

*CERNA - Apprentissage et Intelligence Artificielle:
les vraies questions éthiques*

Apprentissages multiples, substrats neuronaux et modèles

Benoît Girard
ISIR
CNRS / UPMC



Les Apprentissages

- ❖ 3 grandes classes :
 - ❖ apprentissage supervisé
 - ❖ apprentissage non-supervisé
 - ❖ apprentissage par renforcement

Les Apprentissages

- ❖ 3 grandes classes :
 - ❖ apprentissage supervisé
 - ❖ apprentissage non-supervisé
 - ❖ apprentissage par renforcement
- ❖ Pour chaque action, un signal d'erreur est produit indiquant comment se corriger.

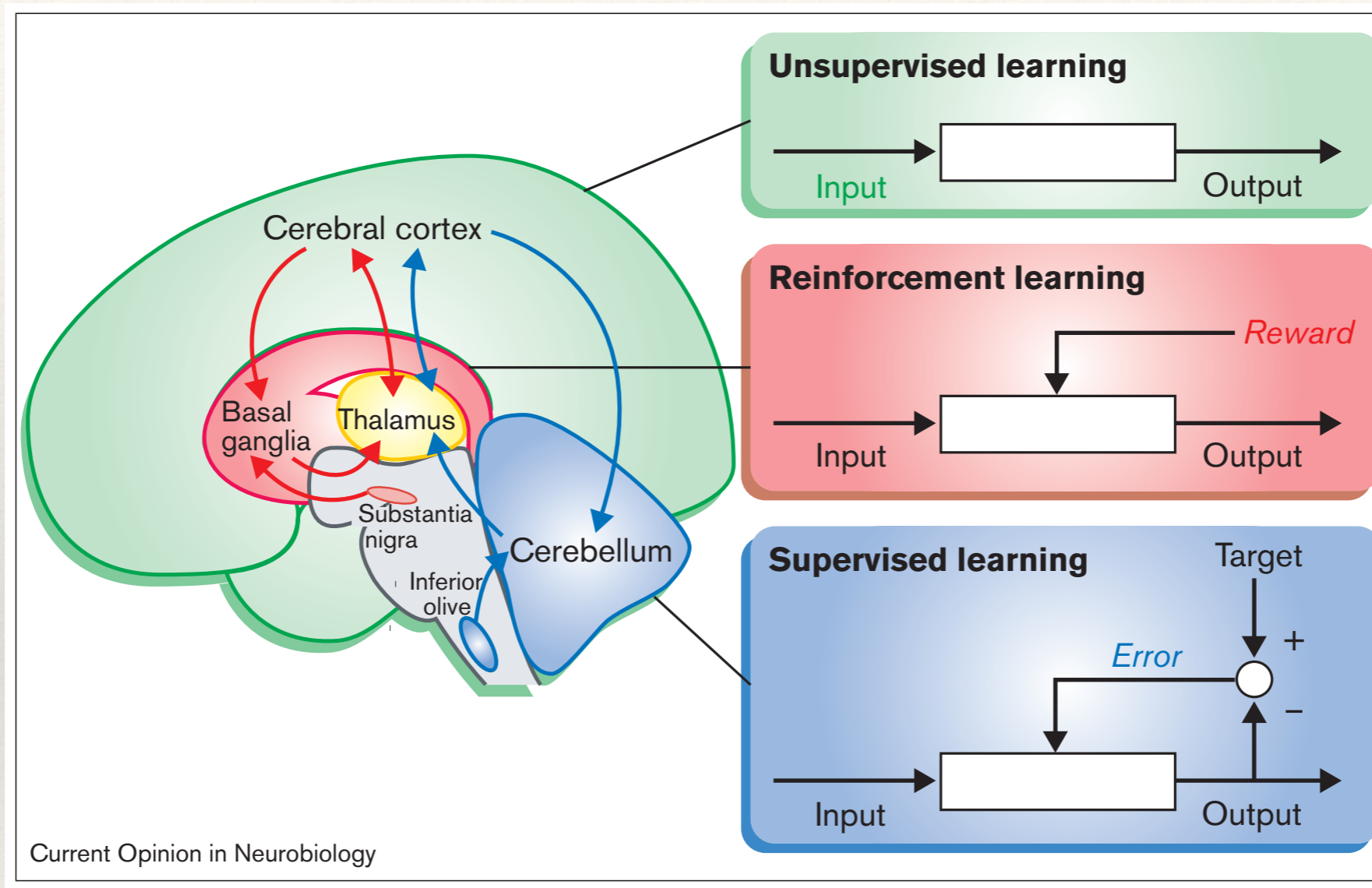
Les Apprentissages

- ❖ 3 grandes classes :
 - ❖ apprentissage supervisé
 - ❖ apprentissage non-supervisé
 - ❖ apprentissage par renforcement
- ❖ Aucun signal autre que sensori-moteur,
- ❖ Identification des régularités statistiques du flux sensori-moteur

Les Apprentissages

- ❖ 3 grandes classes :
 - ❖ apprentissage supervisé
 - ❖ apprentissage non-supervisé
 - ❖ apprentissage par renforcement
- ❖ Un signal de récompense / punition est disponible occasionnellement,
- ❖ Longues séquences d'actions sans information,
- ❖ Pb : attribution des responsabilités dans les résultats obtenus

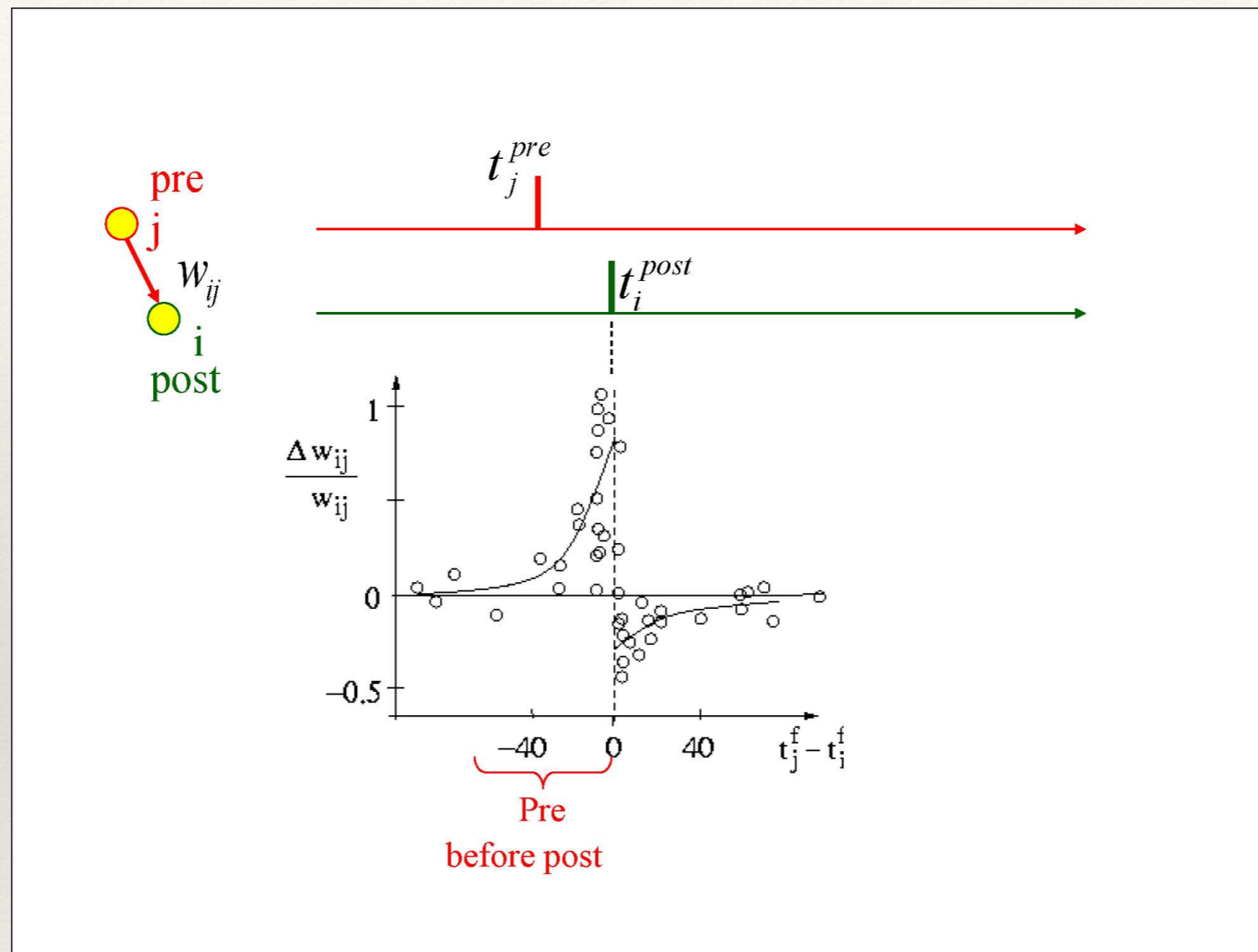
Les apprentissages & le cerveau



❖ Substrats distincts des apprentissages

(Doya, 2000), mais (Ohmae & Medina, 2015) : cervelet & renforcement

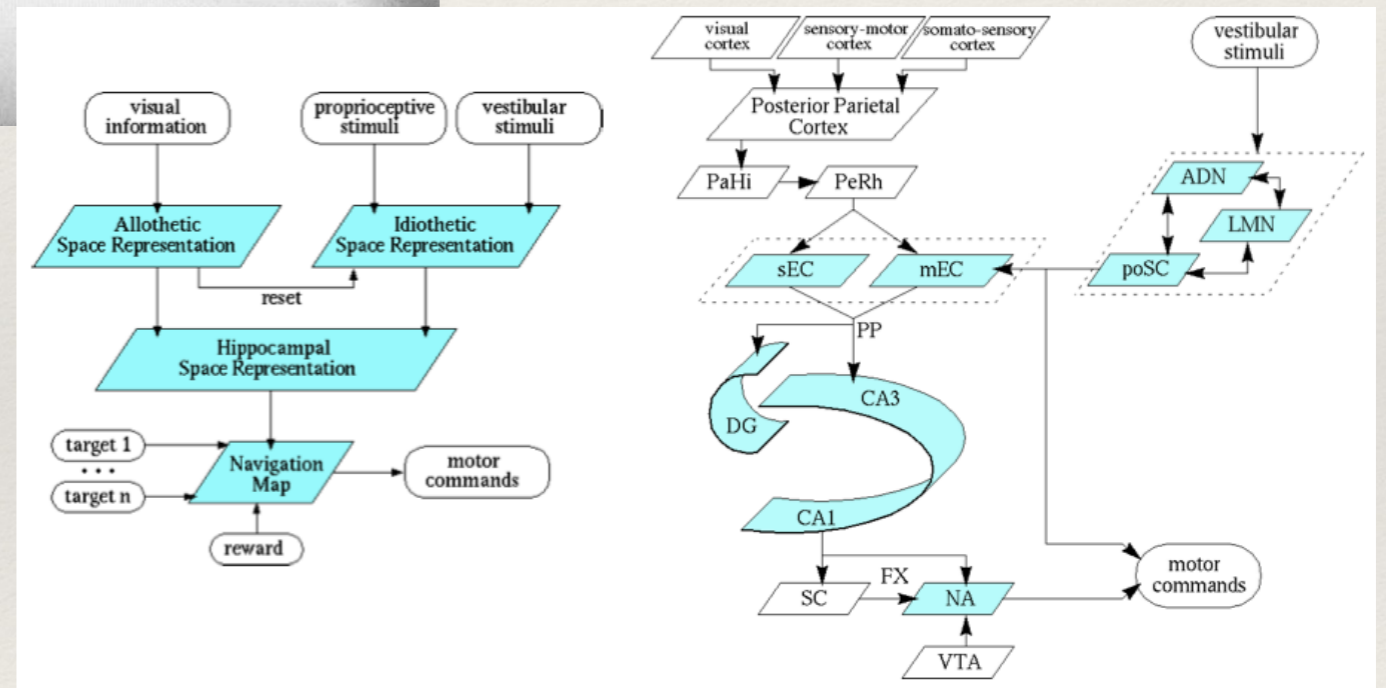
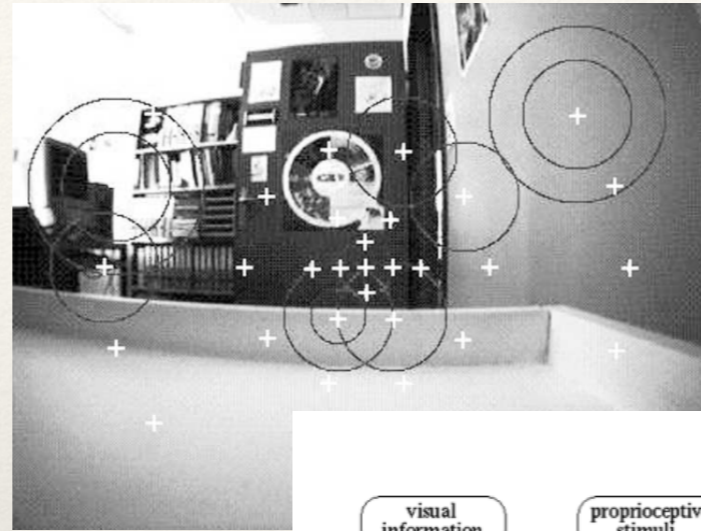
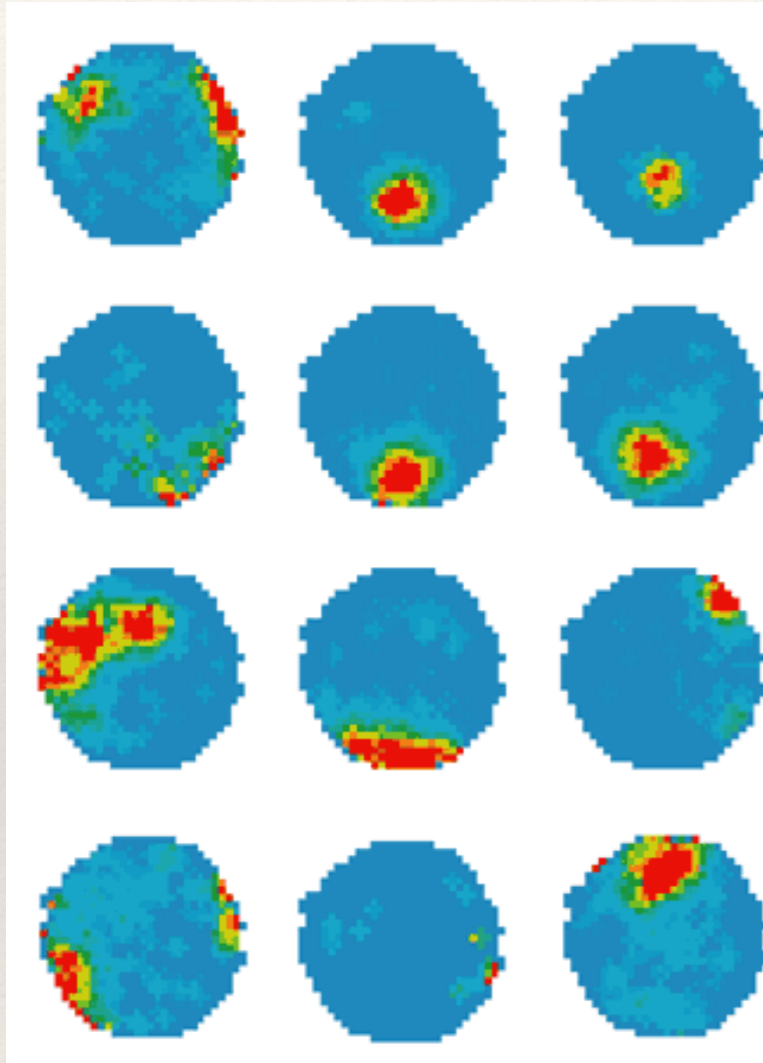
Apprentissage non-supervisé



- ❖ Apprentissage Hebbien,
- ❖ Plasticité dépendant du temps d'occurrence des impulsions (STDP)

(Hebb, 1949 ; Bi & Poo, 1998)

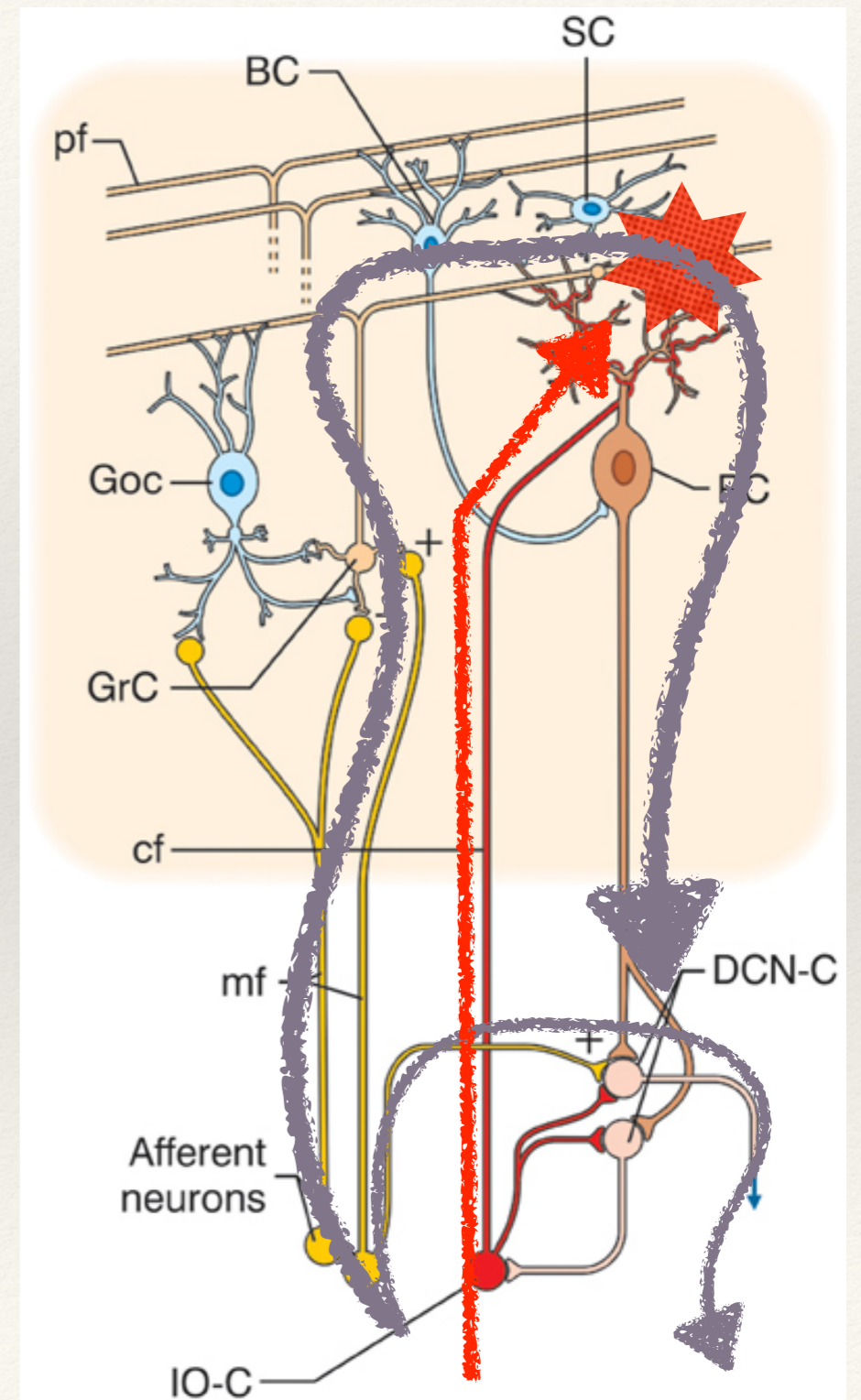
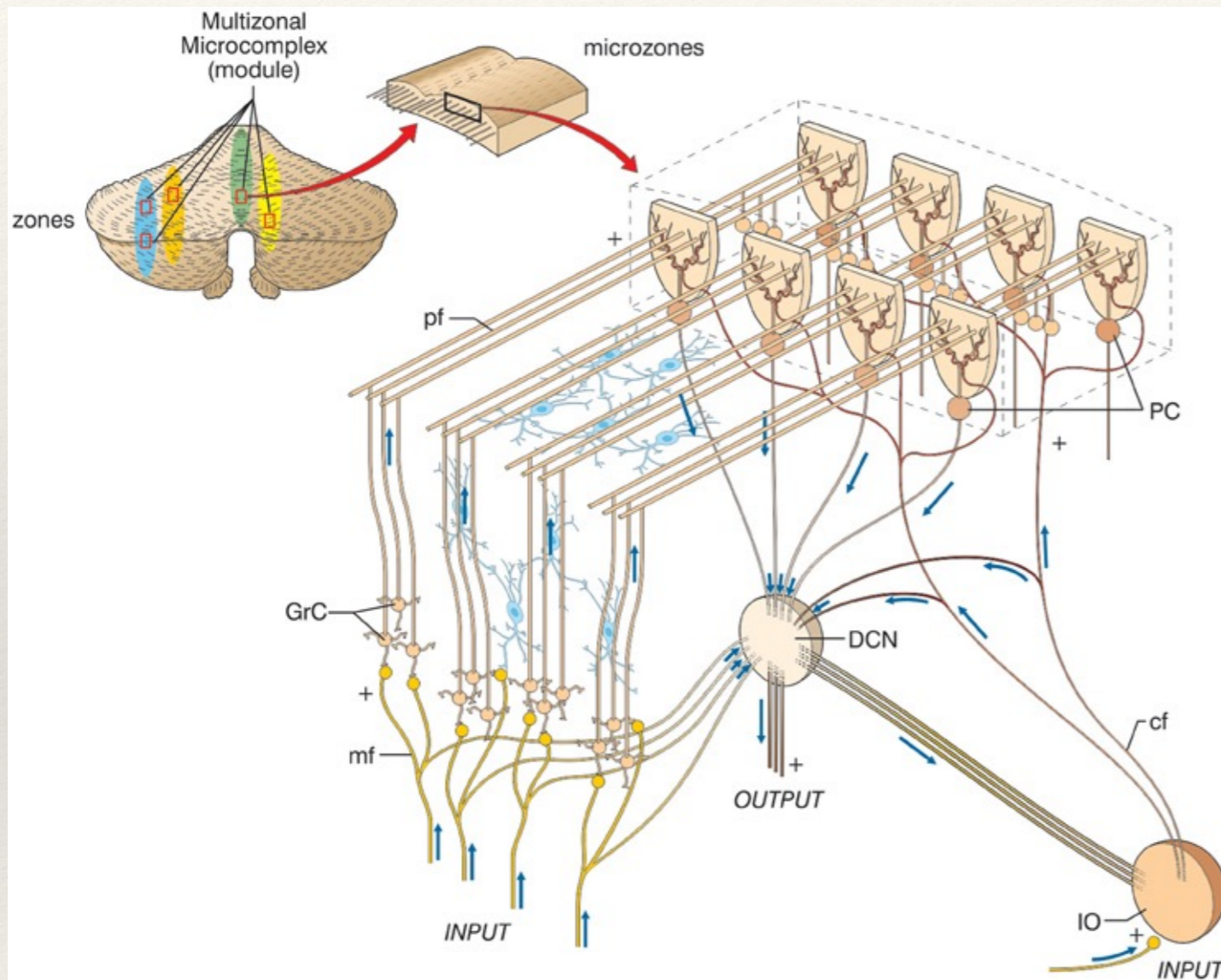
Apprentissage non-supervisé



- ❖ Constituant essentiel de l'apprentissage des cellules de lieux
- ❖ Prix Nobel physiologie ou médecine 2014 (O'Keefe, Moser & Moser)

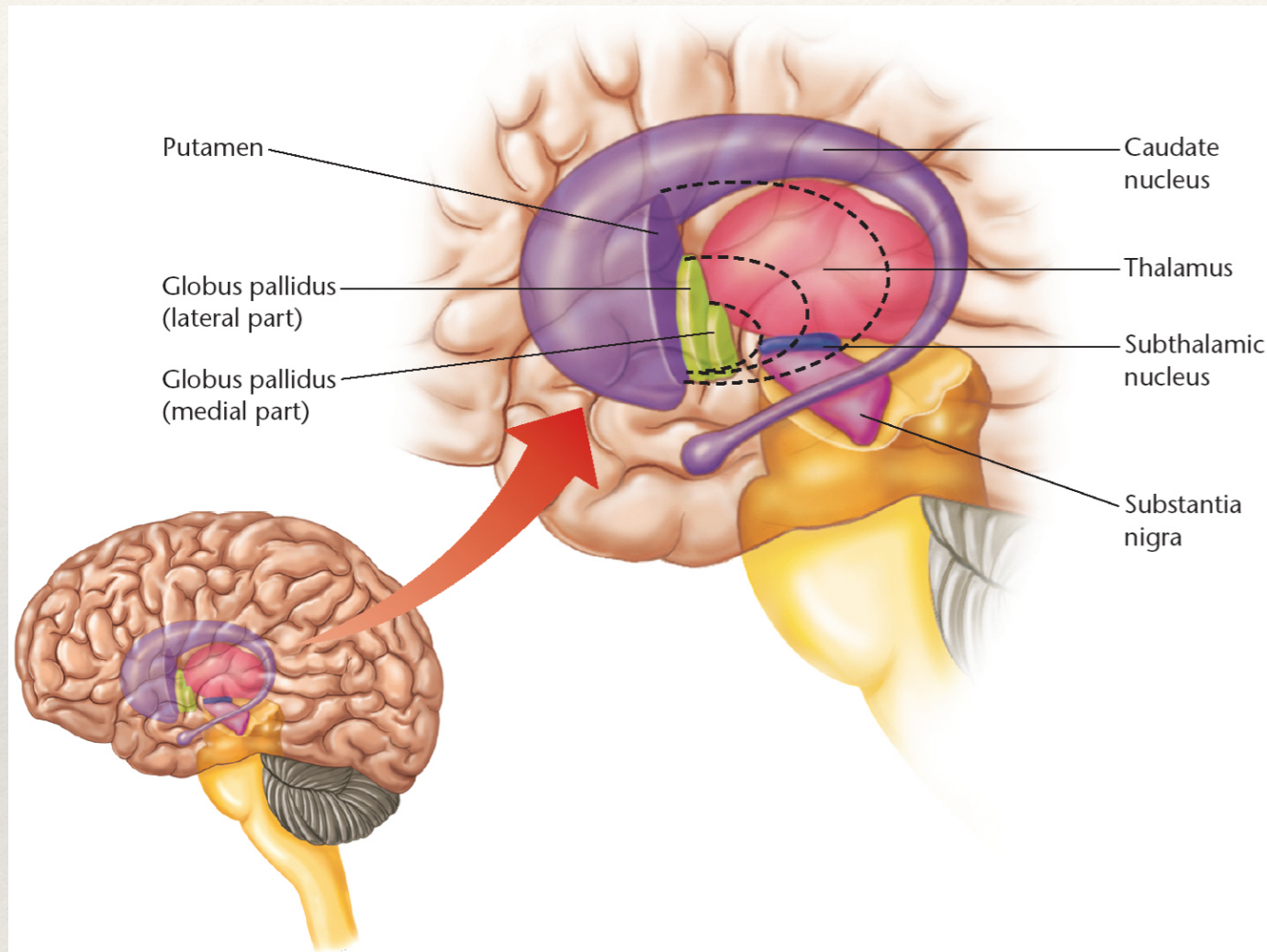
Modèle : (Arleo et al., 2000)

Apprentissage supervisé



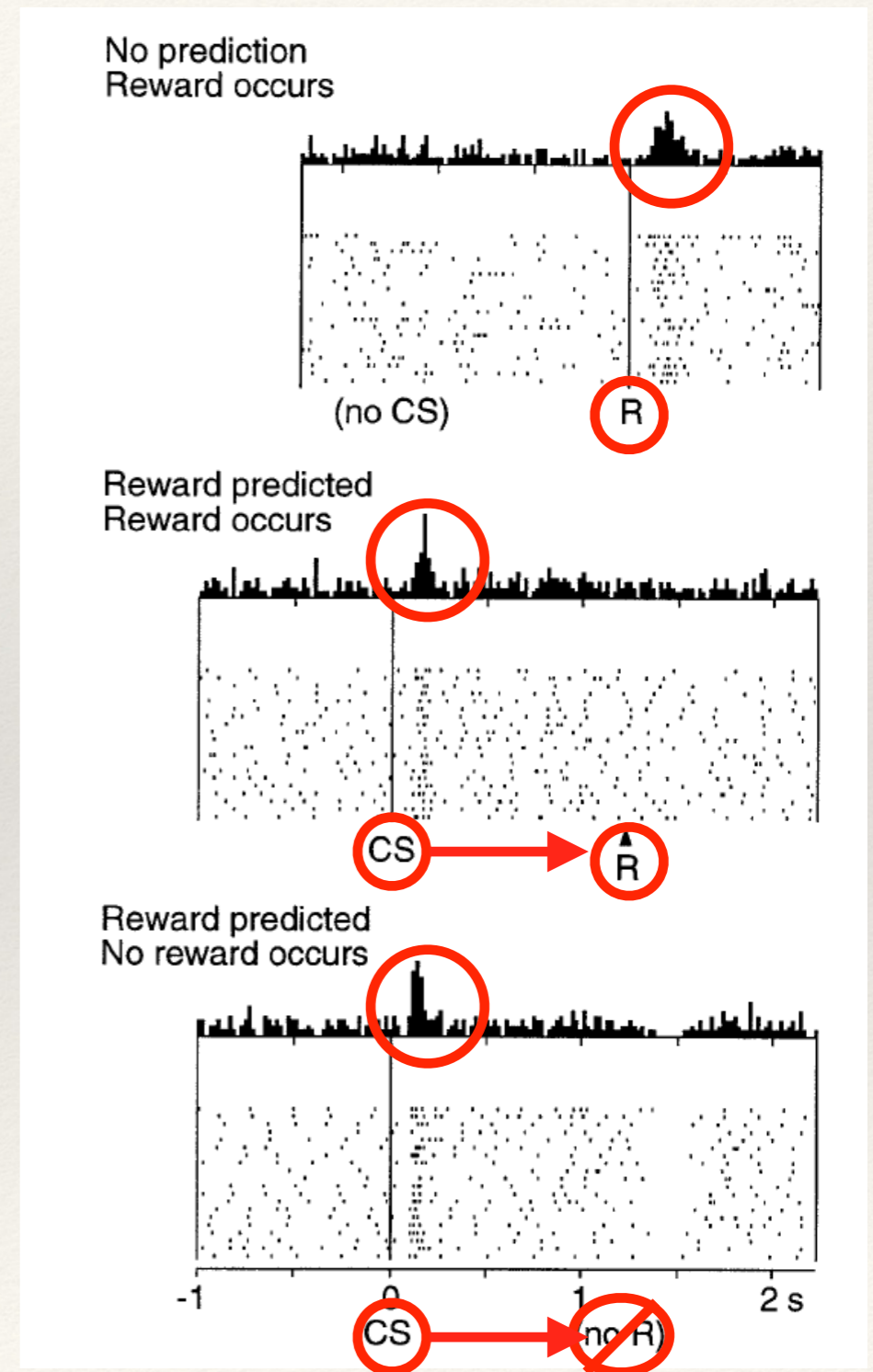
- ❖ Perceptron !
 - ❖ Cellules en grains (50M, 3/4 des neurones du cerveau !)
- (Marr, 1969)

Apprentissage par renforcement



❖ Dopamine et erreur de prédiction de récompense

(Schultz et al., 1997)



Apprentissage instrumental

- ❖ Comportement dirigé vers un but :
 - ❖ apprentissage modéré
 - ❖ hésitations (*vicarious trials & errors*)
 - ❖ temps de réactions longs
 - ❖ sensible à la dévaluation
- ❖ Comportement habituel :
 - ❖ apprentissage prolongé
 - ❖ réponses rapides,
 - ❖ stéréotypées
 - ❖ peu sensible à la dévaluation

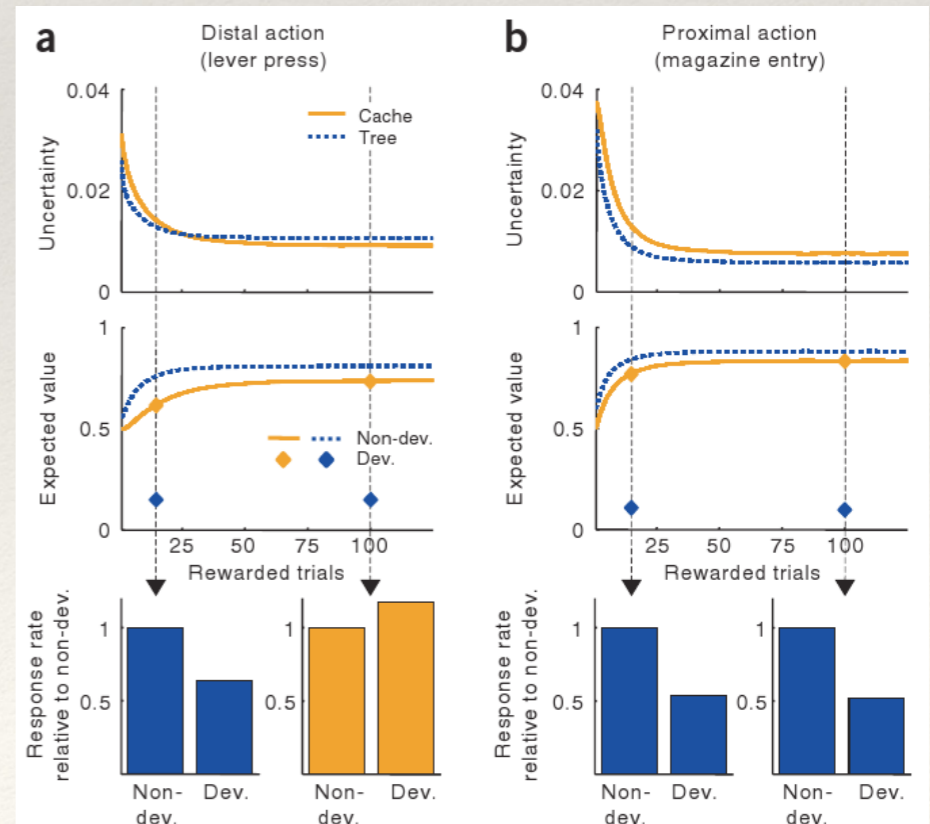
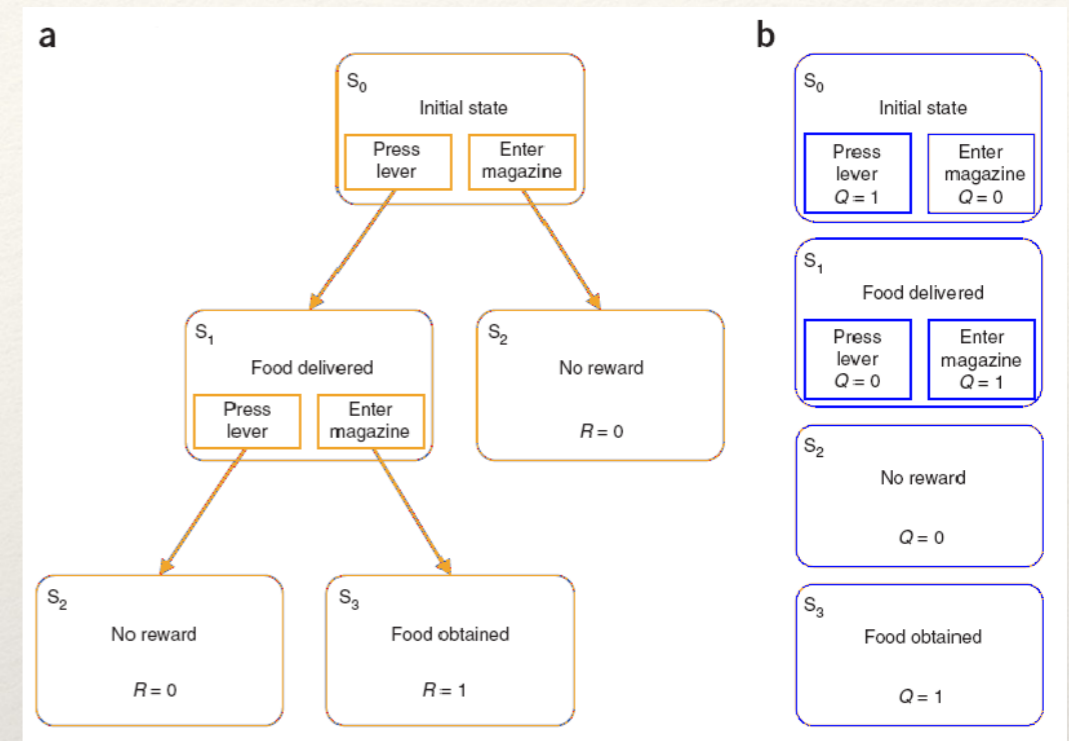
(Dickinson, 1985)

Modèles

- ❖ Théorie dominante :
 - ❖ Cpt dirigé vers un but = apprentissage par renforcement avec modèle du monde
 - ❖ Habitude = apprentissage par renforcement sans modèle (lien avec Dopamine)

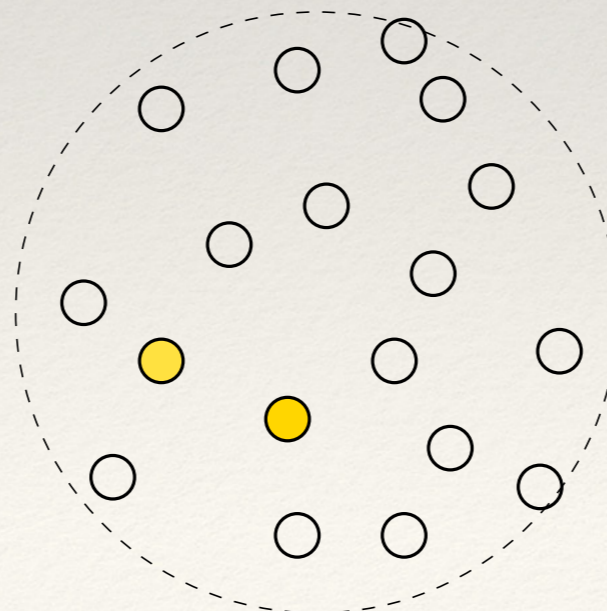
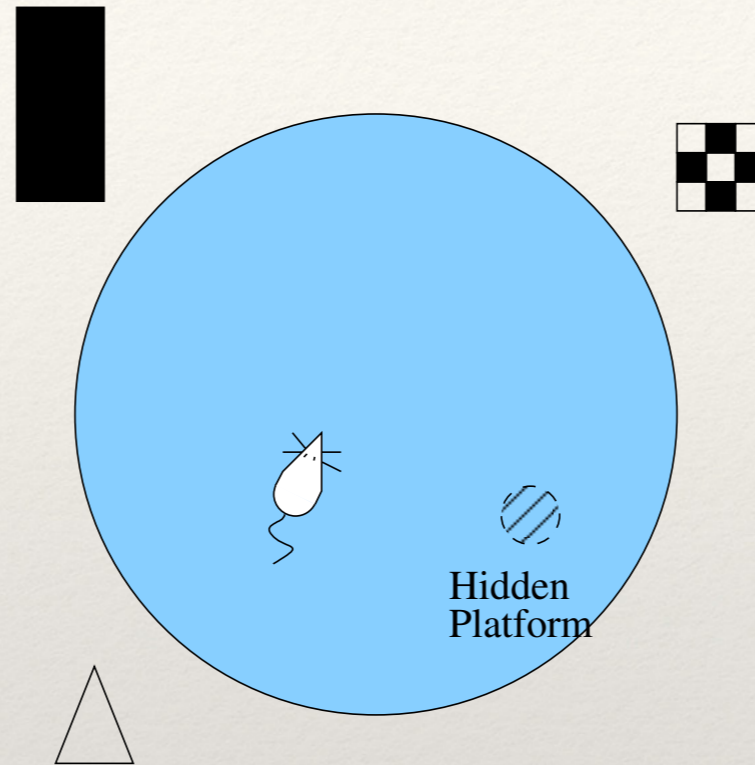
(Daw et al., 2005),

mais (Dezfouli & Balleine, 2012)



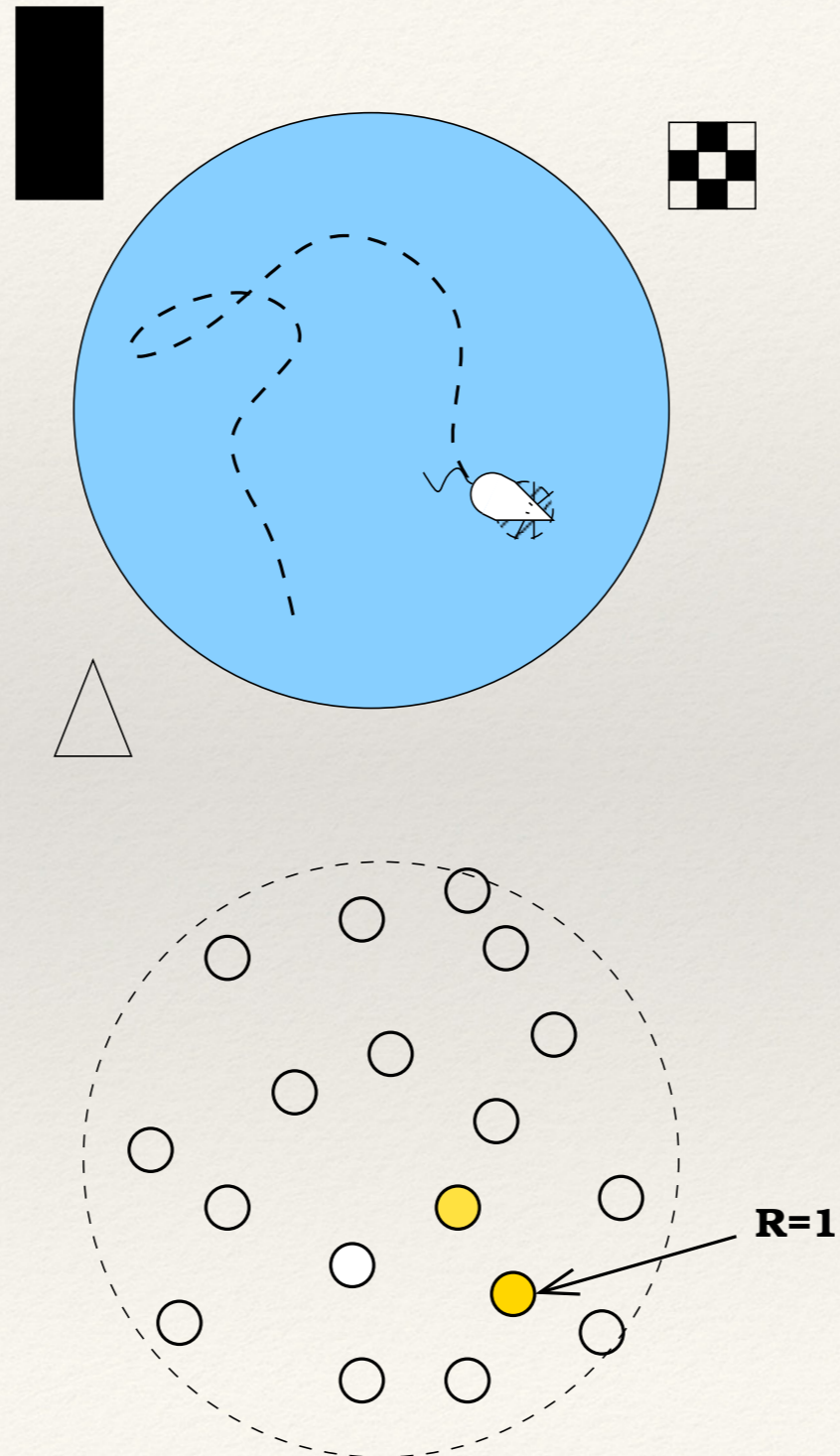
Apprentissage par renforcement sans modèle

Distal Cues



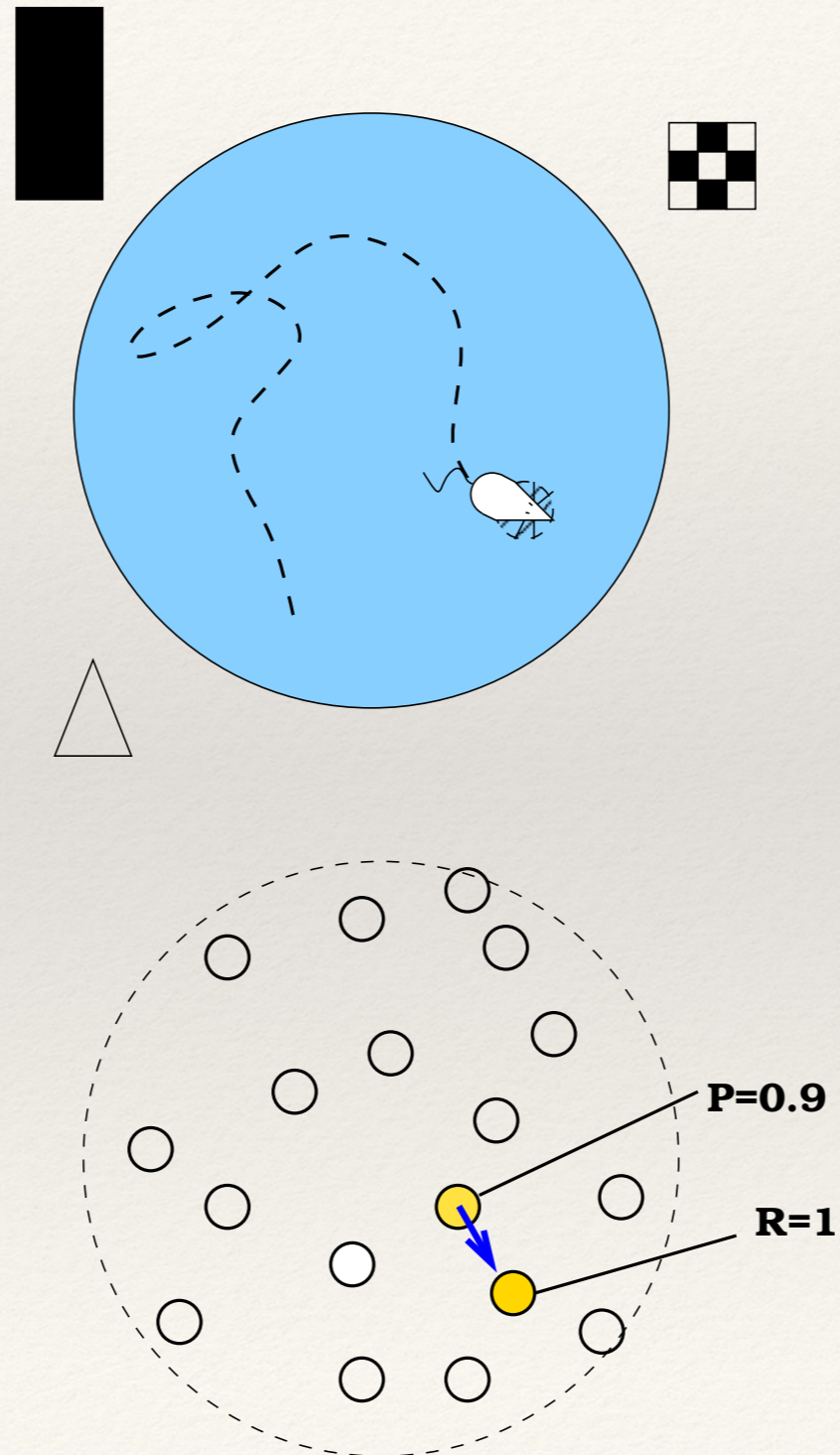
Apprentissage par renforcement sans modèle

Distal Cues



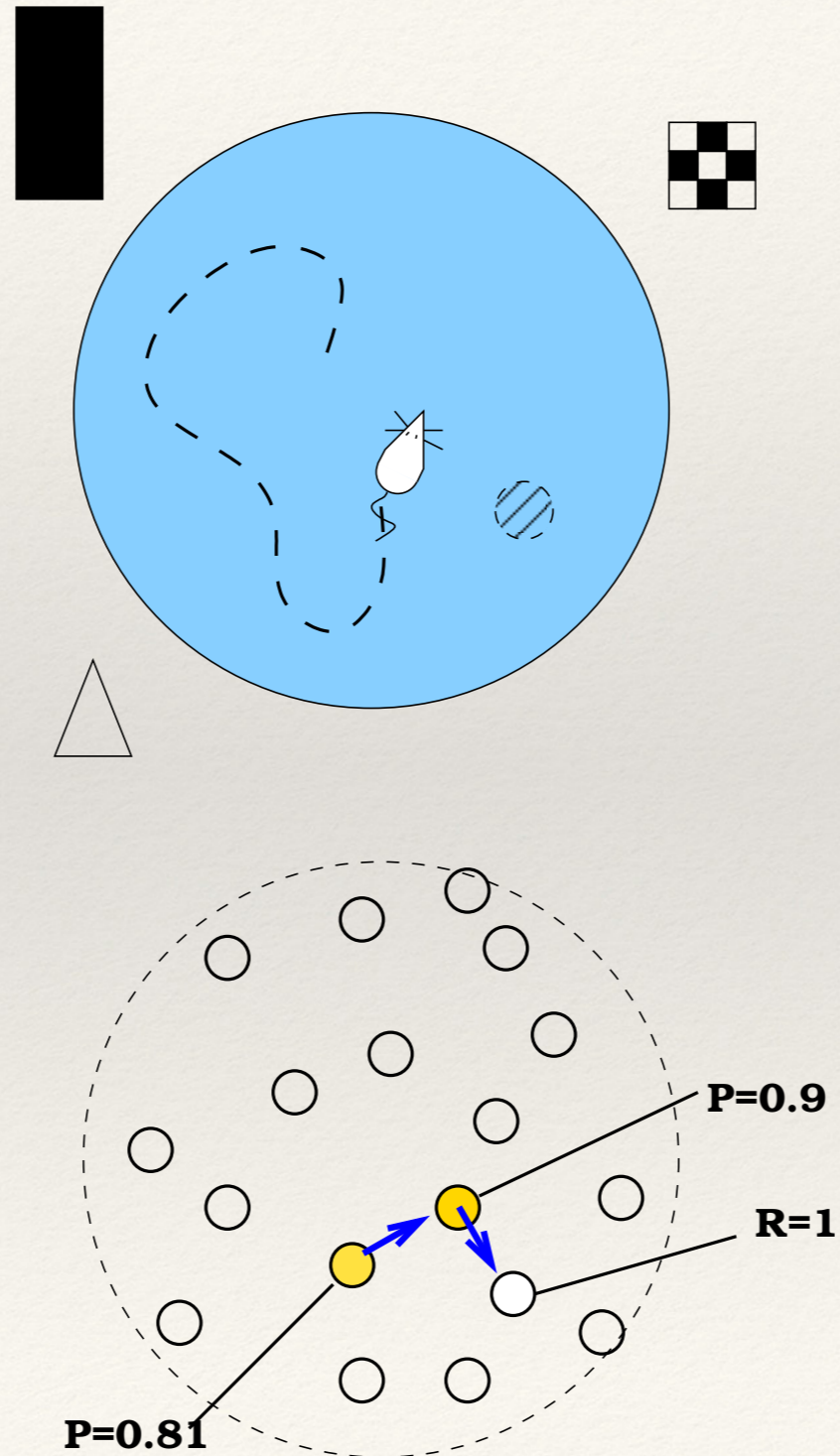
Apprentissage par renforcement sans modèle

Distal Cues



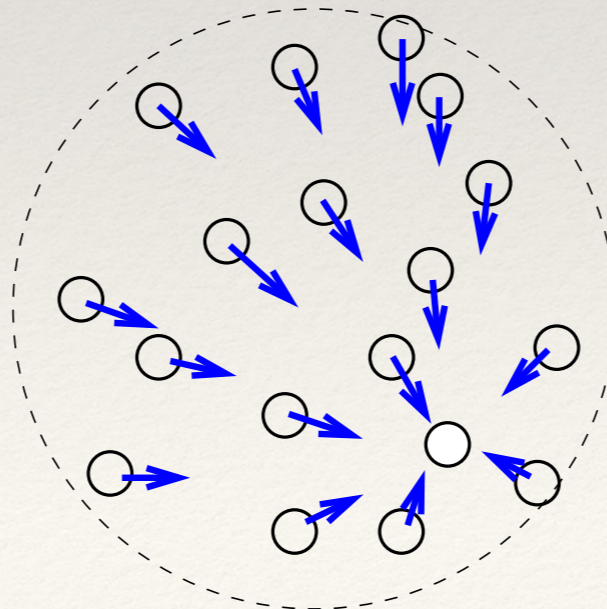
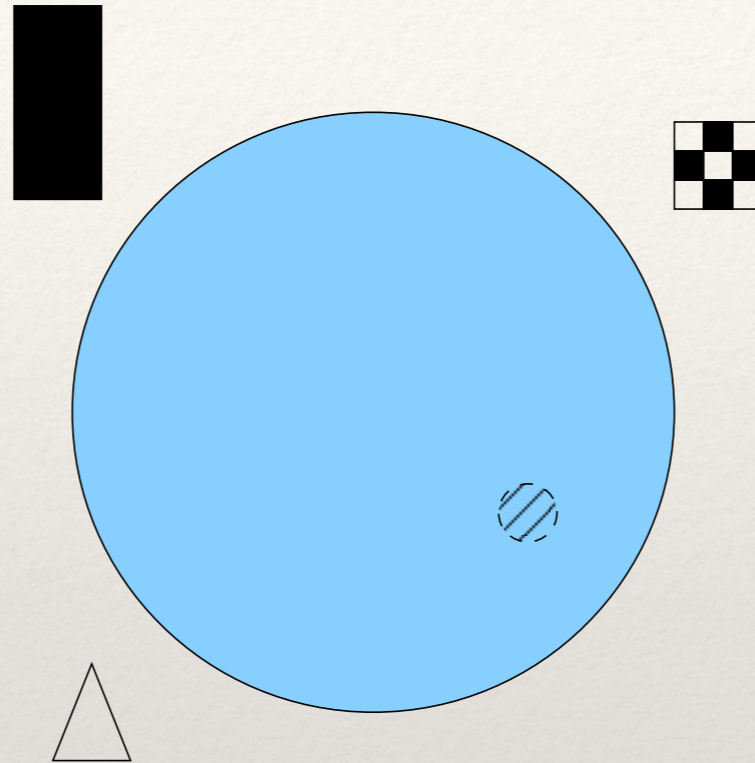
Apprentissage par renforcement sans modèle

Distal Cues



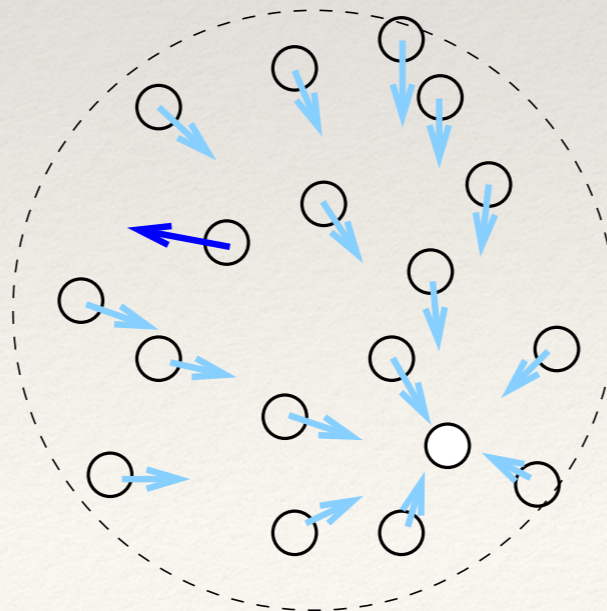
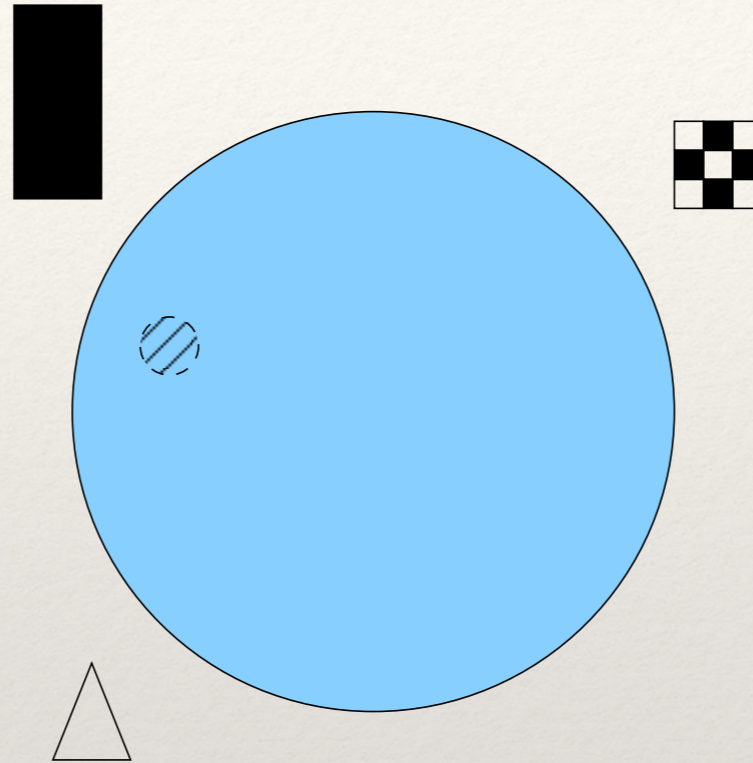
Apprentissage par renforcement sans modèle

Distal Cues



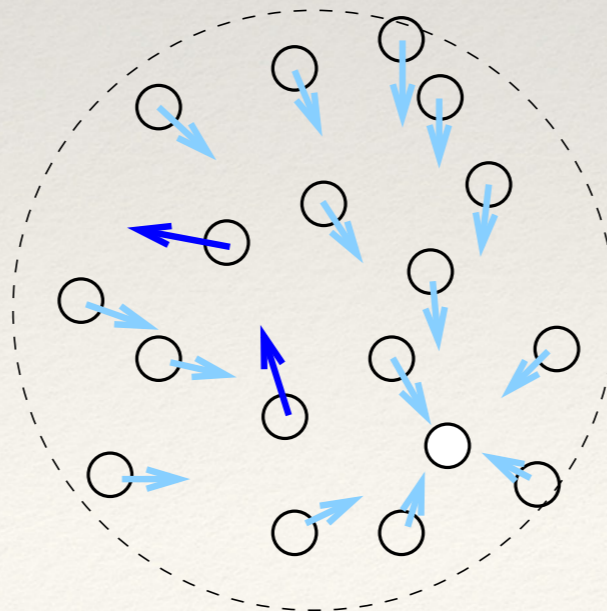
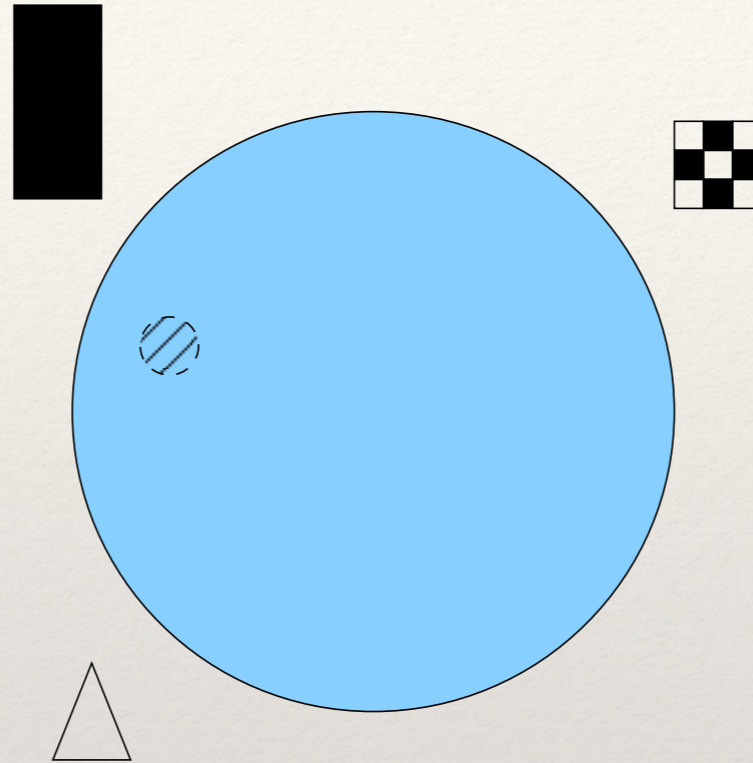
Apprentissage par renforcement sans modèle

Distal Cues



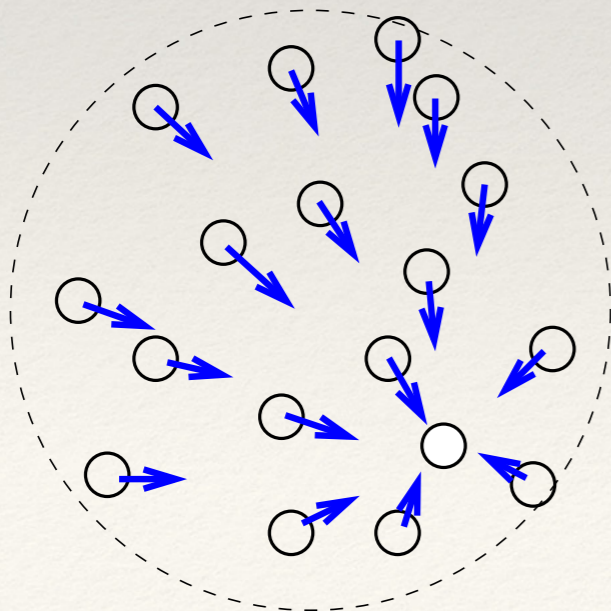
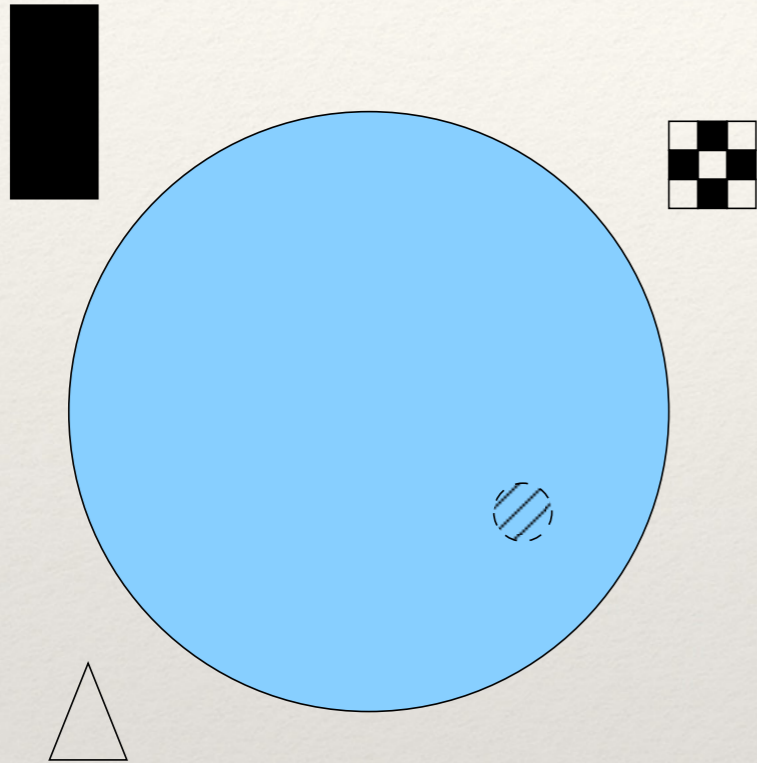
Apprentissage par renforcement sans modèle

Distal Cues



Apprentissage par renforcement sans modèle

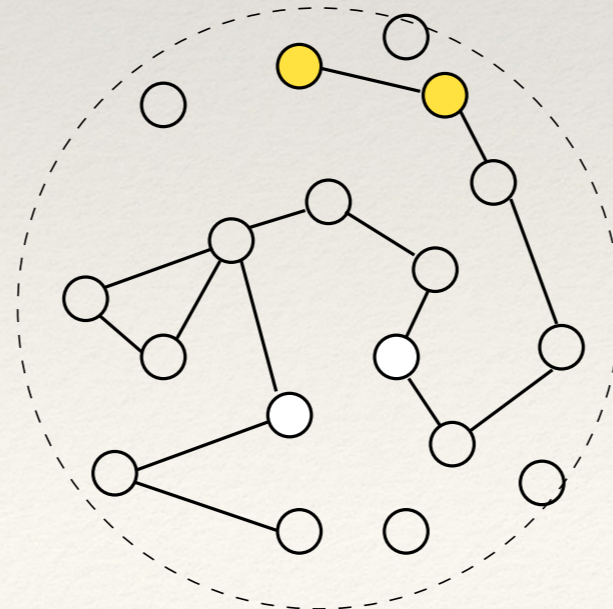
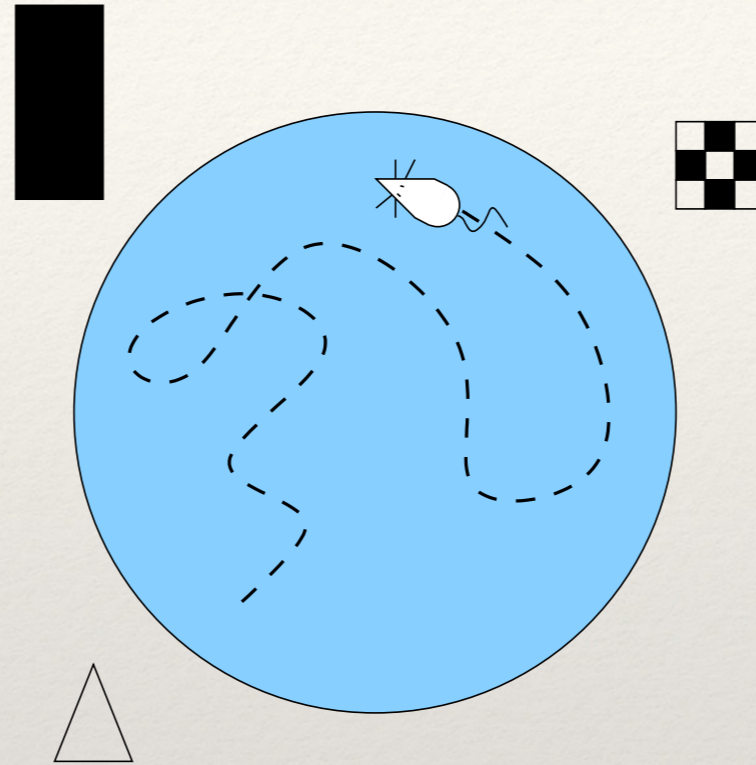
Distal Cues



- ❖ Lent à apprendre,
 - ❖ Nécessite de la récompense,
 - ❖ Rapide à calculer,
 - ❖ Encore plus lent à ré-apprendre.
-
- ❖ Substrat neural :
 - ❖ Cortex sensori-moteur,
 - ❖ Ganglions de la base (dorso-latéral),
 - ❖ système dopaminergique

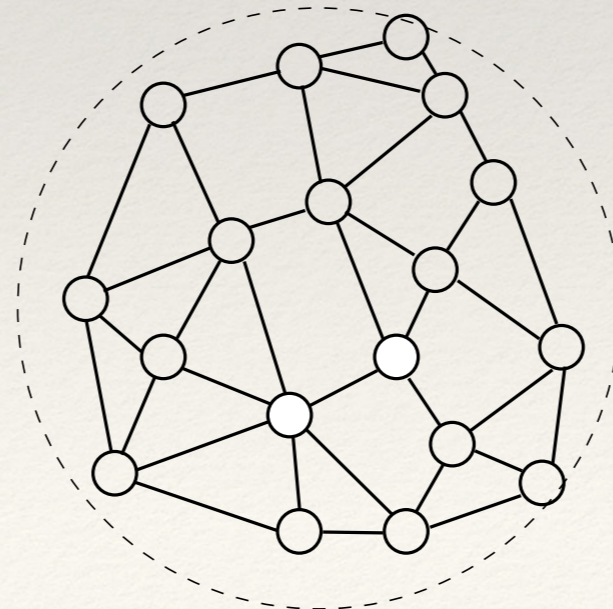
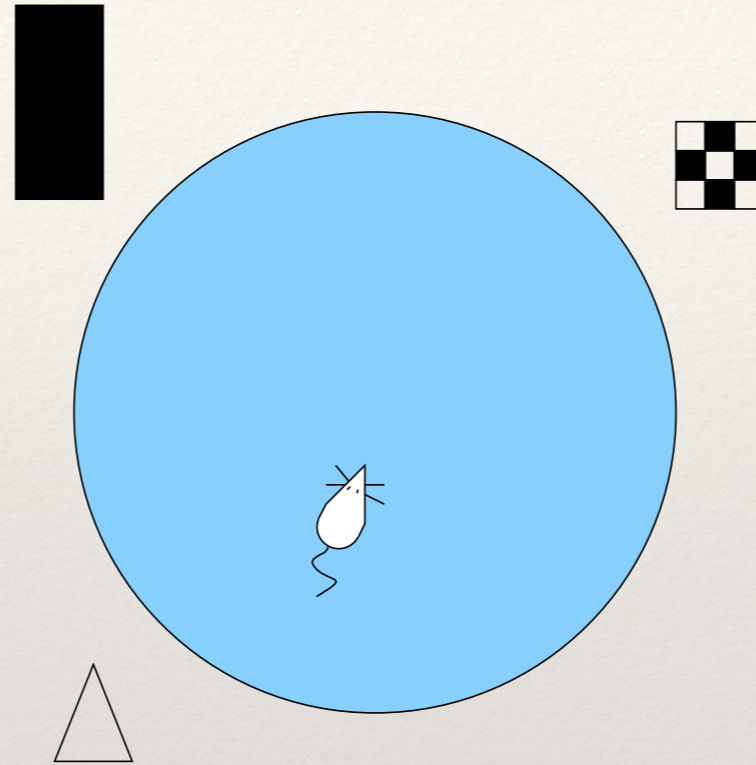
Apprentissage par renforcement avec modèle

Distal Cues



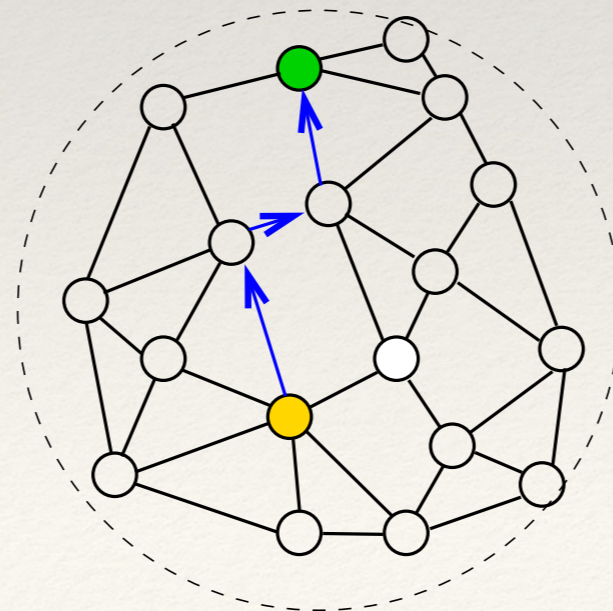
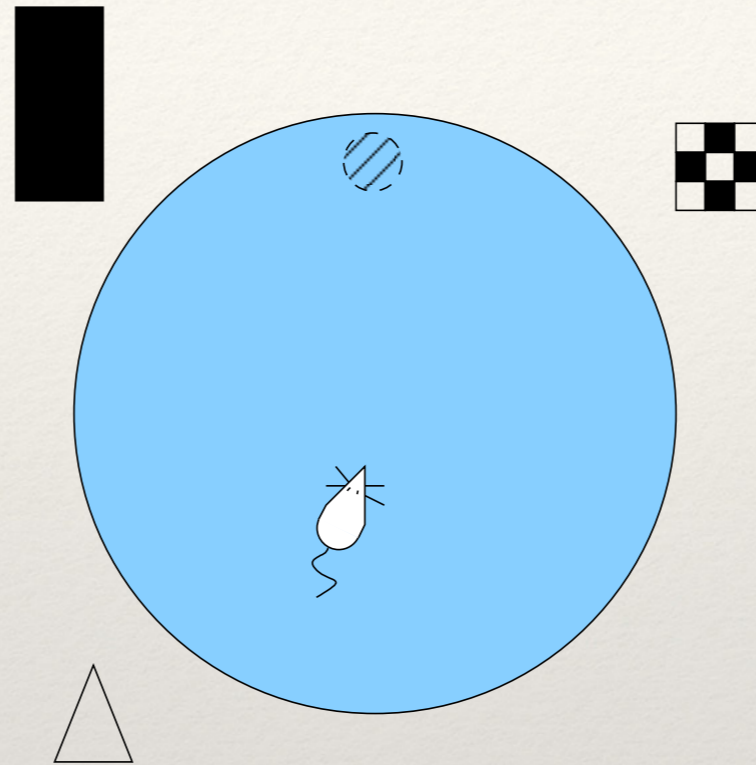
Apprentissage par renforcement avec modèle

Distal Cues

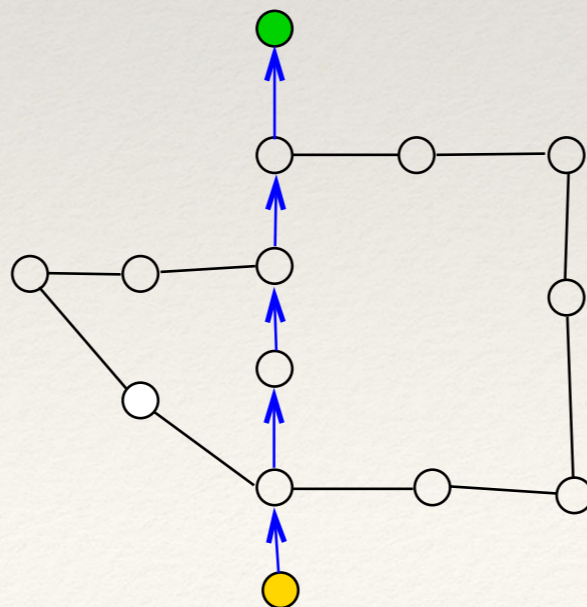
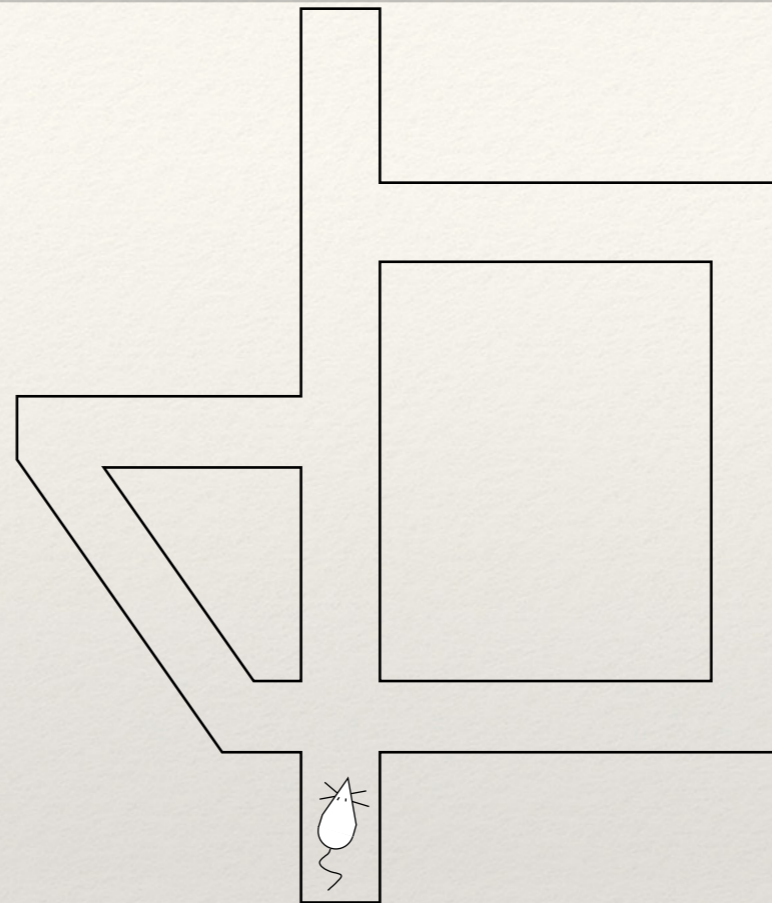


Apprentissage par renforcement avec modèle

Distal Cues

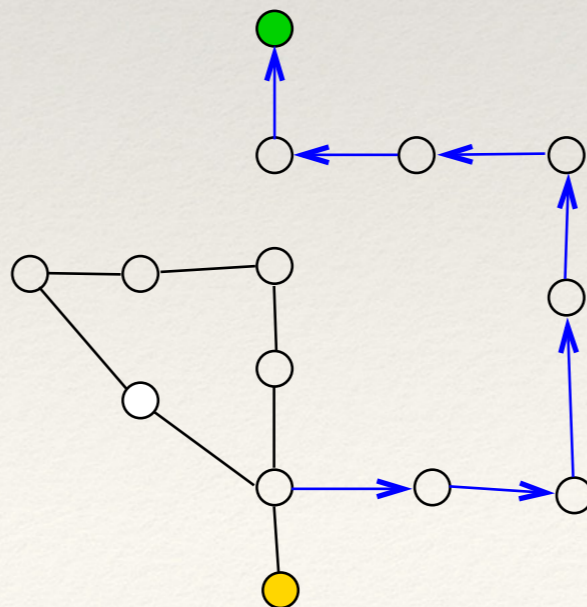
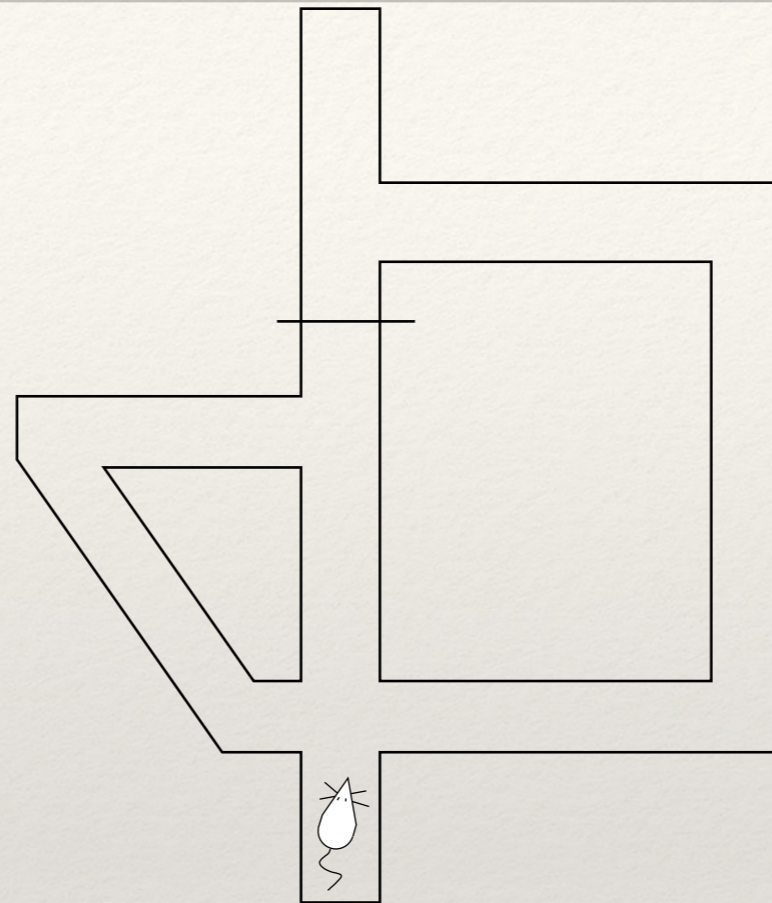


Apprentissage par renforcement avec modèle



(Tolman, 1938)

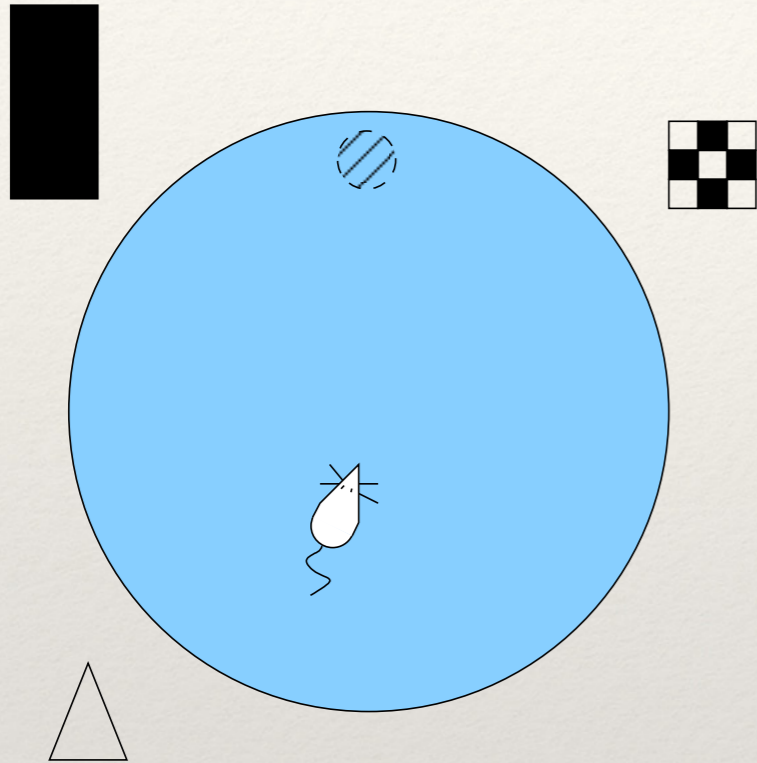
Apprentissage par renforcement avec modèle



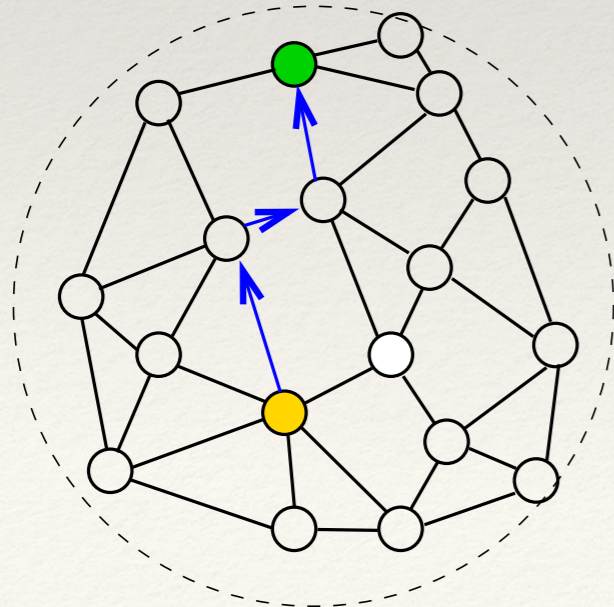
(Tolman, 1938)

Apprentissage par renforcement avec modèle

Distal Cues



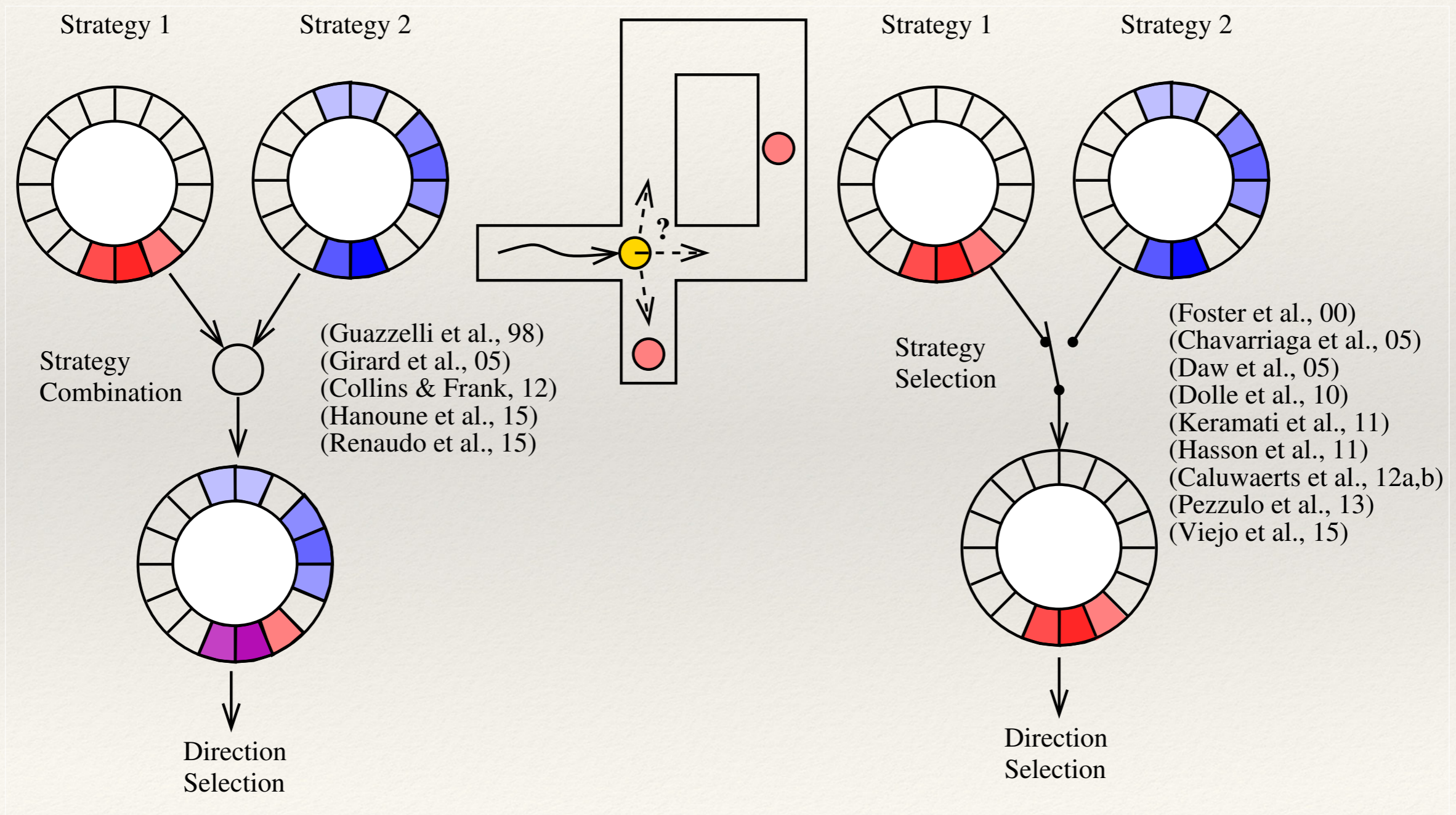
- ❖ Rapide à apprendre,
 - ❖ Composante latente (hebbienne),
 - ❖ Long à calculer,
 - ❖ Rapide à ré-apprendre.
-
- ❖ Substrat neural :
 - ❖ Cortex préfrontal (OFC, vmPFC, dlPFC),
 - ❖ Ganglions de la base (dorso médial)



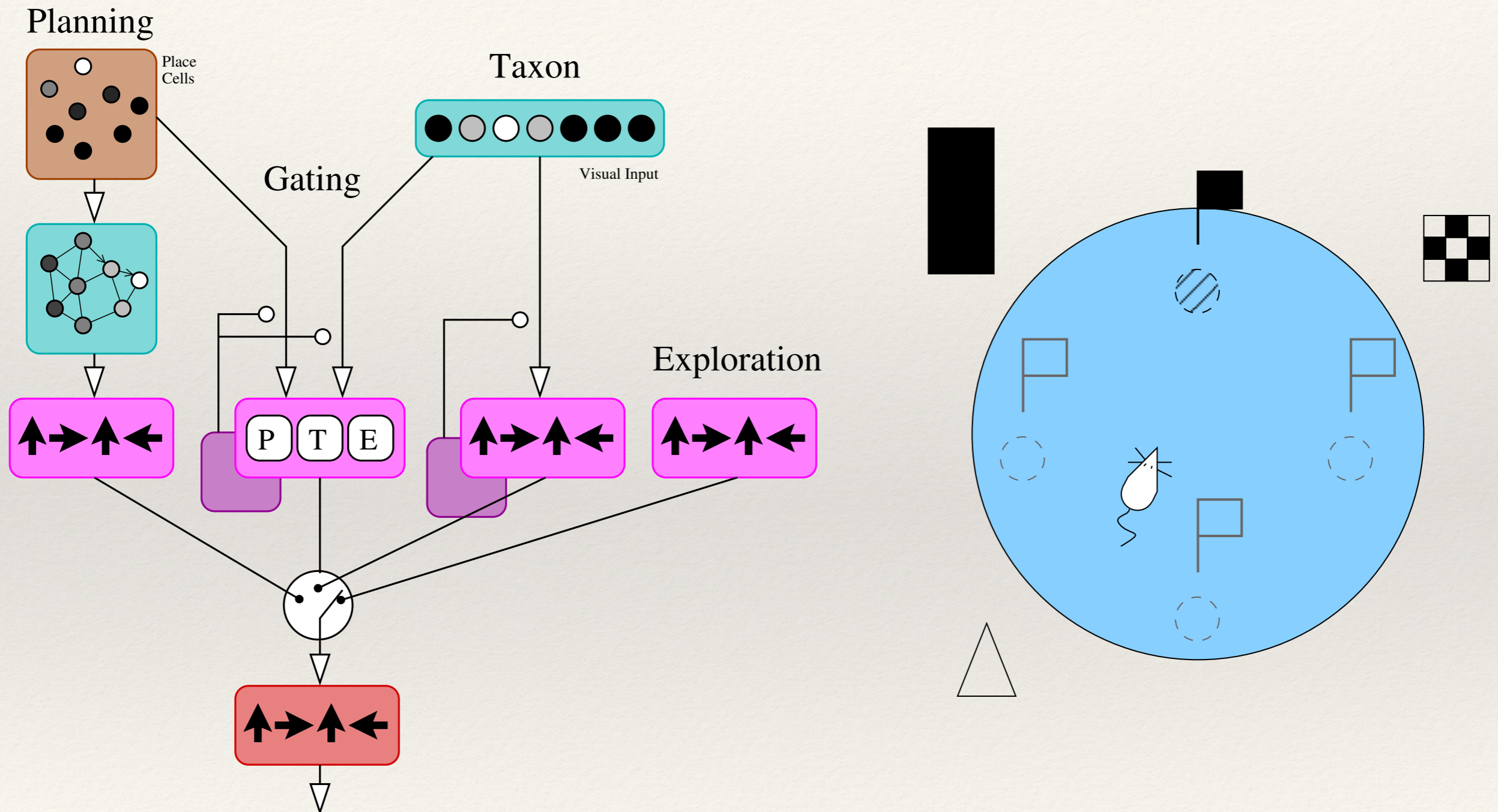
Combinaisons

Fusion (méthodes d'ensemble)

Sélection

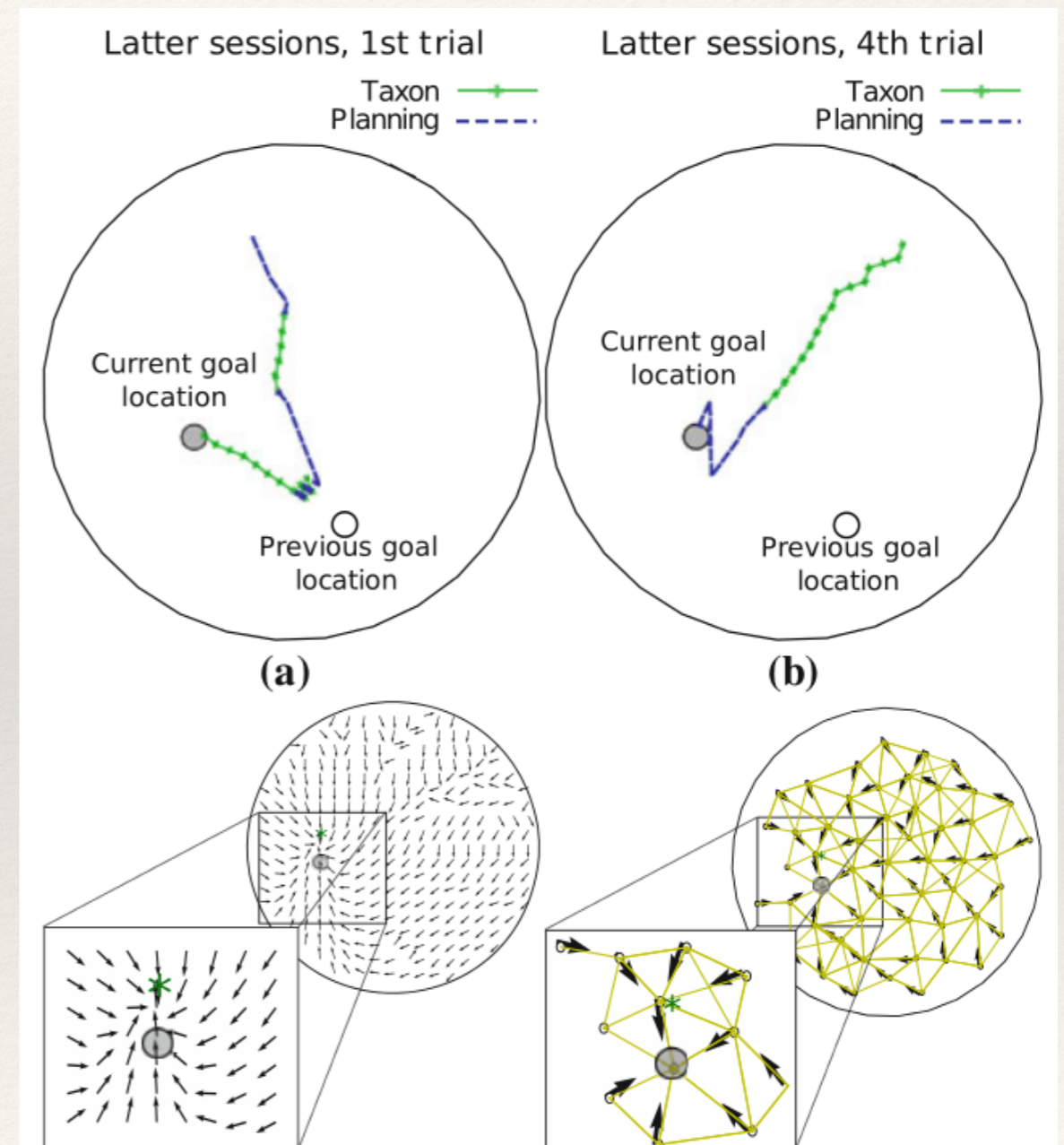
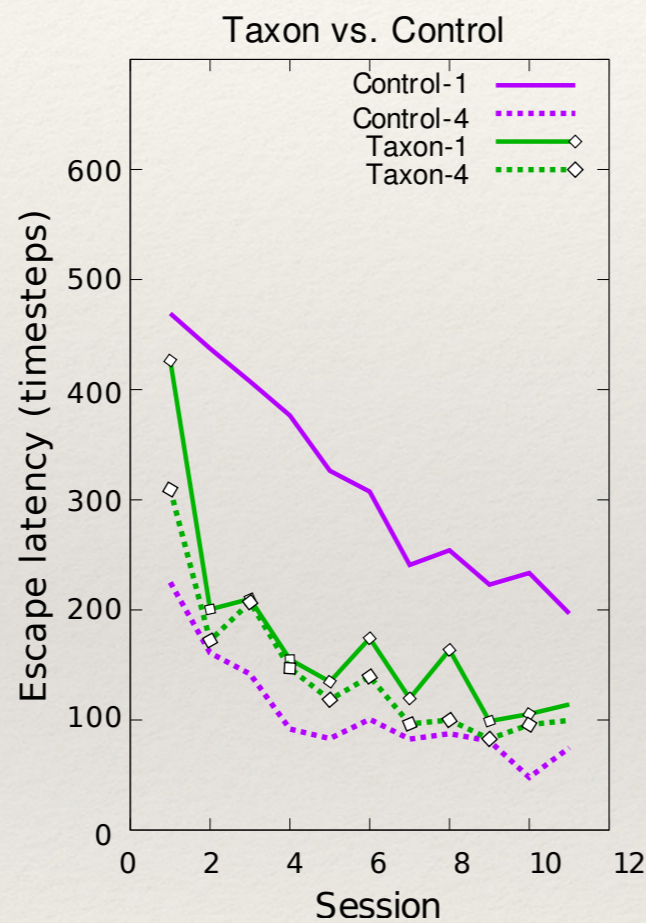
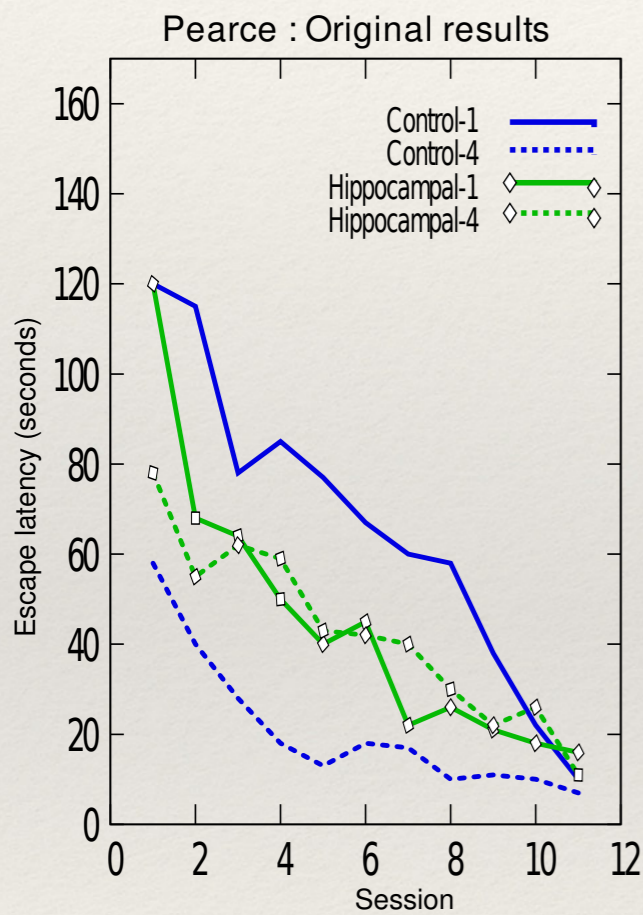


Apprendre à sélectionner - rats



(Dollé et al., 2010)

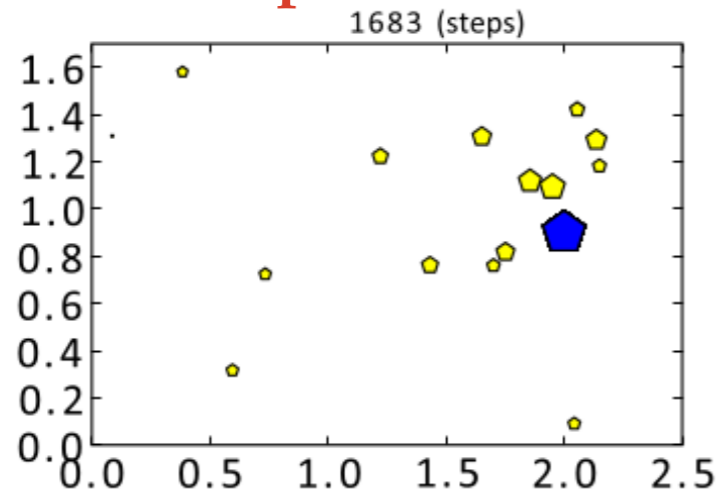
Apprendre à sélectionner - rats



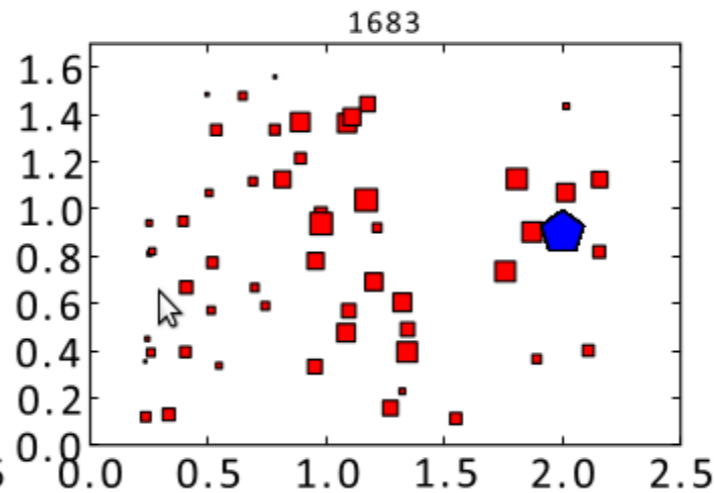
(Dollé et al., 2010)

Apprendre à sélectionner - robots

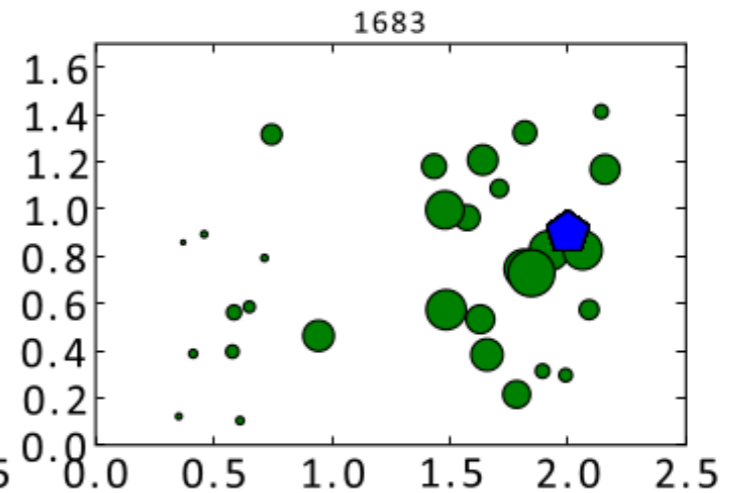
Exploration



Renf. avec modèle

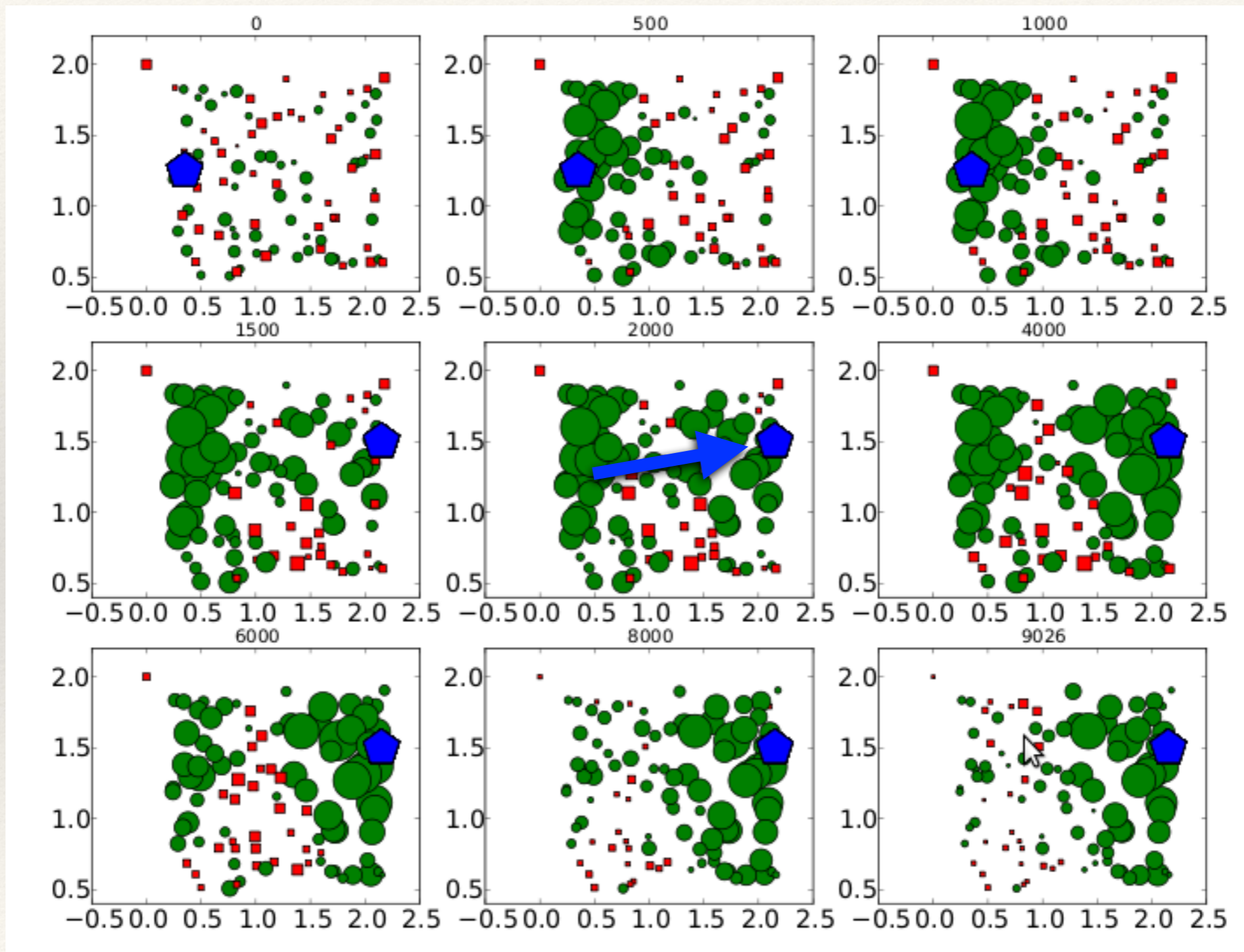


Renf. sans modèle



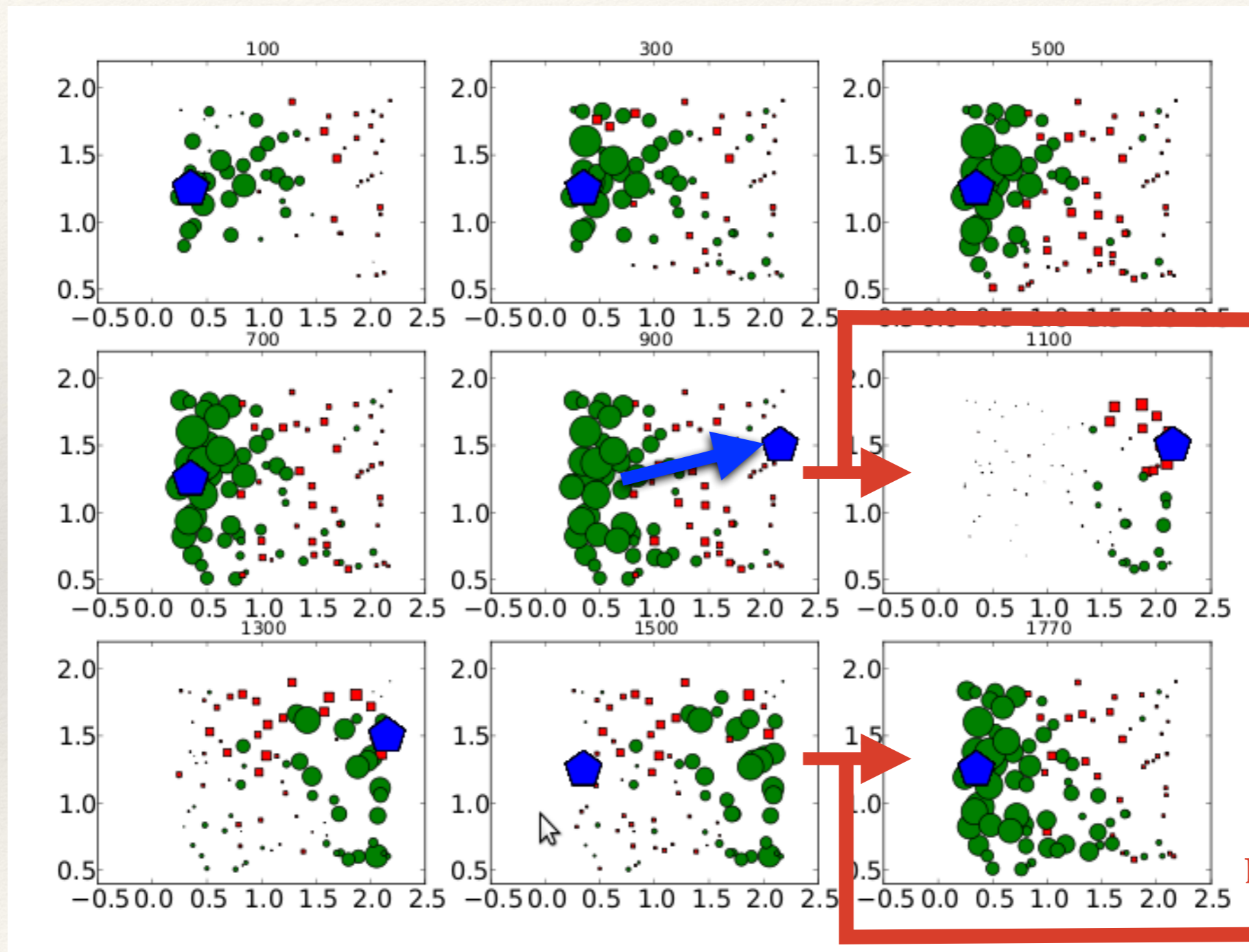
(Caluwaerts et al., 2012)

Apprendre à sélectionner - robots



(Caluwaerts et al., 2012)

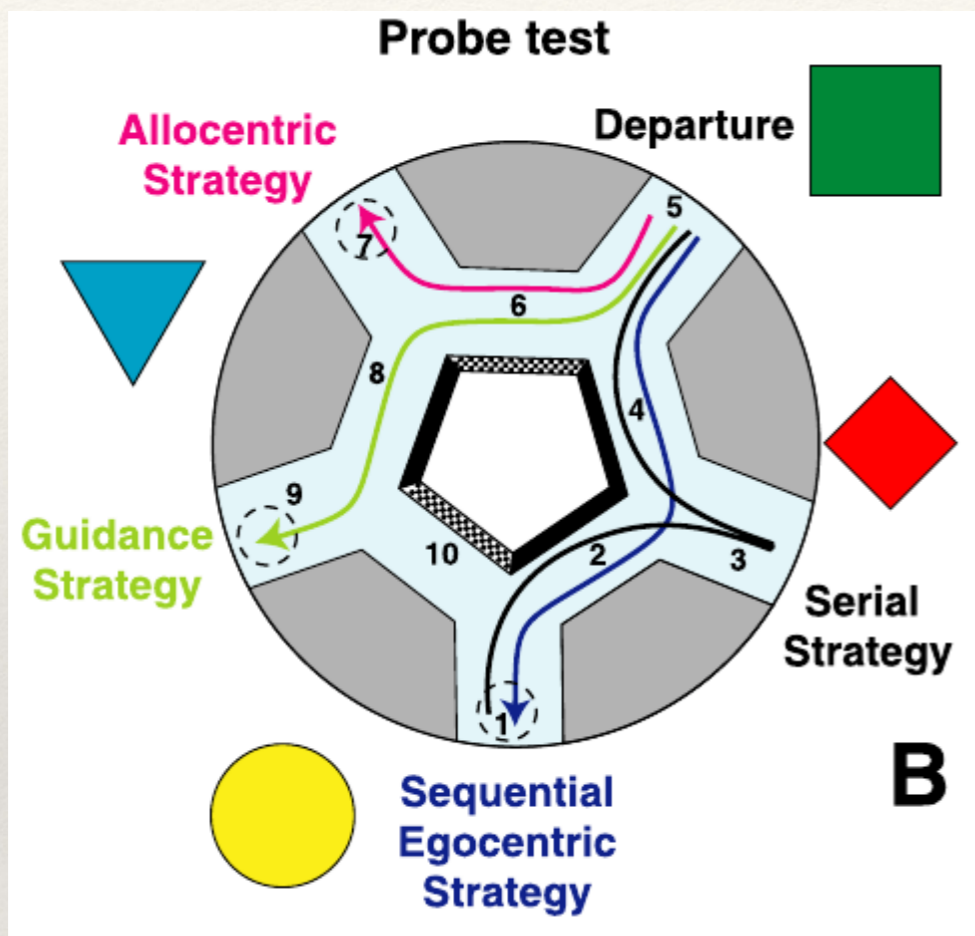
Apprendre à sélectionner - robots



Détection
changement,
Sauvegarde,
Création nouveau
contexte

Détection contexte
initial
Récupération
mémoire

Multiplicité de systèmes complémentaires



- ❖ La multiplicité est la règle,
- ❖ C'est flagrant en navigation.

(Rondi-Reig et al., 2006)

Stratégie :	Réactive	Délibérative
Entrées sensorielles brutes	Réponse à un stimulus	Arbre de décision
	Approche du but	
	Approche d'indice	
	Praxique	
Cellules de lieux	Intégration de chemin	
	Réponse associée à un lieu	Planification topologique
		Planification métrique