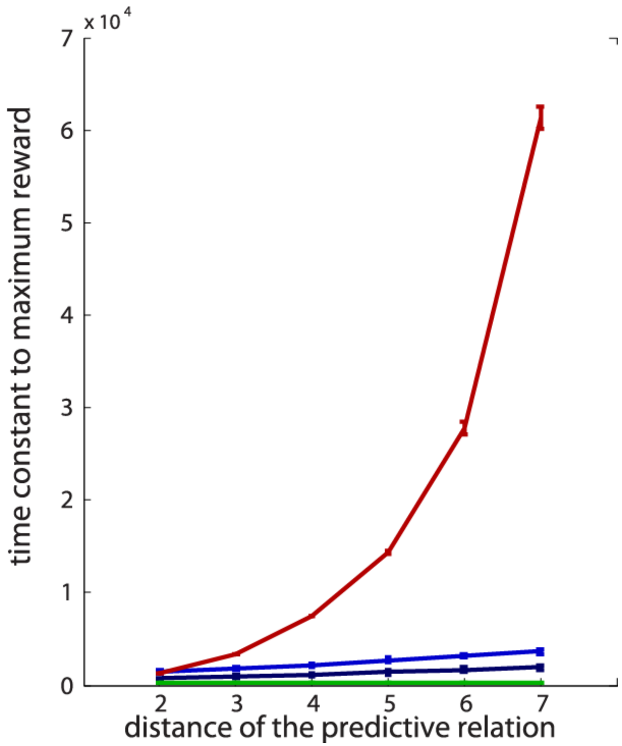
Sudden discovery of the policy
Stabilization of the policy before the stabilization of the conditional transition probability estimate

The advantage of starting small

Comparison of the predictive performances of the reinforcement learning model with 2 learning rates compared to

- A model that tracks the entire relevant history (last past 4 stimuli)
- A model that has one slot but updates at random



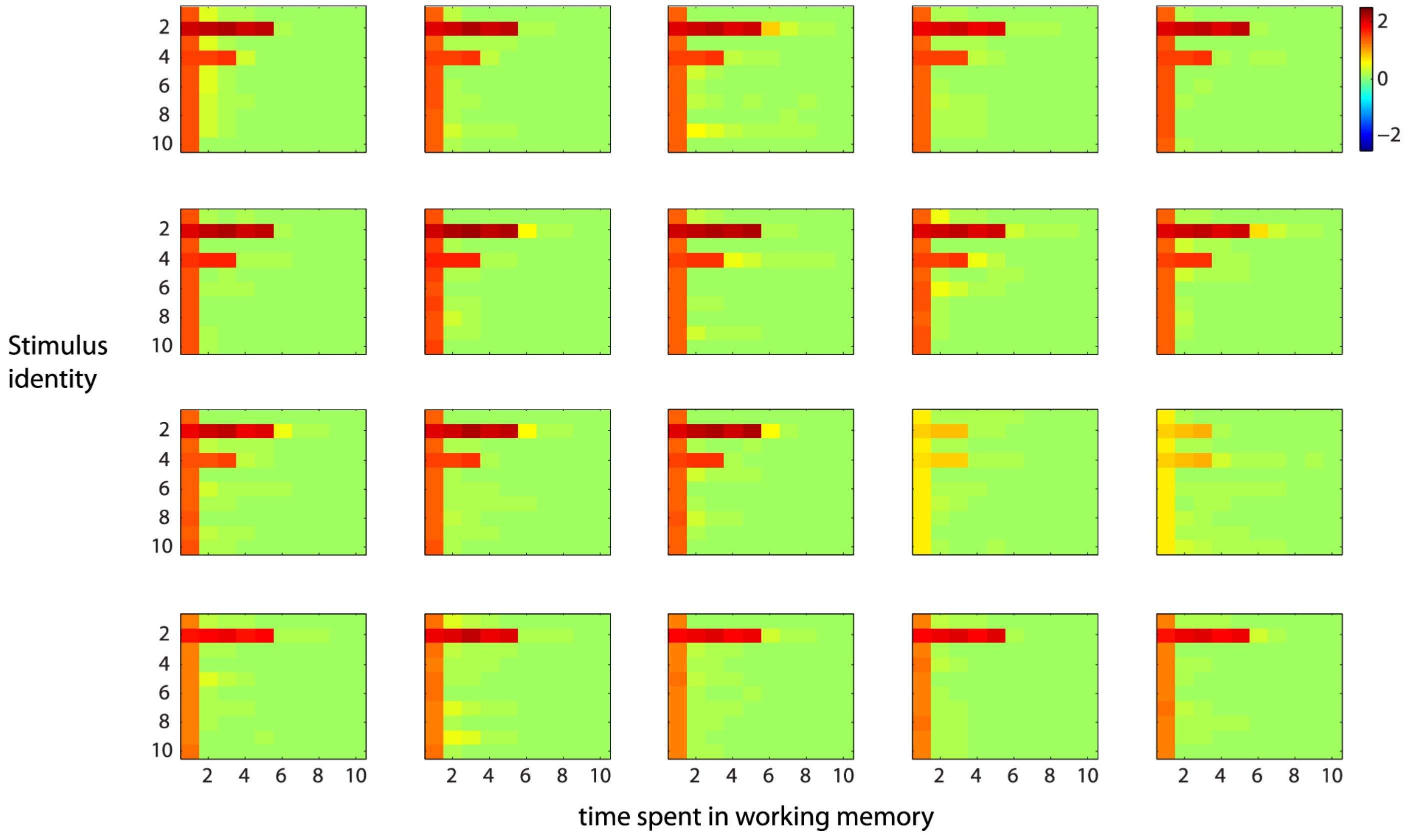


➔ Good tradeoff between final performance and speed of learning

Competition between temporal regularities

20 simulations with the same dependencies:

- "stim2" is predictive 3 and 5 time steps later
- "stim4" is predictive 3 time steps later



➔ Exploration/exploitation tradeoff

Conclusion

- **Predicting** accurately future events is one of the brain's **intrinsic** goals
- Limitations in working memory capacity imposes to **select** relevant information
- Working memory can be managed to achieve this goal in the absence of extrinsic rewards
- A reinforcement learning mechanism is suitable for the discovery and exploitation of long distance dependencies
- Having a small working capacity can be advantageous in terms of learning speed
- The model predicts impairment of long distance dependencies learning in patients with striatal dysfunction
- It also predicts a competition between external tasks, and intrinsic goals like prediction

Acknowlegments

- Stanislas Dehaene
- Jean-Pierre Changeux