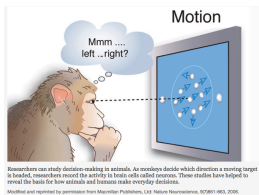


Attention, motivation, apprentissage

Eclairage de la recherche fondamentale



Driss Boussaoud

Directeur de Recherche au CNRS



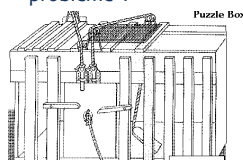
3. Apprentissage par renforcement

- Historique sur le renforcement
- Conditionnement, apprentissage
- Apprentissage associatif, basé sur le renforcement
- Apprentissage par imagerie mentale
- Apprentissage social

Apprentissage, conditionnement

The law of effect

- Le premier psychologue à étudier le conditionnement opérant.
- Son expérience la plus célèbre avec des chats placés dans la «boite-problème».



Thorndike, E. L. (1898). Animal intelligence.



Edward L. Thorndike
Williamsburg, Mass.
1874 - New York 1949

La « boîte-problème » de Thorndike

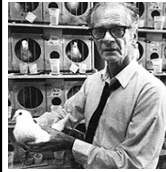
- Thorndike conclut que tout comportement qui a été suivi d'une **conséquence plaisante** tend à être reproduit, et inversement tout comportement suivi d'une **conséquence déplaisante** tend à être **inhibé**: loi de l'effet (law of effect)



Deux formes d'effet : renforcement et punition, bases du conditionnement

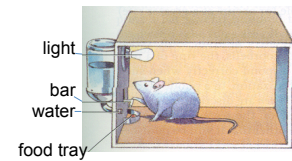
	Renforcement La probabilité de réponse augmente	Punition La probabilité de réponse diminue
Positif	Récompense	Stimulus aversif appliqué
Négatif	Stimulus aversif retiré	Récompense retirée

Burrhus Frederic SKINNER



Susquehanna, Penn. 1904,
Cambridge, Mass. 1990

Cage de Skinner



Conditionnement classique versus opérant

Classique
Réflexe

Récompense dispensée indépendamment du comportement du rat

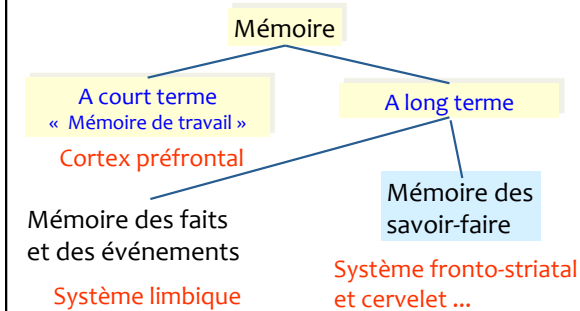
Opérant
Volontaire

Le comportement du rat détermine la récompense

De l'apprentissage à la mémoire

- Le conditionnement conduit à la mémorisation des relations entre stimuli ou entre stimuli et comportements;
- Le conditionnement opérant conduit à la formation de la mémoire procédurale, ou la mémoire des habitudes (habit memory);
- Une mémoire à long terme, dont le support neuronal est le cortex frontal, les ganglions de la base et les structures du lobe temporal médian;

Les différents types de mémoire



Habiletés motrices humaines



Le concept d'apprentissage

- L'apprentissage est un changement du comportement qui résulte de l'expérience passée;
- La capacité d'apprendre existe chez les espèces animales (animal learning theory) et peut être programmée dans des machines (machine learning);
- L'apprentissage vise à réduire l'écart entre la conséquence attendue d'un comportement et la conséquence produite;
- Il résulte de la plasticité neuronale, ou neuroplasticité.

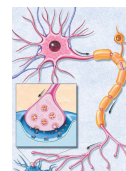
La neuroplasticité

- Elle existe même dans le cerveau adulte:

Les neurones peuvent modifier leurs propriétés individuelles et leurs interactions, par le biais de modifications synaptiques (loi de Hebb)

Ces modifications permettent d'apprendre et de former de nouveaux souvenirs, mais aussi la récupération fonctionnelle.

Hebb's law: "neurons that fire together, wire together... they form cell assemblies"



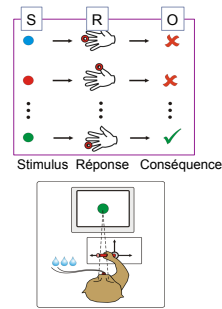
L'apprentissage, un champ multidisciplinaire

- Neurosciences, sciences cognitives, psychologie, médecine de la réhabilitation, sciences de l'éducation...
- L'intelligence artificielle, la robotique, sciences de l'ingénieur, l'informatique, mathématiques et physique;
- Il est basé sur le **renforcement**.



Principaux types d'apprentissage

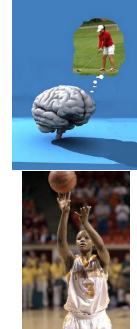
Expérience individuelle



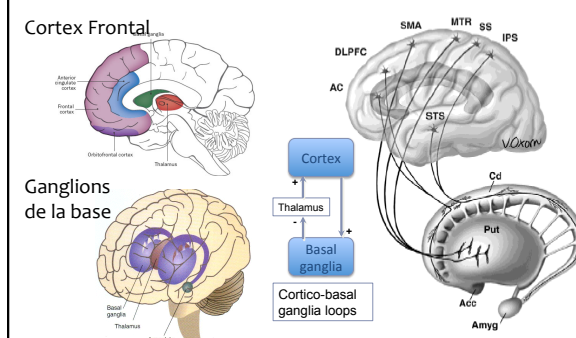
Apprentissage social



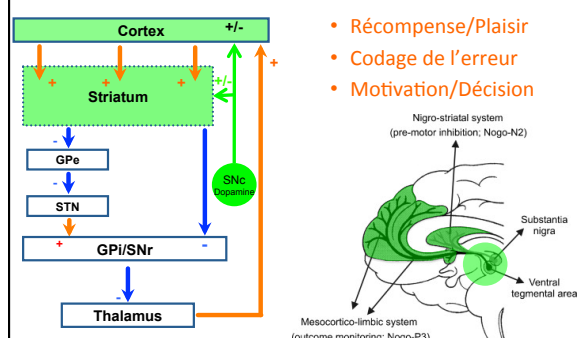
Apprentissage par imagerie mentale



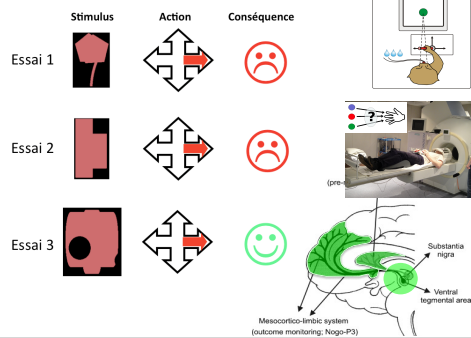
Circuits de l'apprentissage par renforcement



Neuroplasticité & circuits de la dopamine

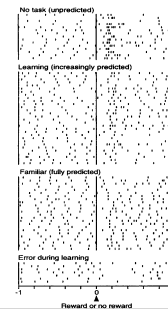


Neuroplasticité & circuits de la dopamine Apprentissage associatif



Dopamine codage de l'erreur de prédiction

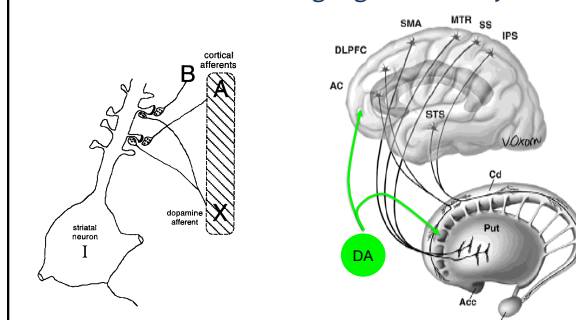
L'erreur de prédiction est l'écart entre la conséquence attendue et la conséquence produite par le comportement.



- La décharge des neurones dopaminergiques déchargent d'autant plus fortement que l'erreur de prédiction est grande;
- Ils projettent sur le striatum et le cortex frontal, dont ils modifient l'activité pour adapter le comportement, de sorte à réduire l'erreur de prédiction.

Wolfram Schultz and Anthony Dickinson (2000). NEURONAL CODING OF PREDICTION ERRORS. Annu. Rev. Neurosci. 2000. 23:473-500.

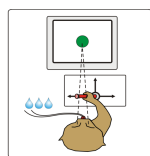
Dopamine modifies synaptic efficacy (Hebb) in frontal cortex and basal ganglia - Memory



A retenir:

- Les neurones dopaminergiques jouent un rôle central dans l'apprentissage par renforcement;
- Ils déchargent en relation avec la récompense (**renforcement**) et tout événement de l'environnement qui la prédit (**motivation**);
- La DA modifie la transmission synaptique dans les régions cibles (cortex frontal cortex et ganglions de la base) → **mémorisation**.

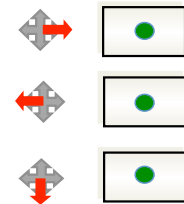
Etude de l'apprentissage par renforcement chez le singe



Stimulus

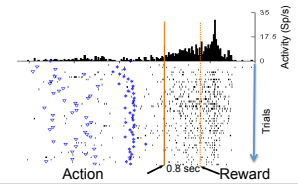
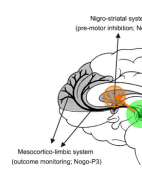
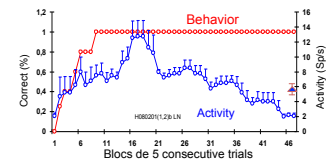
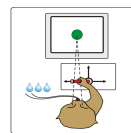


Action Conséquence

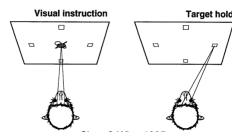


Hadj-Bouziane & Boussaoud, 2003

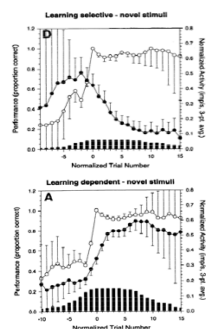
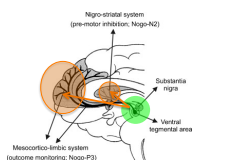
Modifications neuronales dans le striatum



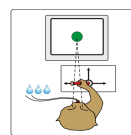
Modifications neuronales dans le cortex frontal



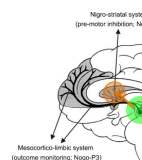
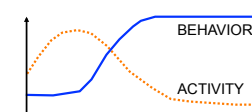
Chen & Wise 1995.



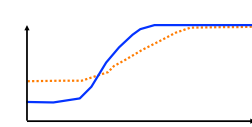
Deux catégories de neurones

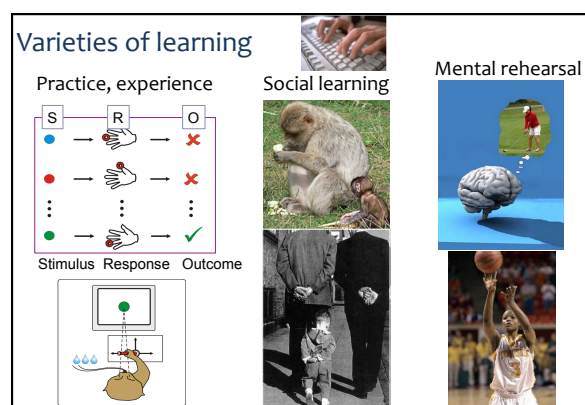
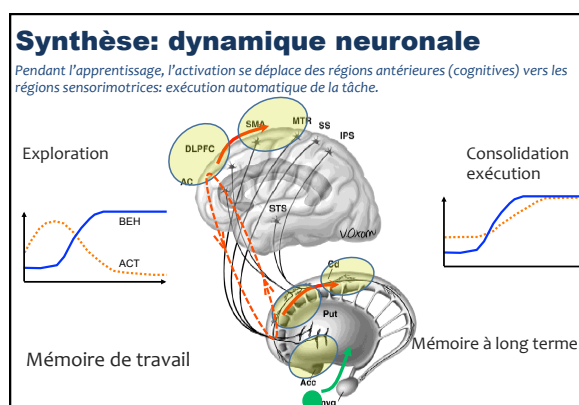
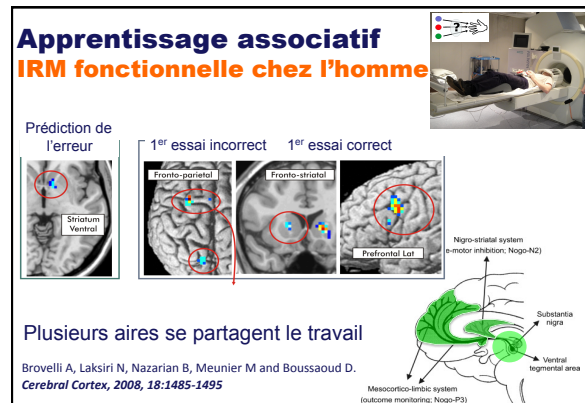
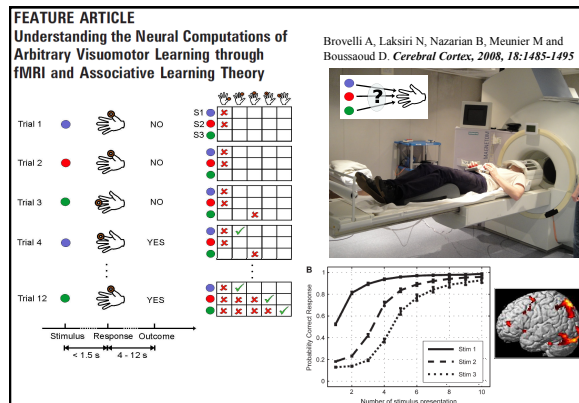


Les neurones de l'exploration



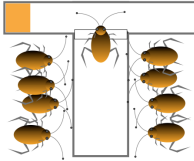
Les neurones de la consolidation





Apprentissage social

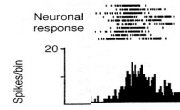
Psychologie sociale



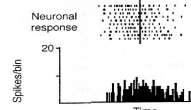
Des blattes à l'homme, le comportement est régulé par le contexte social.

La découverte des neurones miroirs

Monkey does



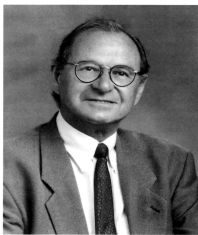
Monkey sees doing



Rizzolatti et al., 1992.

La facilitation sociale

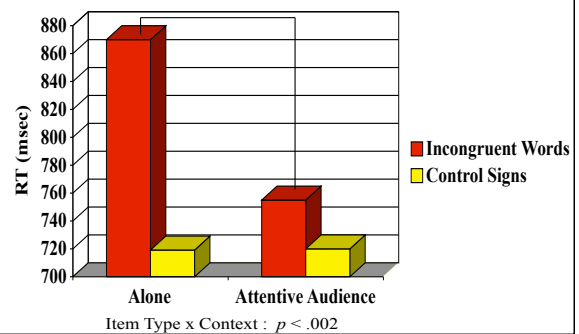
Un phénomène ubiquitaire



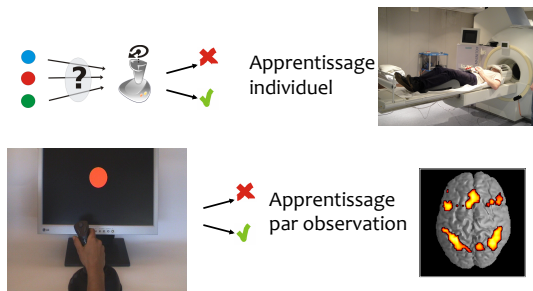
- La présence des congénères facilite le comportement de la blatte pour la réponse dominante, automatique (tâche facile).
- Elle ralentit la réponse non dominante (tâche difficile)

Conclusion de Zajonc (1965)

Effet de l'audience sur le stroop

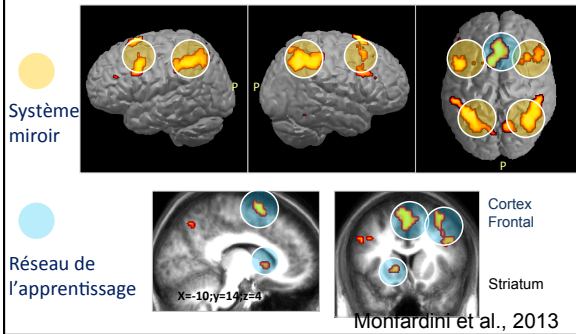


Quel réseau cérébral est impliqué?

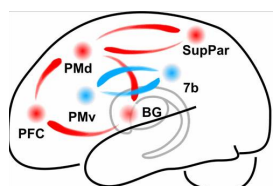


Monfardini et al., 2013

Même réseau que l'apprentissage par renforcement Avec le système miroir



Apprentissage social & individuel, même réseaux cérébraux

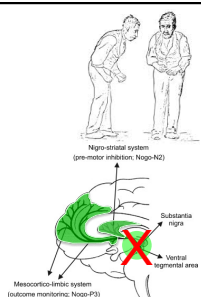
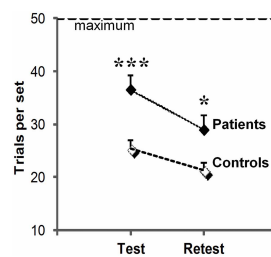


L'apprentissage social active deux systèmes:

- Le système miroir, pour comprendre
- Le système du renforcement, pour apprendre

Sous forte modulation sociale: je me compare à vous en permanence !

Apprentissage sans dopamine?



Sans dopamine, l'apprentissage est fortement altéré.

Hadj-Bouziane, Meunier, Boussaoud, 2013

La dopamine mise en équation

Formule à 3 variables:

S environnement (stimulus)

A actions (réponse)

R conséquence (récompense)

Adaptez A à S pour optimiser R

$$R = r_0 + r_1 + r_2 \dots r_n$$

Modèle de Rescorla-Wagner (1972): $V(t+1) = V(t) + \alpha(\beta - V(t))$

α vitesse d'apprentissage, β la force d'association $S \rightarrow A$ mesurée par R

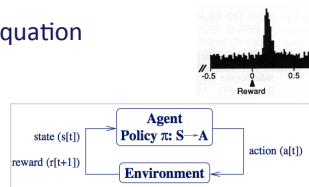
($\beta - V$) représente l'erreur de prédiction (EP): R observée - R attendue

$$EP(t) = 1 - 0.5 = 0.5 \rightarrow R \text{ inattendue}$$

$$EP(t) = 1 - 1 = 0 \rightarrow R \text{ attendue}$$

$$EP(t) = 0 - 0.5 = -0.5 \rightarrow R \text{ omise}$$

Applications: Robotique, ascenseurs, télécom, jeu d'échec ...



Modèle d'apprentissage par renforcement

Todd Hester, Michael Quinlan, and Peter Stone
Department of Computer Science
The University of Texas at Austin
Austin, TX 78712



Generalized Model Learning for Reinforcement Learning on a Humanoid Robot

Interactions entre neurotransmetteurs

Dopaminergic-neuropeptide interactions in the social brain

David H. Skuse* and Louise Gallagher*

Skuse & Gallagher, 2008
Trends in Cognitive Sciences Vol.13 No.1

