

Numéro d'ordre : 4178



THÈSE

PRÉSENTÉE A
L'UNIVERSITÉ BORDEAUX 1

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ
par **Delphine TRAN-TU-YEN**

POUR OBTENIR LE GRADE DE
DOCTEUR
SPÉCIALITÉ NEUROSCIENCES

COORDINATION DES ACTIONS ET DES HABITUDES **Approche neurocomportementale chez le Rat**

Soutenue le 10 décembre 2010

Après avis de :

Mme Marianne AMALRIC (DR, CNRS UMR 6155, Univ. Aix-Marseille I)

M. Philippe OBERLING (MCU/PH, CNRS FRE 3289, Univ. Strasbourg)

Devant la commission d'examen formée de :

Mme Marianne AMALRIC (DR, CNRS UMR 6155, Univ. Aix-Marseille I) **Rapportrice**

M. Philippe OBERLING (MCU/PH, CNRS FRE 3289, Univ. Strasbourg) **Rapporteur**

M. Guillaume FERREIRA (CR, CNRS UMR 5226, Univ. Bordeaux) **Examineur**

M. Anthony DICKINSON (PR, Univ. Cambridge) **Examineur**

M. Etienne COUTUREAU (CR, CNRS UMR 5228, Univ. Bordeaux) **Directeur de thèse**

M. Jean-Louis GUILLOU (PR, CNRS UMR 5228, Univ. Bordeaux) **Président du Jury**

COORDINATION DES ACTIONS ET DES HABITUDES

Approche neurocomportementale chez le Rat

Résumé : L'étude du conditionnement opérant permet une approche théorique et expérimentale des processus cognitifs à l'origine de la prise de décision. De nombreuses études ont établi que la réponse instrumentale, chez les primates et chez les rongeurs, peut dépendre de deux systèmes associatifs. Ainsi, au début de l'apprentissage, les actions dirigées vers un but requièrent une représentation de leur conséquence et dépendent d'une association entre les actions et leurs conséquences (A-O). En revanche, au fur et à mesure de l'entraînement, des réponses de type « habitudes » se mettent en place, engageant alors une association entre les réponses et les stimuli présents dans l'environnement du sujet (S-R). Des données récentes semblent indiquer que ces deux systèmes n'interviennent pas obligatoirement de façon séquentielle, mais seraient présents en parallèle et en compétition à tout moment de l'entraînement. Mon travail de thèse visait à contribuer à ce champ de recherche, en s'intéressant aux mécanismes de sélection des systèmes « action dirigée » et « habitudes » chez le Rat. Le premier ensemble de résultats obtenus n'indiquent pas d'effet du nombre de séances d'entraînement instrumental sur la sensibilité de la performance instrumentale à la dévaluation de la récompense. Malgré le prolongement de leur entraînement, les sujets testés expriment une réponse dirigée vers un but. Une deuxième série d'expériences montre un effet de la dévaluation de la récompense seulement lorsque la phase de dévaluation a lieu dans un contexte différent du contexte opérant. Ces résultats mettent en évidence, pour la première fois, le rôle des stimuli contextuels dans le contrôle de la réponse instrumentale de type « action dirigée ». Un troisième ensemble de résultats obtenus par inactivation ciblée du cortex prélimbique (PL) permet de conclure sur son implication temporaire dans le contrôle du système « action dirigée ». En effet, l'activité neuronale au sein du PL est nécessaire seulement au moment de l'acquisition d'une réponse instrumentale médiée par le système A-O. En revanche, une fois acquise la réponse dirigée vers un but ne requiert plus l'activité du PL mais dépendrait essentiellement du striatum dorso-médian comme le suggèrent les données de la littérature. Un quatrième et dernier ensemble d'expériences visait à étudier l'activité neuronale au sein de ces boucles au cours d'un conditionnement opérant. Ainsi, l'expression du gène précoce Fos a été quantifiée au sein des structures préfrontales et striatales chez divers groupes de sujets sacrifiés à différents moments de l'entraînement. Les résultats obtenus sont conformes aux données actuellement disponibles dans la littérature. L'activité au sein des cortex prélimbique et infralimbique est importante au début de l'entraînement puis diminue avec la répétition des séances d'entraînement. Au contraire, les niveaux d'activation au sein des ganglions de la base sont plus stables, avec une quantification plus élevée au sein du striatum dorso-médian. L'ensemble des résultats obtenus au cours de ce travail de thèse contribue à la vaste entreprise qu'est l'étude des interactions entre les systèmes « action dirigée » et « habitudes ».

Mots Clés : conditionnement opérant, action dirigée, habitudes, récompense, dévaluation, contexte, prélimbique, muscimol, striatum, Fos.

COORDINATION OF ACTIONS AND HABITS

A neurobehavioural approach in Rats

Summary: Previous research has established that instrumental conditioning, in both primates and rats, is mediated by two concurrent associative systems. In early stages of training, instrumental response is thought to be mediated by an association between the action and the outcome (A-O). While training proceeds however, as the response becomes less sensible to the outcome value, it is conceived as being mediated by an association between the stimulus and the response (S-R). Recent evidences suggest that the both systems operate in tandem and/or competition from the beginning of training. This work aimed at studying the mechanisms that coordinate the control of the instrumental response by the goal-directed system or the habit system. A first batch of results indicates no effect of the amount of training sessions on the goal-directed nature of the conditioned instrumental response. Indeed, the outcome devaluations by CTA or selective satiety reduced the instrumental performances, independently of the training procedure applied. The instrumental responses resulting from our 3 training procedures depend of an actualized representation of their outcomes. A second batch of results indicates that information about the context of instrumental conditioning is included in the incentive representation of the outcome. Indeed, we observed no sensitivity to outcome devaluation when devaluation occurred outside the training context. These results offer new original hypotheses about context encoding and the nature of instrumental responding. A third batch of experiments investigates the role of the prelimbic cortex in acquisition vs. expression of goal-directed instrumental behaviour, using reversible neuronal inactivation. The results show that the prelimbic cortex plays a transient but crucial role in the acquisition of goal-directed responding and that the A-O and S-R systems can operate in a competitive fashion early in training. Using ex-vivo imaging, a last batch of experiments aimed to study the temporal cerebral activation throughout instrumental training with a focus on prefrontal and striatal regions. Results show levels of Fos expression that vary with regions. At the beginning of conditioning, the density of Fos positive nuclei is high in the prefrontal regions. It decreases with training. Labelling is denser in the dorsomedial striatum than in the dorsolateral striatum. The weak activation in the dorsolateral striatum appears consistent with the absence of habit. These data are in accordance with data of the literature concerning dynamics of activation in cortico-striatal circuits. Furthermore, they are in agreement with the suggestion that activity in the prelimbic cortex could promote the acquisition of goal-directed action by the induction of neuronal plasticity in the dorsal striatum.

Key words: operant conditioning, goal-directed actions, habits, reward, outcome devaluation, context, prelimbic cortex, muscimol, striatum, FOS.

Remerciements

Ce travail de recherche a été effectué au sein du laboratoire dirigé par le Dr Georges Di Scala, le Centre de Neurosciences Intégratives et Cognitives, UMR5228/Universités de Bordeaux 1 et 2, Avenue des Facultés, 33405 Talence Cedex.

Cette thèse est le fruit d'un travail de réflexion, mené au sein de l'équipe de recherche du Dr Georges Di Scala, comprenant les Docteurs Alain Marchand, Jean-Rémi Pape et Etienne Coutureau qui fut le directeur de cette thèse. Je tiens à tous les remercier pour avoir permis à ce travail d'aboutir, sans oublier bien sûr l'apport des étudiants/postdoc de l'équipe, Aurore San-Galli, Frédéric Esclassan, Fabien Naneix et Nathan Holmes.

Mes remerciements vont également à tous les membres (étudiants, techniciens, chercheurs, stagiaires...) du CNIC que j'ai eu la chance de côtoyer pendant ces 4 années de M2R et de thèse. Même si je ne les nommerai pas tous ici, chacun dans son individualité a fait de cette période de ma vie, une aventure à dimension humaine. Merci à Dominique, Nathalie et Jennifer pour le soin apporté aux animaux. Merci à Marie-Laure et Karine pour toute l'aide administrative. Et un grand merci à Angélique et Laurence, pour leur aide technique évidemment, mais surtout pour leurs oreilles amicales et leurs conseils. Angélique notre collocation de bureau me manquera, entre autres...

Aurore, Edith et Fred, merci pour avoir accueilli la petite M2 que j'étais et pour lui avoir fait découvrir le fabuleux monde du rat de laboratoire.

Nathan, thank you very much for your advice and your incredible kindness. Thank you & Kelly for the custody of Lili, the "wake up up up!!!" Hungry Cat. I will take good care of her.

Susanna, sono stata molto lieta di incontrarti. Grazie mille per avermi fatto praticare il mio poverino italiano... Ricordati, domani dobbiamo andare nuotare!

Malorie, Laurent et Yves vous êtes ceux qui m'ont vu évoluer au plus près au cours de ces 4 années. Jour après jour, vous avez su m'épauler et votre présence a signifié beaucoup pour moi. Yves, j'ai trouvé en toi un écho à mon amour pour l'enseignement bien fait et responsable. Malo & Lolo, même si vous avez souvent peiné à trouver les mots, je mesure l'effort que cela a dû vous coûter de briser votre monde de taiseux. J'espère que j'aurai dans le futur de nombreuses occasions de vous noyer sous mes bavardages intempestifs et délirants.

Je tiens également à remercier l'équipe pédagogique de la Licence Sciences de la Vie et de la Terre de l'Université Bordeaux 1. Ces heures d'enseignement parmi vous, m'ont permis d'assouvir ma soif de transmission et de me confronter pour la première fois au monde de l'enseignement. Merci pour votre accueil chaleureux à Isabelle Rondot, Jacques Puiroux, Anja Gellida-Knoll, Sophie Javerzat, Pascal Fossat et aux pétillantes techniciennes Flora et Alexia.

Enfin, mes derniers remerciements iront tout naturellement à mes familles de « sang » et de « cœur », pour leur soutien passé, présent et futur. J'ai vécu ces dernières années loin de vous, et si parfois j'avais un sentiment de solitude, jamais je ne me suis sentie abandonnée. J'ai toujours eu une mère, une sœur, une amie à qui me confier. Merci pour les longs moments au téléphone ou sur msn, les visites, les petits mots d'encouragement... Pardon, d'avoir si souvent été source d'inquiétude. J'espère dans le futur être plus souvent source de rires, de joie et de couleurs. Chaque jour vous m'avez donné la force d'avancer, vous êtes en quelque sorte mon château des ducs. Vous connaître, avoir la chance de partager ces moments de vie avec vous, c'est cela dont je suis le plus fière aujourd'hui. MERCI.

Table des matières

Introduction générale _____ **14**

1. Aspects comportementaux des réponses dirigées vers un but et des habitudes _____ **17**

2. Aspects neurobiologiques des réponses dirigées vers un but et des habitudes _____ **27**

3. Objectifs du travail de thèse _____ **38**

Chapitre I _____ **42**

Introduction _____ **42**

Expérience I.1 _____ 44
Action dirigée et aversion gustative conditionnée

Expérience I.2 _____ 55
Extinction de l'AGC et Action dirigée

Expérience I.3 _____ 59
Procédure Habitude, 1^{er} essai

Expérience I.4 _____ 64
Procédure Habitude, 2^{ème} essai

Chapitre I : Discussion _____ **71**

Chapitre II _____ **780**

Introduction _____ **80**

Expérience II.1 _____ 82
Action dirigée et spécificité de l'encodage de la récompense

Expérience II.2 _____ 92
Action dirigée et spécificité de l'encodage du contexte

Expérience II.3 _____ 100
Action dirigée et rôle du contexte dans la mise à jour de la valeur de la récompense

Expérience II.4 _____ 108
Action dirigée et rôle du contexte dans la signalisation de la valeur de la récompense

Chapitre II : Discussion _____ **118**

Chapitre III	124
Introduction	124
Expérience III.1	126
Implication du cortex prélimbique dans l'apprentissage d'une action dirigée	
Expérience III.2	138
Implication du cortex prélimbique dans l'expression d'une action dirigée déjà acquise	
Chapitre III : Discussion	145
 Chapitre IV	 152
Introduction	152
Expérience IV	154
Circuits neuronaux de l'action	
Chapitre IV : Discussion	164
 Discussion générale	 170
 Références bibliographiques	 174
Annexes	188

Introduction générale

Introduction générale

Ce travail s'appuie sur un parti pris théorique fort, en cela qu'il s'inscrit dans le courant de recherche qui considère que les processus de haut niveau résultent de l'intégration de processus cognitifs élémentaires pouvant être étudiés dans le cadre théorique et méthodologique des apprentissages associatifs (Rescorla, 1988 ; Wasserman et Miller, 1997 ; Kirsch et al., 2004 ; Holland, 2008 ; De Wit et Dickinson, 2009 ; Shanks, 2010).

Ces dernières années, cette recherche a mis en évidence que l'adaptation d'un sujet à un environnement en perpétuel changement nécessitait l'intégration de processus de prédiction, de sélection de l'action et enfin d'évaluation de l'action.

Les travaux portant sur l'acquisition des liens de prédiction résultant de l'association entre plusieurs événements ont conduit à des élaborations théoriques qui toutes ont souligné la richesse de cette représentation. Divers travaux ont mis en évidence l'intégration au sein de la représentation des propriétés sensorielles, motivationnelles, hédoniques ou encore temporelles des événements (Delamater et Oakeshott, 2007 ; Holland, 2008 ; Balleine et O'Doherty, 2010). Mais si le sujet n'était qu'une machine à produire des prédictions, il ne pourrait agir efficacement sur son monde, et un nombre important d'études tente à l'heure actuelle de comprendre les mécanismes par lesquels sont utilisées les représentations générées par l'acquisition des liens de prédiction (De Wit et Dickinson, 2009). Mon travail de thèse se situe dans le cadre de l'évaluation de l'action en cela qu'il vise, par une approche neurocomportementale, à étudier les mécanismes du contrôle de l'action.

Le cadre théorique et méthodologique des apprentissages associatifs a un impact considérable sur les neurosciences. L'exemple récent le plus remarquable de cette influence est peut-être celle de certains modèles d'apprentissage, pourtant élaborés il y a bientôt 40 ans (Rescorla et Wagner, 1972), sur diverses disciplines en pleine expansion comme la modélisation neurocomputationnelle (Dayan et Daw, 2008) ou la neuroéconomie (Kable et Glimcher, 2009). Le succès de cette entreprise reflète également le fait que les théories des apprentissages associatifs offrent un cadre d'analyse formelle du fonctionnement cérébral. Par exemple, c'est à la rencontre d'Anthony Dickinson avec Wolfram Schultz que l'on doit la proposition la plus approfondie du rôle de la dopamine dans les processus de prédiction et de contrôle des événements (Schultz et Dickinson, 2000). Enfin, ce courant de recherche offre un cadre permettant de prendre en compte les dimensions ontogénétiques et phylogénétiques de

la cognition, en cela qu'il permet d'étudier le comportement en utilisant des tâches structurellement identiques du rongeur à l'humain. De fait, un nombre important d'études récentes (ex : De Wit et Dickinson, 2009 ; Balleine et O'Doherty, 2010) ont montré une grande conservation des mécanismes élémentaires de la cognition des rongeurs jusqu'à l'homme. Ces travaux ont clairement fourni aux neurosciences des pistes pour l'étude des circuits neuronaux à la base de la cognition normale et pathologique (Dalley et al, 2004, Ashby et al., 2010).

1. Aspects comportementaux des réponses dirigées vers un but et des habitudes

Ces dernières années, un nombre important d'études a démontré que la performance des actions liées à l'obtention d'une récompense reflète l'interaction entre deux processus d'apprentissage. Le premier contrôle l'acquisition d'un comportement lié **à la représentation du but** et le second contrôle celle des **habitudes**.

- **Représentation de la valeur du but**

La première démonstration que les animaux encodent les conséquences de leur action date d'environ 30 ans maintenant (Adams, 1982 ; Adams et Dickinson, 1981). Cette étude classique comprend différentes phases qui sont schématisées dans la figure 1. Dans une première phase, les animaux étaient placés dans une cage de conditionnement et pouvaient librement appuyer sur un levier pour obtenir une récompense alimentaire (dans le cas présent une quantité fixée d'une solution d'eau sucrée). Dans une seconde phase, la phase de dévaluation, la valeur de la récompense est altérée en associant la consommation de la récompense avec un malaise digestif résultant de l'injection intra-péritonéale d'un agent nauséux, le Chlorure de Lithium (LiCl). A l'issue de la dévaluation, les animaux sont replacés dans la cage de conditionnement et peuvent de nouveau appuyer librement sur le levier. Néanmoins, l'appui est inefficace puisqu'il n'engendre aucune conséquence particulière. De fait, cette procédure de test assure que la réponse de l'animal résulte bien de l'utilisation d'une représentation de la valeur de la récompense.



Figure 1 Schématisation des phases d'une procédure de dévaluation de la récompense par aversion gustative et résultats d'un test en extinction (d'après Adams et Dickinson, 1981).

La dévaluation se tient après l'acquisition de la tâche instrumentale (A1). La récompense est proposée en accès libre au sujet et l'animal reçoit une injection intra-péritonéale, immédiatement après la prise alimentaire (A2). Les sujets du groupe P reçoivent du chlorure de Lithium qui va induire un malaise digestif, ceux du groupe U une solution contrôle inoffensive. Lors d'un test non récompensé (B) les sujets P appuient très peu sur le levier en comparaison avec les sujets U, ils sont donc capables d'adapter leur réponse aux changements de valeur (A3).

Le panneau B de la figure illustre les résultats de cette étude. En comparaison du groupe contrôle U (dans lequel la consommation de la récompense et le malaise digestif étaient dissociés), les animaux du groupe P exprimaient une fréquence d'appuis sur le levier beaucoup plus faible. Ce résultat, fondateur pour ce domaine de recherche, démontre d'une part que l'animal encode l'association entre l'appui et la récompense durant l'entraînement et surtout qu'il est capable d'ajuster sa réponse en fonction d'une représentation de la valeur de la récompense. Depuis, ce résultat a été reproduit dans des situations de choix, comme schématisé dans la figure 2 (voir Colwill et Rescorla, 1985 ; Rescorla, 1991 ; Balleine et Dickinson, 1998a ; Holland, 2004 ; Naneix et al., 2009 ; Balleine et O'Doherty, 2010).



Figure 2 Schématisation d'une procédure de dévaluation de la récompense par aversion gustative et résultats d'un test en extinction mettant en compétition deux actions (d'après Colwill et Rescorla, 1985).

Les sujets apprennent deux tâches instrumentales simultanément, chacune d'elles est associée à une récompense spécifique (A). La dévaluation est apprise pour la récompense associée à la réponse 1 pour la moitié des sujets (B, ND2-D1). L'autre moitié des sujets acquièrent une aversion pour l'autre récompense (B', ND1-D2). Lors du test non récompensé les animaux appuient préférentiellement sur le levier précédemment associé à la récompense qui n'a pas été dévaluée.

Le contrôle de la réponse par le système dépendant de la représentation a également été obtenu en utilisant une méthode alternative de dévaluation, à savoir la satiété sensorielle spécifique, telle que schématisée dans la figure 3 (voir Balleine et Dickinson, 1998b ; Killcross et Coutureau, 2003 ; Yin et al, 2005b). Cette procédure de dévaluation consiste, à l'issue de la phase d'entraînement, à fournir à l'animal d'importantes quantités de la nourriture ayant servie à entraîner les animaux. Comme le montre le panneau B, les animaux ainsi dévalués présentent une activité d'appuis largement inférieure à celle d'animaux contrôles ayant reçu jusqu'à satiété une nourriture contrôle.



Figure 3 Schématisation d'une procédure de dévaluation de la récompense par satiété sensorielle spécifique et résultats d'un test en extinction (d'après Balleine et Dickinson, 1998).

Suite à la phase de conditionnement instrumental (A1), on fournit aux sujets une quantité abondante de la récompense et on les laisse s'en sustenter pendant une heure (A2). Immédiatement après la prise alimentaire les sujets sont placés en présence du levier (A3), pour un test en extinction. On peut voir sur le panneau B que leurs performances instrumentales sont alors beaucoup plus faibles que celles de sujets n'ayant pas été nourris au préalable avec la récompense.

Il est important de noter que, même si les processus neurobiologiques sous jacents divergent, des données de la littérature indiquent que la dévaluation obtenue par aversion ou par satiété sensorielle altère les mêmes propriétés de la récompense. En effet, des données issues de l'observation des mimiques oro-faciales (Berridge, 1991 ; Berridge et al., 2009) ont montré que ces deux traitements induisent un changement hédonique. En effet, alors que la présentation d'une récompense appréciée induit des mimiques oro-faciales d'appétence, la dévaluation de cette même récompense change la valeur qui lui est attribuée puisque les mimiques oro-faciales traduisent un dégoût. De même, lorsqu'un rat dans un état de satiété sélective pour une solution sucrée est mis en contact forcé (infusion intra-orale) avec celle-ci, il ne présente pas les mimiques caractéristiques du dégoût, mais ses expressions oro-faciales caractéristiques du plaisir sont considérablement réduites. De plus, les données d'électrophysiologie obtenues par Rolls (2005) chez le primate lors d'un état de satiété sélective mettent en évidence une diminution des trains de décharge des neurones uniquement

dans les aires secondaires gustatives. Enfin, les sujets humains rapportent que l'intensité gustative de la récompense n'est pas modifiée, alors que la sensation de plaisir se trouve diminuée par la satiété (voir Rolls, 2005 pour revue).



Figure 4 Réactions oro-faciales hédoniques ou aversives chez le Rat, l'Orang-outang et le nouveau né Humain (d'après Berridge et Kringelbach, 2008).

Les tests de réactions oro-faciales développés chez les rongeurs par Grill et Norgren en 1978 mesurent le degré de palatabilité d'une substance. Le contact avec une substance agréable au goût comme une solution sucrée provoque l'extension de la langue, le léchage des babines/lèvres et des pattes/mains. L'injection d'une solution de quinine au goût très amer provoque des réactions typiques d'une aversion gustative : l'ouverture de la bouche, des frissons, des soubresauts de la tête.

La réponse ne dépend pas tout le temps de la représentation. En effet, Adams (1982) fut le premier à démontrer que la sensibilité de la réponse aux changements de valeur dépendait du degré d'entraînement de la réponse.



Figure 5 Schématisation d'une procédure de surentraînement et résultats d'un test en extinction (d'après Adams, 1982 ; Expérience I).

Deux groupes de sujets sont constitués. Un groupe de sujets suit un entraînement modéré pour l'obtention de 100 récompenses (non représenté). Un autre groupe de sujets suit un entraînement intensif qui leur permet d'obtenir 500 récompenses (A1). L'ensemble des sujets est soumis à la phase de dévaluation par aversion gustative (A2). Lors du test en extinction les animaux surentraînés et ayant reçu du LiCl (500-P) ne sont pas sensibles à la dévaluation, au contraire leur nombre d'appuis est très élevé. (U animaux injectés avec du solvant ; P animaux injectés avec du LiCl ; 500 surentraînement, 100 entraînement modéré.)

Comme l'illustre la figure 5, les niveaux de performance étaient sensibles à des changements dans la condition modérée mais pas dans la condition intensive, démontrant ainsi que sous l'effet de l'entraînement, la nature de la réponse change. Ultérieurement, A. Dickinson (voir Dickinson, 1985) a proposé que la réponse soit sous le contrôle de deux influences : les

réponses dirigées vers un but (Action-Outcome, A-O) et les habitudes (Stimulus-Response, S-R).

- **Représentation du lien causal**

En plus de ces différences dans la structure associative, telle que démontrée par leur sensibilité différente à la dévaluation, un corpus important de données suggère que actions et habitudes sont appris selon des règles différentes. En particulier, il est apparu que les actions, au contraire des habitudes, étaient sensibles aux changements de contingence entre la réponse et le résultat de celle-ci. La contingence peut être formalisée comme la probabilité d'obtenir le résultat (récompense) *en effectuant l'action* moins la probabilité d'obtenir ce même résultat *en n'effectuant pas l'action*.

Une des premières démonstrations de l'influence de la contingence sur l'apprentissage instrumental provient d'une étude de Hammond (1980). Dans cette étude, schématisée dans la figure 6, cet auteur a pu démontrer que des variations de la probabilité d'obtenir le résultat de l'action sans l'avoir effectuée avaient des conséquences importantes sur les performances des animaux.

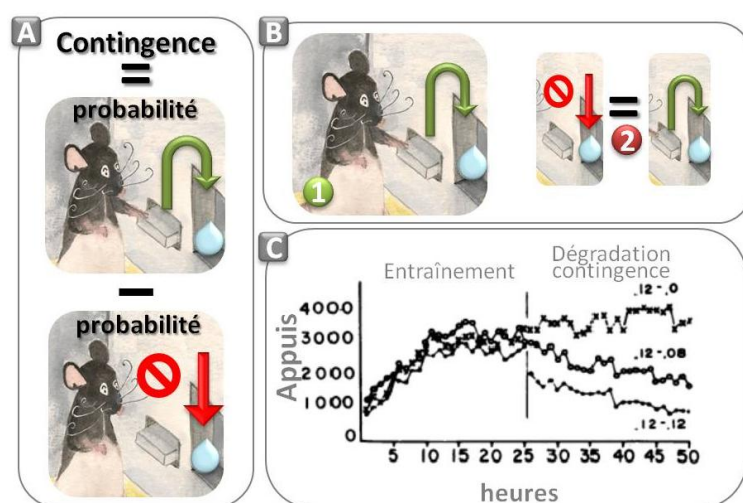


Figure 6 Définition de la contingence (A), schématisation et résultats d'une procédure de dégradation de la contingence (B et C) (d'après Hammond, 1980).

Suite à l'entraînement (B1), deux groupes sont soumis à des conditions où la récompense peut être délivrée suite à l'action mais également en absence d'appui (B2). Les performances instrumentales diminuent alors peu à peu (C).

Ce résultat a été reproduit à de nombreuses reprises (voir Dickinson et Mulareto, 1989 ; Balleine et Dickinson, 1998a ; Ostlund et Balleine, 2005) et surtout en utilisant des procédures mieux contrôlées. En particulier, la distribution de récompense non contingente est susceptible de favoriser l'émergence de comportement alternatif (ex : visite à la mangeoire) qui pourrait entrer en compétition avec l'appui. De fait, un raffinement (ex : Balleine et Dickinson, 1998a) de ces procédures a consisté à comparer deux conditions, une condition « différente » et une condition « identique » telle que schématisée dans la figure 7.

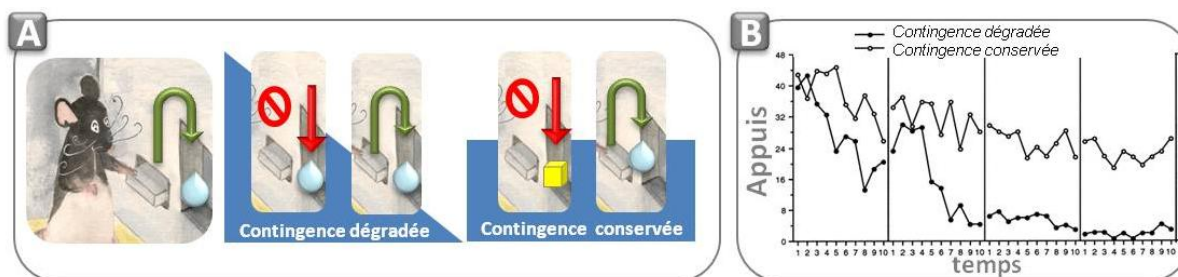


Figure 7 Dégradation de la contingence avec deux types de récompenses (d'après Balleine et Dickinson, 1998).

Suite à l'entraînement, deux groupes sont soumis à des changements de contingence (A). Pour un groupe, la récompense distribuée automatiquement correspond à celle précédemment associée à l'action. L'autre groupe reçoit une récompense d'un type différent. Seuls les animaux recevant automatiquement la récompense associée au levier diminuent leur performance (B).

Comme illustré, la dégradation de la contingence a un impact déterminant sur la performance de l'animal (voir Ostlund et Balleine, 2005 ; Lex et Hauber, 2010a ; Naneix et al., 2009). Les données disponibles concernant l'impact de la dégradation de contingence sur une réponse habitude sont largement moins abondantes et concernent des procédures d'imposition d'omission. Ce type de procédures consiste à imposer un changement brutal de contingence, passant d'une contingence positive à une contingence négative. Ces études ont montré que des réponses qui avaient été entraînées de manière intensive n'étaient pas sensibles à de tels changements de contingence (Balleine et Dickinson, 1998a ; Yin et al., 2004).

Il découle de l'ensemble de ces données la conception séquentielle du contrôle de la réponse. Selon cette conception, proposée par A. Dickinson en 1985, les systèmes de contrôle de la réponse opèrent de manière séquentielle, les habitudes remplaçant peu à peu les actions dirigées au fil de l'entraînement.

Actions et habitudes : conditions de contrôle

Programmes de renforcement

Le premier facteur à prendre en compte est celui des programmes de renforcement mis en œuvre lors de l'entraînement. Deux types de programme de renforcement sont généralement utilisés. Dans un programme en Ratio, un certain nombre de réponses doivent être émises avant qu'une réponse soit récompensée. Au contraire, c'est le temps qui constitue la contrainte dans un programme en intervalle puisque, suite au renforcement d'une réponse, un laps de temps est imposé pour qu'une réponse émise ultérieurement soit récompensée. Par ailleurs, les valeurs des ratios et intervalles peuvent être fixes ou variables. La mise en œuvre de ces types de programme de renforcement a des conséquences importantes sur le patron d'appuis comme l'ont démontré de nombreux auteurs (Ferster et Skinner, 1957 ; Catania et

Reynolds, 1968 ; Reed, 2007). Ces patrons sont illustrés dans la figure 8 adaptée de Reynolds (1975).

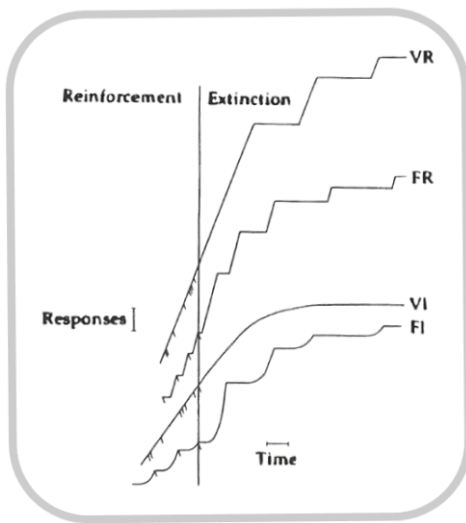


Figure 8 Evolution du nombre total d'appuis au cours des 4 programmes de renforcement VR, FR, VI et FI, suivis par une phase d'extinction.

Les 4 programmes de renforcement classiquement utilisés ne donnent pas les mêmes profils évolutifs. Les programmes en ratio favorisent de hautes performances, qui s'éteignent de façon abrupte. Les niveaux de performance sont plus modérés sous les programmes imposant un intervalle de temps. L'extinction suit une évolution en continu suite à un programme en intervalle variable.

VR : ratio variable, FR : ratio fixe, VI : intervalle variable, FI : intervalle fixe.

(d'après Reynolds, 1975)

Au delà de patrons différents d'activité opérante, ces programmes de renforcement induisent des différences importantes sur un facteur, à savoir la corrélation entre taux d'appuis et taux de récompenses. Dans un article désormais classique (Dickinson et al., 1983 ; Dickinson, 1985), Dickinson avait proposé que les programmes en intervalle accélèrent la mise en place des habitudes (voir également Yin et Knowlton, 2006). Dans cette étude, cet auteur avait représenté les variations des taux d'appuis en fonction des taux de récompense (voir figure 9). Comme le montre cette figure, au début d'un entraînement effectué en intervalle, la fonction en retour correspondant à la corrélation entre taux d'appuis et taux de récompense est très importante. En effet, puisque l'animal appuie peu sur le levier, il expérimente très fortement la corrélation. Néanmoins, avec l'entraînement, la perception de cette corrélation diminue rapidement. Lorsque l'entraînement s'effectue selon un programme en Ratio, la perception de la corrélation demeure forte plus longtemps puisque le taux de récompense dépend directement du taux d'appuis. Selon cette analyse, une réponse de type habitude apparaît lorsque l'animal atteint un taux d'activité important pour lequel les variations de la corrélation sont faibles.

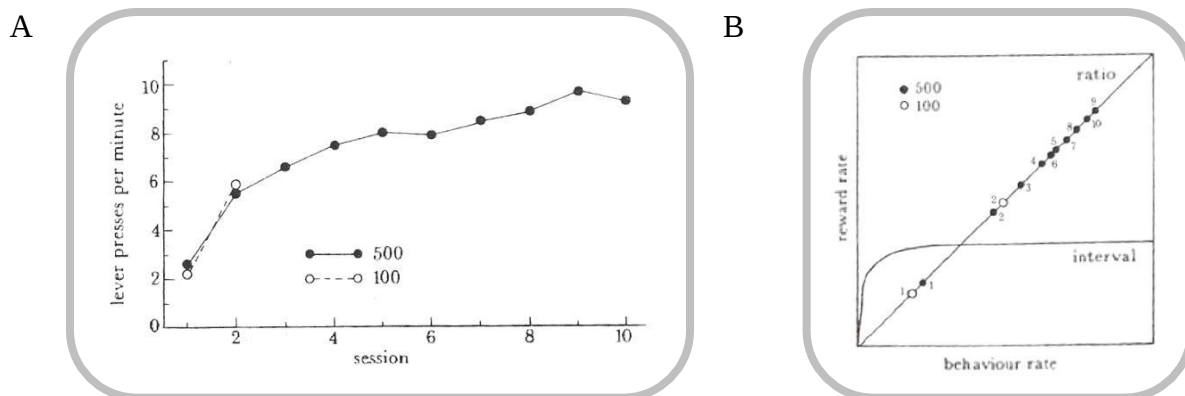


Figure 9 Evolution des taux d'appuis (A) selon l'intensité de l'entraînement instrumental et évolution estimée des taux de récompenses en fonction des taux d'appuis (B).

Au début de l'entraînement en ratio fixe, les taux d'appuis augmentent rapidement ainsi que les taux de récompenses. Néanmoins, avec le surentraînement il y a peu de variation dans le comportement et donc dans ses conséquences. Une telle situation serait propice au développement des habitudes. Sous un programme en intervalle variable, le taux de récompenses devient rapidement stable et n'évolue pas en fonction du taux de réponses. Un tel programme favoriserait le passage en mode habitude.

Disques blancs : 100 appuis récompensés ; disques noirs : 500 appuis récompensés (d'après Dickinson, 1985).

Bien qu'extrêmement séduisante, cette proposition n'a pas trouvé à l'heure actuelle de démonstration définitive, et des travaux systématiques sont requis. Dans une étude récente néanmoins, DeRusso et al. (2010) ont montré que l'incertitude liée aux intervalles variables d'un programme en VI pourrait constituer un facteur déterminant pour le contrôle comportemental.

Quantité et mode de l'entraînement

Au delà du programme de renforcement, il peut sembler intuitif de prendre en compte la quantité d'entraînement. Un examen approfondi de la littérature indique une grande disparité dans le nombre de séances d'entraînement comme l'illustrent les tableaux de la figure 10. En plus de la quantité d'entraînement, reflété dans le nombre de séances, Adams, dès 1982, avait noté que le mode d'entraînement avait un impact considérable sur le contrôle par les actions ou les habitudes. En particulier, cet auteur avait montré que la réponse était sensible à la dévaluation dans des conditions d'entraînement massé mais pas dans des conditions d'entraînement espacé dans le temps.

FR ou VR → Action dirigée		
Nb séances	Procédure	Références
2	FR1, VR9	Adams & Dickinson 1981
2	2x FR1	Adams 1982
2	FR1, VR9	
1	1x VR10	
7	FR1, 3x VR5, 3x VR10, 3x VR20	Corbit & Balleine 2003
6	FR1, 5x VR20	Killcross & Coutureau 2003
11	2x FR1, 3x VR5, 3x VR10, 3x VR20	Ostlund & Balleine 2005
8	2x FR1, 2x VR5, 2x VR10, 2x VR20	Yin et al. 2005a
7	3x FR1, VR10, 3x VR20	Hilario et al. 2007

FR ou VR → Habitudes		
Nb séances	Procédure	Références
3	3x FR1	Wilson et al. 1981
9	9x FR1	Adams 1982
10	1x FR1, 9x VR9	
10	10x VR10	

FI ou VI → Action dirigée		
Nb séances	Procédure	Références
5	FR1, 2x VI30, 2x VI60	Holman 1975*
14	4x FI1, 6x FI5, 14x FI20	Hammond 1980
3	FR1, 2x VI60	Colwill & Rescorla 1988
17	FR1, VI30, 15x VI60	
8	VI2, VI15, VI30, 5x VI60	Balleine & Dickinson 1991
8	VI2, VI15, VI30, 5x VI60	Lopez et al. 1992
4	VI2, VI15, 2x VI30	Dickinson et al. 1995
3	FR1, VI30, VI60	Colwill 2006
5	2x FR1, 3x VI30	Nelson & Killcross 2006
6	2x FR1, 2x VI30, 2x VI60	Coutureau et al. 2009
11	3x FR1, 2x FI20, 6x FI60	DeRusso et al. 2010
5 à 21	VI30, (4, 8, 12 ou 20)x VI60	Jonkman et al. 2010*
5	VI30, 4x VI60	

VI → Habitudes		
Nb séances	Procédure	Références
5	FR1, 2x VI30, 2x VI60	Holman 1975*
12	VI2, VI15, 10x VI30	Dickinson et al. 1995
5	VI2, VI15, 2x VI30	
12	VI2, VI15, 10x VI30	Coutureau & Killcross 2003
23	FR1, 22x VI60	
21	FR1, 15x VI60, 5x VR20	Killcross & Coutureau 2003
7	FR1, 2x VI10, 2x VI30, 2x VI60	Yin et al. 2004
8	4x FI20, 2x VI30, 2x VI60	
7	3x VI15, 2x VI15, VI30, VI60	
10	5x FI20, VI30, 4x VI60	Yin et al. 2006
7	3x FR1, VI10, 3x VI20	Hilario et al. 2007
9	FR1, VI15, VI30, VI45, 5x VI60	Hitchcott et al. 2007
10	3x FR1, VI20, 6x VI60	DeRusso et al. 2010
5 à 21	VI30, (4, 8, 12 ou 20)x VI60	Jonkman et al. 2010*

Figure 10 Tableaux indiquant la nature de la réponse instrumentale obtenue suite à des entraînements appliquant des programmes de renforcements selon des ratios ou des intervalles.

Dans la littérature, les études faisant état d'une réponse instrumentale de type habitude emploient de façon préférentielle des programmes de renforcement selon des intervalles variables (VI). En revanche, le nombre de séances nécessaires varie d'une étude à l'autre. L'action dirigée est obtenue suite à des procédures d'entraînement diversifiées employant aussi bien des programmes à ratios (fixes, FR et variables, VR) que des programmes à intervalles (fixes FI et variables).

*La présence d'un astérisque signale une procédure d'entraînement qui aboutit à une réponse dirigée ou une réponse de type habitude selon la méthode choisie pour explorer sa nature.

Facteurs de l'environnement

Le rôle des facteurs environnementaux dans la sélection du mode de contrôle de la réponse est à l'heure actuelle largement inconnu. Néanmoins, une étude récente (Dias-Ferreira et al., 2009) a montré que le stress chronique favorisait l'expression des habitudes.

Par ailleurs, il semble que le contexte joue un rôle de premier plan dans les mécanismes de contrôle. Le contexte pourrait intervenir à plusieurs niveaux. Premièrement, il pourrait jouer un rôle de signallement. C'est probablement une étude de Killcross et Coutureau (2003) qui illustre le mieux cette notion. Ces auteurs avaient entraîné de manière intensive une réponse dans un contexte et de manière modérée une autre réponse dans un autre contexte. Au

moment du test, et après dévaluation, ces auteurs avaient observé que la réponse étaient de type habitude dans le contexte d'entraînement intensif et de type action dans l'autre contexte. Deuxièmement, le contexte pourrait jouer un rôle plus complexe d'intégration. Par exemple, un ensemble de données de la littérature suggèrent que le contexte de dévaluation joue un rôle très important dans la sensibilité à la dévaluation (Yin et al., 2006). Cette notion sera explorée en détail dans le chapitre 2.

Codage de la valeur

Un nombre important d'études ont montré ces dernières années que les processus de récompense à la base de l'établissement de la valeur dépendent de la capacité du sujet à évaluer les propriétés affectives et motivationnelles du but (Balleine et Dickinson, 1992 ; Dickinson et Balleine, 2002). Cette capacité est la résultante d'un apprentissage communément nommé apprentissage incitatif (voir Dickinson et Balleine, 1994). Par exemple, les rats ne modifient pas immédiatement leur choix de réponse associé avec une récompense plus (ou moins) calorique lorsque leur état de privation est augmenté (ou au contraire diminué). Au contraire, ils modifient leur réponse après être entrés en contact direct avec la récompense dans ce nouvel état motivationnel (Balleine et Dickinson, 1992 ; Balleine et al., 1995). Ce type d'apprentissage incitatif est nécessaire à la performance d'une réponse dirigée qui constitue de fait une « force brute » par laquelle l'action est dirigée vers l'option avec la plus forte valeur (Niv et al., 2006). Nous l'avons vu précédemment, ce type de processus est à l'œuvre dans les effets de dévaluation par satiété sensorielle (Dickinson et Balleine, 1998) ou aversion gustative mais il joue assurément un rôle de premier plan dans les effets de changements motivationnels dans des domaines variés allant des récompenses sexuelles (Everitt et Stacey, 1987 ; Everitt et al., 1987) ou alimentaires (Lopez et al., 1992). S'il ne fait aucun doute que l'apprentissage incitatif joue un rôle dans l'ajustement des réponses dirigées, son rôle dans les habitudes est sujet à discussion (voir Dickinson et al., 1995).

Processus en retour

Au delà de l'aspect formel, les processus de choix résultant de l'initiation de la réponse activent un processus d'évaluation. De fait, la réponse (R) active une représentation du résultat (ou *Outcome* O), qui elle-même peut influencer sur l'action ou la séquence d'actions. Ces processus R-O/O-R sont actuellement considérés comme essentiels à l'adaptation de l'animal à son environnement (voir Balleine et O'Doherty, 2010). De plus, lorsque la récompense est

physiquement présente (comme c'est le cas par exemple dans des tests de réacquisition où la récompense préalablement dévaluée est présentée à l'animal), ce processus en retour agit de manière instantanée pour influencer sur l'émission de la réponse (Adams, 1982 ; Corbit et Balleine, 2003 ; Killcross et Coutureau, 2003).

2. Aspects neurobiologiques des réponses dirigées vers un but et des habitudes

Généralités: Aspects phylogénétiques

L'encéphale des mammifères présente une morphologie externe qui est globalement commune à l'ensemble des vertébrés. En effet, quel que soit l'ordre considéré, l'encéphale comprend les trois parties suivantes : un tronc cérébral, un cervelet et un cerveau (cf. figure 11). Le cerveau est composé de deux régions ayant la même origine embryologique : le télencéphale (cerveau hémisphérique) et le diencephale (cerveau intermédiaire). Le diencephale, région comprenant différents noyaux thalamiques, unit le tronc cérébral au télencéphale. Ce dernier est scindé en deux hémisphères qui comprennent en surface un cortex cérébral et à leur base des noyaux gris centraux.

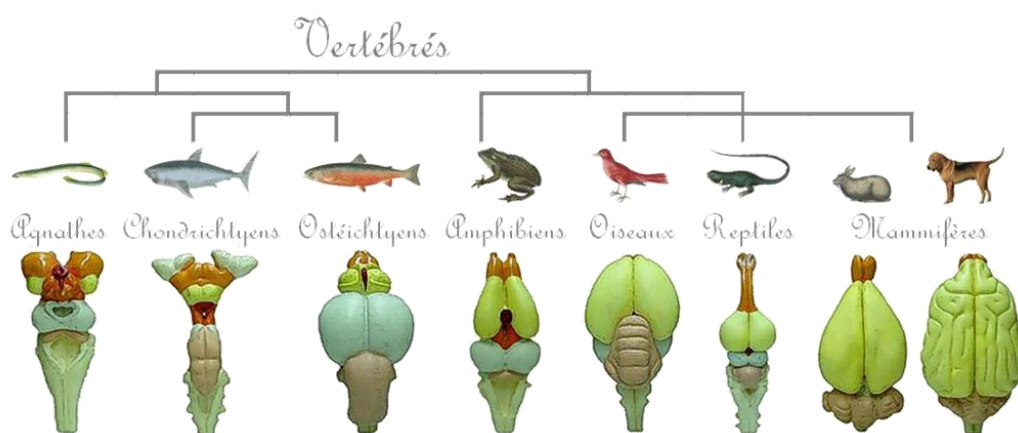


Figure 11 Phylogénie de l'encéphale des vertébrés.

L'encéphale des vertébrés est organisé selon un schéma commun : le tronc cérébral (blanc) surmonté du cervelet (saumon) est lié au diencephale (bleu) lui-même raccordé aux structures du télencéphale (jaune). Néanmoins, au gré des divergences évolutives, chaque espèce a adopté les innovations optimisant sa survie et sa propagation. Les hémisphères télencéphaliques se sont développés de plus en plus jusqu'au recouvrement total du diencephale (thalamus et hypothalamus) chez les mammifères. L'apparition du néocortex et son expansion aboutissent chez les plus grands mammifères à la formation de circonvolutions corticales ; elles permettent un accroissement de la masse cérébrale malgré une limitation de l'espace imposée par la boîte crânienne. Au contraire, les lobes olfactifs (orange) occupent un volume de plus en plus petit chez les espèces développant préférentiellement le sens de la vision. En effet, chez les mammifères, le traitement des informations visuelles est pris en charge par le cortex cérébral, contrairement aux oiseaux qui conservent des lobes optiques très développés. (Représentation de animaux d'après différentes planches naturalistes des XVIIIème et XIXème siècles ; Illustrations des encéphales obtenue par transformation de photographies tirées de la plateforme collégiale DECclic www.deccllic.qc.ca).

Les mammifères se distinguent des autres vertébrés par la différenciation d'un néopallium entre le pallium olfactif et le pallium hippocampique. Il s'agit du néocortex qui est organisé en six couches cellulaires. Comme l'illustre la figure 12, le cerveau des mammifères présente une forte hétérogénéité à la fois en taille et en aspect. De fait, la lissencéphalie observée chez le rongeur a peu à peu laissé sa place à une complexification croissante liée à la gyrencéphalie. Au-delà de la complexification d'aspect, il semble évident que des parties entières du cerveau soient développées de manière prédominante en fonction du type de capacités cognitives et sensorielles de l'espèce. On notera par exemple la place prépondérante des bulbes olfactifs chez le rongeur.

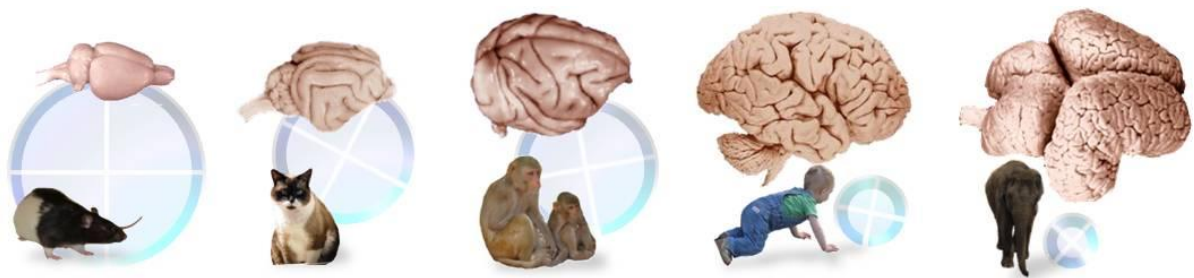


Figure 12 Phylogénie du cerveau mammalien (*Rat Long Evans, Chat domestique, Macaque Rhésus, Humain, Eléphant d'Asie* (les disques de 5 cm de diamètre sont à la même échelle que les cerveaux auxquels ils sont appariés). (photographies des cerveaux d'après www.brainmuseum.org ; photographies des animaux par DTTY, sauf macaques par Mleciu K2.)

L'apparition de la gyrencéphalie chez les grands mammifères a permis de s'affranchir des limites spatiales imposées par la boîte crânienne. L'expansion du néocortex est favorable à la différenciation et à la multiplication d'aires à modalités fonctionnelles diverses ; on les regroupe usuellement selon leur rôle moteur, sensoriel ou multimodal. Les aires multimodales sont dites associatives, elles permettent l'intégration de multiples informations. La surface occupée par les aires associatives s'est de plus en plus étendue au cours de l'évolution. Ainsi, elles constituent 80% du néocortex humain et seulement 10% de la masse néocorticale chez le rat. Les aires associatives situées rostralement sont celles dont la taille a augmenté de façon la plus disproportionnée. Dans la littérature, cette expansion est corrélée à l'émergence d'aptitudes cognitives de plus en plus complexes (pour revue consulter : Fuster, 2002 ; Striedter, 2006 ; Rakic, 2009).

Lobe frontal chez le primate

Chez l'homme, le cortex préfrontal constitue la partie antérieure du lobe frontal. De manière schématique, cette région comprend trois territoires situés médialement, ventralement et dorsolatéralement. La figure 13 fait état des différentes délimitations et dénominations que nous utiliserons ultérieurement. Elle propose comme repère spatial la cartographie cytoarchitectonique établie par Brodmann (1909). Comme le montre cette figure, le cortex préfrontal dorsolatéral comprend un ensemble de régions de Brodman (Brodman Area, BA) situées sur les faces dorsales (BA 8, 9) et latérales (BA 46). Au contraire, le cortex orbitofrontal est situé plus ventralement et comprend les régions BA 47, 12 et 11. Enfin, le cortex préfrontal médian comprend de la face ventrale vers la face dorsale, les régions BA 25, l'ensemble des régions du cortex cingulaire BA 32, BA 24 (a'/b', 24 a/b; c/c') (voir Walton et al., 2007).

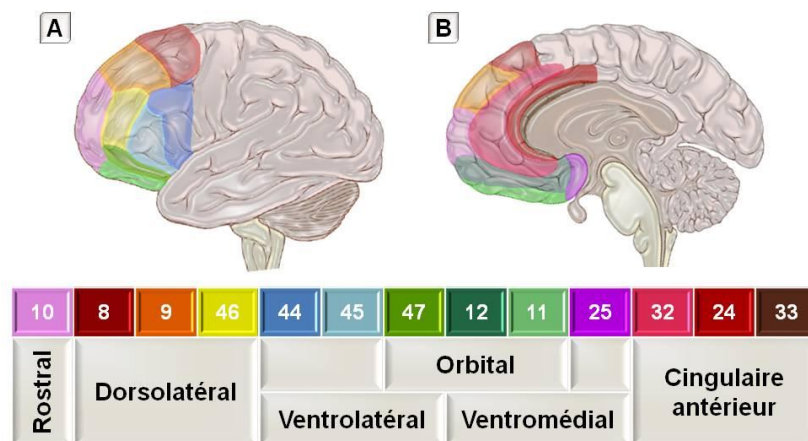


Figure 13 Les régions du cortex préfrontal de l'Homme, vue latérale (A) et coupe sagittale médiane (B) de l'encéphale.

Le code couleur appliqué correspond aux aires de Brodmann. La lecture du tableau permet de faire correspondre le regroupement de plusieurs aires à la dénomination des régions corticales préfrontales. La délimitation de ces régions peut varier quelque peu d'une publication à l'autre. Le cortex cingulaire malgré son état agranulaire est considéré comme faisant partie du cortex préfrontal dans sa partie antérieure. L'aire 25 ainsi que la région corticale qui lui est directement adjacente (ici une partie de l'aire 12 ou parfois indiquée comme 32 a') sont proposées comme étant de possibles analogues aux régions prélimbique et infralimbique du cortex préfrontal du Rat. (Images obtenues par transformations d'illustrations de l'encéphale par Patrick J. Lynch.)

Anatomie du cortex préfrontal chez le rongeur

- Topologie

Chez le Rat, la partie la plus antérieure du cortex peut être considérée comme un homologue du cortex préfrontal des primates. L'existence même d'un cortex préfrontal chez le rongeur, a été longtemps débattue (voir Preuss 1995 ; Wise, 2008) mais les données actuelles plaident largement pour l'existence de régions fonctionnellement comparables chez le Rat.

L'existence d'une homologie, initialement proposée sur la base de la connectivité avec le thalamus dorsomédian (Rose et Woolsley, 1948), est de nos jours basée sur une approche multicritère. Ces critères (Uylings et al., 2003) correspondent (1) au développement embryonnaire, (2) aux caractéristiques cytoarchitectoniques (chez les espèces phylogénétiquement proches), (3) aux substances neuro-actives et leurs récepteurs, (4) à la connectivité et la densité neuronale et enfin (5) aux propriétés fonctionnelles.

Plusieurs zones distinctes ont ainsi pu être identifiées chez le Rat. La figure 14 illustre la partition du PFC en une région latérale (cortex insulaire antérieur IC), une région orbitale (cortex orbital CO) et une région médiane (cortex préfrontal médian, mPFC). Le cortex préfrontal médian (mPFC) se subdivise selon un axe dorsoventral en un cortex précentral médian (PrCm ou Fr2 ou AGm), un cortex cingulaire antérieur (Cg), un cortex prélimbique (PL) et enfin un cortex infralimbique (IL) (Gabbott et al., 1997). Au sein de la vaste région du cortex orbital (OC) on distingue les sous-régions médiale (MO), latérale (LO) et ventrolatérale (VLO) (Schilman et al., 2008).

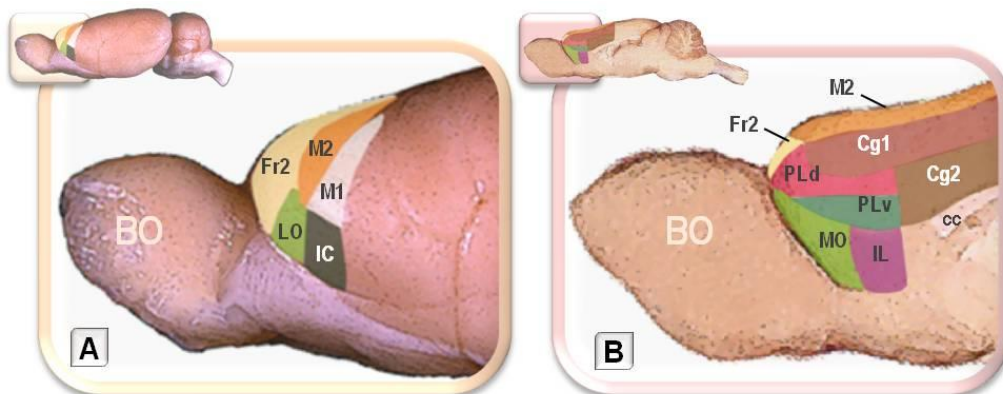


Figure 14 Les régions du cortex préfrontal du Rat, vue latérale (A) et coupe sagittale (B) de l'encéphale.

Le cortex préfrontal du Rat se trouve juste à l'arrière du bulbe olfactif (BO). Ont été délimités à la surface corticale : le cortex précentral médian (Fr2), la partie latérale du cortex orbital (LO) ainsi que la partie antérieure du cortex insulaire (IC). On distingue également les parties antérieures des régions M2 et M1, dénommées aussi Fr1 et Fr3, parfois considérées comme faisant partie du PFC du Rat. La vue médiane correspond à une coupe située latéralement à 0,40 mm du bregma. Elle permet de localiser les positions par rapport au corps calleux (cc) du cortex cingulaire (Cg1 et Cg2), du prélimbique dorsal et ventral (PLd et PLv), de l'infralimbique (IL) et de la partie médiane du cortex orbital (MO). (Illustration obtenues par transformations de photographies de l'encéphale par Adam C. Puche.)

Connectivité

Les régions du cortex préfrontal sont fortement interconnectées. Néanmoins, la principale particularité des ces régions est d'être en relation étroite avec un ensemble de structures appartenant au système limbique (voir Vertes, 2006) et aux ganglions de la base. Les ganglions de la base comprennent un ensemble de structures : le striatum, le pallidum (externe et interne) et le noyau sous thalamique. Les régions contenant les corps cellulaires

dopaminergiques (substance noire et aire tegmentale ventrale) font également partie de cet ensemble de structures. Le striatum peut être divisé en une partie ventrale qui correspond au noyau accumbens et une partie dorsale qui correspond au noyau caudé-putamen. Un ensemble important de données ont montré que le noyau accumbens pouvait être subdivisé en une partie coquille (shell), une partie cœur (core) et une partie pole rostral (Meredith et al., 2008). De même, le noyau caudé putamen comprend une partie dorsomédiane et une partie dorsolatérale (Reiner et al., 1998).

- Afférences

Les régions du cortex préfrontal reçoivent des entrées d'un vaste ensemble de structures (Hoover et Vertes, 2007). Contrairement au cortex prélimbique, le cortex infralimbique reçoit peu d'afférences corticales, elles proviennent préférentiellement des cortex prélimbique et insulaire. Au contraire, le cortex prélimbique reçoit des informations d'origine plus vaste. Localement, elles proviennent des cortex orbitofrontal (médian et ventral), infralimbique et cingulaire. Des afférences corticales d'origine plus éloignées arrivent des cortex insulaire postérieur, périrhinal et enthorinal. Les afférences du système limbique proviennent de l'amygdale et de l'hippocampe. En particulier, la région CA1 projette massivement sur les cortex prélimbique et infralimbique (voir Jay et Witter, 1991 ; Laroche et al., 2000 ; Heidbreder et Groenewegen, 2003). Ces régions sont également des régions de fortes intégrations neuronales puisqu'elles reçoivent des afférences importantes du système noradrénergique en provenance du Locus Coeruleus (Nutt et al., 1997), dopaminergique en provenance de l'aire tegmentale ventrale (Seamans et Yang, 2004), sérotoninergique en provenance du raphé.

- Efférences

Les systèmes de projection du cortex prélimbique et infralimbique sont largement dissociés (Vertes, 2004). Le cortex infralimbique projette de façon diffuse sur de nombreuses régions. Les projections locales touchent les cortex prélimbique, orbitofrontal et cingulaire. Plus éloignées elles atteignent les structures limbiques telles que l'hypothalamus, les différents noyaux de l'amygdale, le thalamus et divers noyaux rentrant en jeu dans des fonctions viscérales comme le noyau parabrachial (cf. figure 15, panneau A). Concernant les ganglions de la base, les projections sont majoritairement dirigées vers la partie ventrale du striatum et concernent en quantité moindre le striatum dorsal. En comparaison au cortex infralimbique, les projections du cortex prélimbique sont restreintes et majoritairement en

direction des régions qui projettent sur lui (cf. figure 15, panneau B). Ainsi, elles atteignent les cortex orbitofrontal, insulaire et cingulaire ainsi que l'aire tegmentale ventrale. La majorité des projections à destination sous-corticale sont en direction des ganglions de la base et assimilés, notamment le striatum dorsomédian, le noyau basolatéral de l'amygdale et le noyau accumbens.

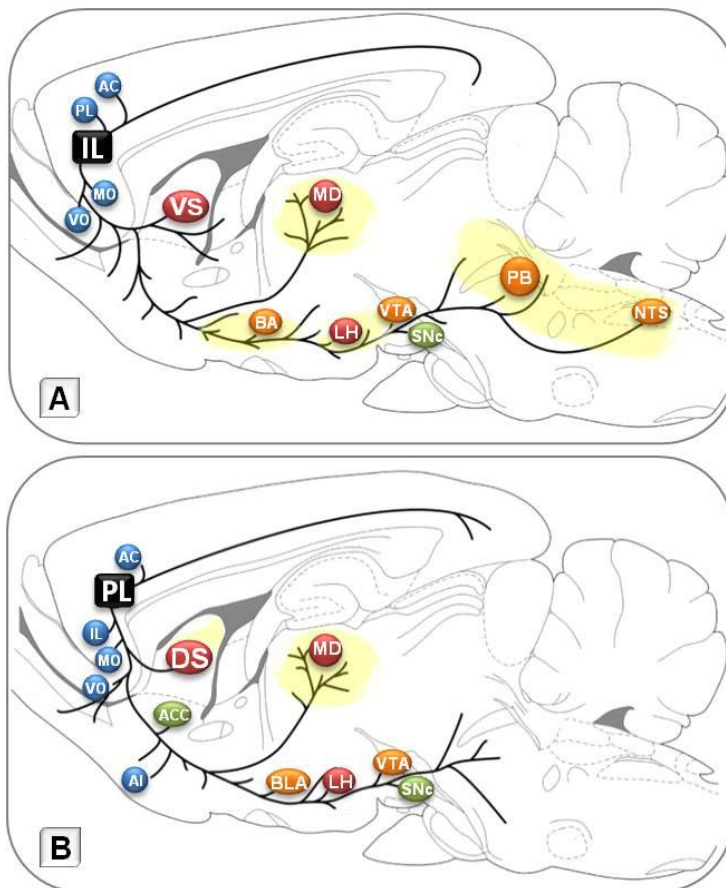


Figure 15 Représentation schématique des projections corticales infralimbique (A) et prélimbique (B).

Intensité des projections : faible en vert, moyenne en orange et forte en rouge. Structures préfrontales cibles en bleues. Abréviations : cortex cingulaire AC, noyau accumbens ACC, cortex insulaire agranulaire AI, noyaux de l'amygdale BA et basolatéral BLA, striatum dorsal DS, cortex infralimbique IL, aire hypothalamique latérale LH, cortex orbitofrontal médian et ventral MO et VO, thalamus médiodorsal MD, noyaux du tractus solitaire NTS, noyaux parabrachiaux PB cortex prélimbique PL, substance noire pars compacta SNc, striatum ventral VS, aire tegmentale ventrale VTA.

(Figure d'après Vertes, 2004 modifiée à partir des données de Gabbott et al., 2005.)

- Principes d'organisation

Les données de connectivité actuelles font apparaître une organisation topographique remarquable au sein de boucles cortico-striato-corticales, comme l'illustre la figure 16. (voir Voorn et al., 2004). Cette organisation suggère une topographie selon laquelle à un axe dorsoventral des régions frontales correspond un axe latéro-médian au sein du striatum. Ainsi les projections des aires préfrontales les plus ventrales, l'infralimbique et l'insulaire, projettent massivement en direction du striatum latéral et ventral. La topographie des projections du prélimbique respecte également son organisation en deux sous régions dorsale et ventrale. Enfin, le striatum latéral est connecté de façon préférentielle avec le cortex

sensorimoteur. Cette topographie est également conservée au niveau des structures de projection du striatum comme le thalamus et l'amygdale.

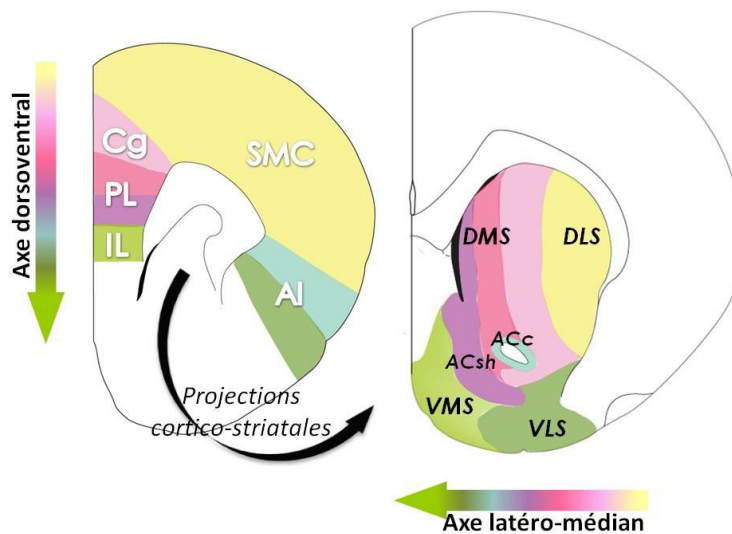


Figure 16 Topographie des projections corticales préfrontales sur les régions striatales. Aires corticales : insulaires latérale et ventrale AI, cingulaire Cg, infralimbique IL, prélimbique PL, sensorimoteur SMC. Aires striatales : dorso-latérales DLS, dorso-médianes DMS, ventro-latérales VLS, ventro-dorsale VMS. Cœur et coquille de l'accumbens ACc et ACsh. (figure inspirée de Voorn et al., 2004 et réalisée à partir de planches anatomiques de Paxinos et Watson, 1998.)

Ainsi, plusieurs études ont décrit les circuits neuronaux impliquant les ganglions de la base et le cortex cérébral comme organisés en boucle spatialement distinctes et parallèles selon un chemin cortico-striato-pallido-thalamo-cortical (Alexander et al., 1986 ; Donoghue et Herkenham, 1986 ; Parent et Hazreti, 1995).

Cette analyse se base sur le postulat que, en fonction de l'information traitée, des boucles différentes seraient mobilisées de manière à intégrer en continu les informations émotionnelles, cognitives et motrices. Un corpus important de données a montré que ces boucles sont en interaction directe ou indirecte (Homme : Joel et Wiener, 1992 ; Francois et al., 1994 ; Calzavara et al., 2005 ; Rat : Mogenson et al, 1980 ; Heimer et al, 1982 ; Smith et al., 1998 ; Kolomiets et al., 2001). C'est probablement aux travaux de Haber et collaborateurs que l'on doit la proposition la plus originale concernant les interactions entre territoires. En effet, ces auteurs (Haber et al., 2000) ont démontré l'existence d'une organisation en spirale qui pourrait rendre compte des changements de plasticité au sein des différents territoires (voir Belin et al, 2009).

Circuits des actions et habitudes

Les données de la littérature suggèrent que des circuits fronto-striataux contrôlent un certain nombre de fonctions cognitives communément regroupées sous le terme de fonctions exécutives (voir Dalley et al., 2004). Ces fonctions concernent le maintien des informations

en mémoire de travail (Gisquet-Verrier et Delatour, 2006), la capacité de changer de stratégie (Floresco et al., 2009), l'adaptation en fonction de la représentation de l'effort (Walton, 2009) ou encore la capacité à inhiber une réponse (Amalric et al., 1995 ; Robbins, 2007). De fait, la recherche concernant les circuits neuronaux des actions et habitudes s'est naturellement orientée vers l'étude de tels circuits (voir Yin et al., 2008 ; Yin et al., 2009 ; Balleine et O'Doherty, 2010).

Cortex préfrontal médian

En particulier, ces expériences ont montré que le cortex prélimbique jouait un rôle crucial dans la réponse dirigée vers un but. En effet, des lésions excitotoxiques de ce cortex altèrent l'adaptation de l'animal à la fois à des changements de valeur et à des changements de contingence (Balleine et Dickinson, 1998a ; Corbit et Balleine, 2003 ; Kilcross et Coutureau, 2003 ; Ostlund et Balleine, 2005 ; Coutureau et al., 2009). Le rôle du cortex prélimbique est spécifique puisque des lésions soit de la partie infralimbique (Killcross et Coutureau, 2003), soit du cortex orbitofrontal sont sans effet sur le conditionnement opérant (Ostlund et Balleine, 2005). Les données disponibles sur ce thème suggèrent néanmoins que des interactions étroites entre régions du cortex préfrontal jouent un rôle important dans la coordination des actions et habitudes. En effet, dans des conditions d'entraînement intensif où la réponse est insensible à la dévaluation de la récompense et donc contrôlée par le système habitude, des lésions du cortex infralimbique (Killcross et Coutureau, 2003) ou des inactivations de cette région au moment du test (Coutureau et Killcross, 2003) restaurent la sensibilité à la dévaluation. Il semble que le système dopaminergique mésocortical intervienne également dans la coordination des actions et des habitudes. En effet, Hitchcott et al (2007) ont montré que l'injection de dopamine restaure également la sensibilité à la dévaluation. Par ailleurs, des données récentes obtenues au laboratoire ont montré une dissociation remarquable entre représentation de la valeur et représentation de contingence. En effet, des lésions sélectives (par injection intra-préfrontal de 6-OHDA) des terminaisons dopaminergiques du cortex prélimbique (Naneix al., 2009, voir également Lex et Hauber, 2010a) altéreraient la sensibilité des animaux à des changements de contingence mais pas à des changements de valeur.

Striatum

De manière parallèle, un ensemble important de données accumulées par le groupe de B. Balleine a montré que le contrôle de la réponse instrumentale dépendait du fonctionnement

normal des territoires striataux (voir Yin et al., 2009 pour une revue récente). Dans un ensemble de données, ce groupe a montré que des lésions de la partie médiane du striatum entraînaient un déficit de réponse dirigée vers un but. Néanmoins, ces auteurs ont également pris en compte la dimension temporelle et spatiale du contrôle comportemental. En effet, ils ont montré d'une part que c'était l'intégrité de la partie postérieure du striatum dorsomédian (Yin et al, 2005) au moment de l'acquisition qui était nécessaire pour une réponse dirigée vers un but normale (Yin et al., 2005b, voir également Corbit et Janak, 2010). Il semble également que des processus de plasticité précoce se mettent en place dans cette région cérébrale dès le début de l'apprentissage puisque une étude de Shiflett et al (2010) a récemment montré des changements précoces des niveaux de la kinase ERK (extra cellular signal regulated kinase). De manière très intéressante, il a également été montré par ce groupe que l'intégrité de la partie dorsolatérale du striatum était requise dans la production d'une réponse de type habitude (Yin et al., 2006). Enfin, Faure et collaborateurs (Faure et al., 2005) ont montré que des lésions des terminaisons dopaminergiques mésostriatales empêchaient la mise en place des habitudes. Les données concernant le striatum ventral sont moins abondantes. Néanmoins, dans une série d'expériences, Corbit et al. (2001) ont montré que des lésions de la partie coquille du noyau accumbens modifiaient l'adaptation à des changements de valeur. Bien que n'utilisant pas le cadre formel des apprentissages instrumentaux (voir Kelley et al., 2005 ; Lex et Hauber, 2010b) un ensemble d'autres données suggèrent que le noyau accumbens jouent un rôle déterminant durant les premières étapes de l'apprentissage.

Régions temporales

Dans la littérature actuelle, l'effort principal a concerné l'étude des régions préfrontales et striatales. Néanmoins, les régions temporales doivent également être prises en compte. En effet, si des lésions totales de l'hippocampe sont sans effet sur la production des réponses instrumentales (Corbit et Balleine, 2000), il semble que des lésions du cortex entorhinal, qui constitue la voie d'entrée principale de l'hippocampe, altèrent l'adaptation à des changements de contingence (Corbit et al., 2002 ; Lex et Hauber 2010a). Par ailleurs, un nombre important d'études ont montré que l'amygdale jouait un rôle de premier plan, puisque des lésions de la partie basolatérale altèrent l'adaptation de la réponse à des changements de valeur (Balleine et al., 2003 ; Blundell et al., 2003 ; Corbit et al., 2005 ; Ostlund & Balleine 2008 ; Coutureau et al., 2009).

Thalamus

Les expérimentations menées au sein du thalamus ont établi que les noyaux thalamiques antérieurs n'étaient pas impliqués dans le conditionnement instrumental (Corbit et al., 2003). Au contraire, des lésions du noyau thalamique médiodorsal (MDT) abolissent la sensibilité à la dévaluation de la récompense et à la dégradation de la contingence (Ostlund et Balleine, 2008). Ce résultat a été depuis conforté par une étude récente menée par Pickens (voir Pickens et al., 2008)

Neuromodulation

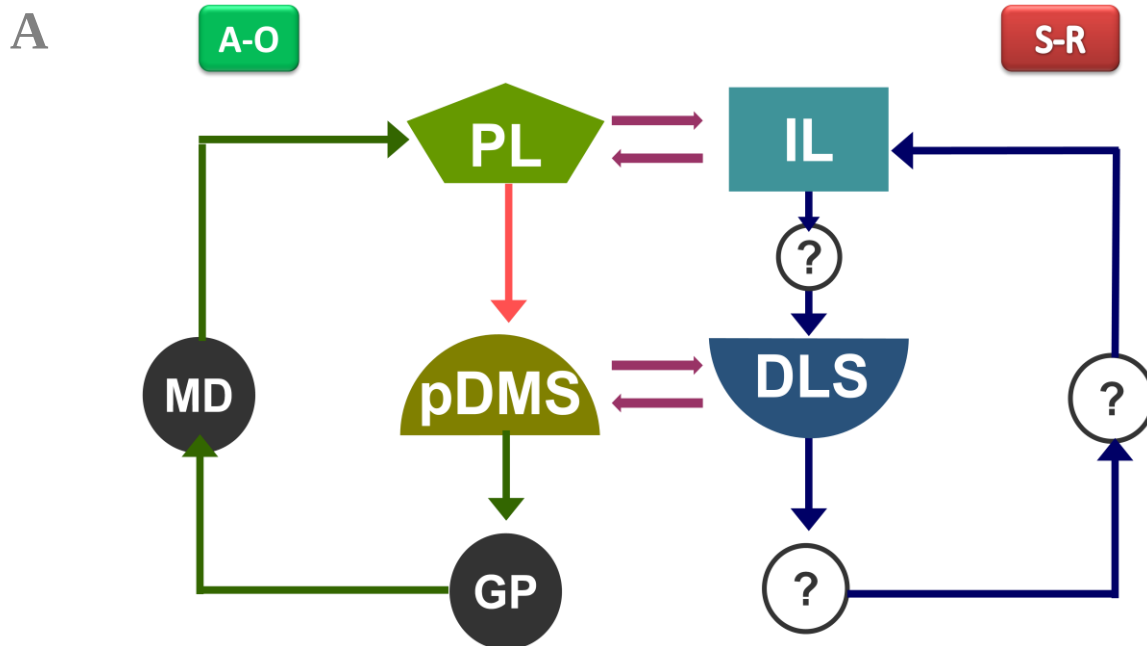
La coordination des actions et habitudes dépend également de l'histoire pharmacologique du sujet. En particulier, deux études ont montré que la sensibilisation du cerveau aux psychostimulants favorisait le contrôle de la réponse par le système habitude (Nelson et Killcross, 2006 ; Nordquist et al., 2007).

Connectivité fonctionnelle

Les données d'exclusion cérébrale obtenues par lésion réversibles ou permanentes ne permettent pas d'accéder à la notion d'ensemble de structures en interaction. De ce point de vue, l'utilisation d'une approche par imagerie cellulaire est susceptible d'apporter des informations importantes. A ce jour, seules deux études sont disponibles. Elles ont toutes deux montré l'activation de régions striatales et corticales en fonction de l'entraînement (Hernandez et al., 2006 ; Faure et al., 2006).

Circuits fonctionnels/Synthèse

Les données synthétisées dans la figure 16 indiquent que la coordination des actions et des habitudes s'effectue au sein de circuits corticostriataux dissociés.



B

Régions		Action dirigée		Habitudes	Références
		Dégradation contingence	Dévaluation	Surentrainement	
Cortex préfrontal	PL	X	X	•	Balleine et Dickinson (1998) ; Corbit et Balleine (2003) ; Kilcross et Coutureau (2003) ; Hitchcott et al. (2005) ; Ostlund et Balleine (2005) ; Coutureau et al. (2009) ; Naneix et al. (2009).
	IL	•	•	X	Coutureau et Killcross (2003) ; Killcross et Coutureau (2003).
	OFC	X	•		Ostlund et Balleine (2007) ; Pickens et al. (2003).
	AI	•	X		Balleine et Dickinson (2000)
Ganglions de la base	pDMS	X	X	•	Yin et al. (2005a) ; Yin et al. (2005b) ; Shiflett et al. (2009).
	DLS	•	•	X	Yin et al., (2004) ; Faure et al. (2005) ; Yin et al. (2005) ; Shiflett et al. 2009
	NACsh	•	•		Corbit et al. (2001).
	NACco	•	X		Corbit et al., (2001).
Thalamus	MDT	X	X		Corbit et al. (2003) ; Ostlund et Balleine (2008) ; Pickens (2008).
	ANT	•	•		Corbit et al. (2003)
Système limbique	BLA	X	X		Balleine et al. (2003) ; Blundell et al. (2003) ; Corbit et al. (2005) ; Ostlund & Balleine (2008) Coutureau et al. (2009)
	HPC	•	•		Corbit et Balleine (2000)
	EC	X	•		Corbit et al. (2002)

Figure 16 Représentation schématique des circuits corticaux striaux impliqués dans les actions (A-O) et les habitudes (S-R) (A) et tableaux de synthèse (B).

Les croix rouges signalent un déficit suite à une manipulation de la région cérébrale (lésion, inactivation...). Les points blancs indiquent une absence d'effet.

Abbréviations : cortex insulaire agranulaire AI, noyau antérieur thalamique ANT, noyau basolatéral de l'amygdale BLA, striatum dorsolatéral DLS, cortex entorhinal EC, globus pallidus GP, hippocampe HPC, cortex infralimbique IL, thalamus médiodorsal MDT, cœur et coquille du noyau accumbens NAc et Nash, cortex orbitofrontal OFC, striatum dorsomédian postérieur pDMS, cortex prélimbique PL.

Circuits des actions et habitudes chez l'Homme

Chez l'homme, les données concernant les circuits des actions et habitudes sont naturellement beaucoup moins abondantes même si un nombre important d'études s'est attaché à décrire les zones cérébrales du traitement des récompenses (O'Doherty et al., 2001 ; Tanaka et al., 2004 ; Daw et al. 2006 ; Hampton et al. 2006 ; Kim et al., 2006). Néanmoins, en utilisant des tâches structurellement identiques à celles utilisées chez le rongeur, les données disponibles suggèrent une grande conservation des mécanismes neurobiologiques et comportementaux du contrôle comportemental. En effet, Valentin et al. (2007) ont montré que chez des sujets préalablement entraînés à toucher un stimulus visuel pour obtenir une récompense alimentaire, la dévaluation de la récompense par satiété sélective entraînait une variation de l'activité au niveau de la partie médiale et centrale du cortex orbitofrontal. Des études récentes ont également mis en évidence l'activation de ces mêmes régions lors de l'adaptation à des changements de contingence (Tanaka et al., 2008 ; Gläscher et al., 2009). Les données d'imagerie disponibles à ce jour suggèrent que actions et habitudes pourraient, comme c'est le cas chez le rongeur, solliciter des activations au sein de boucles cortico-striatales dissociées (voir Tricomi et al., 2009).

Dans cette étude, Tricomi et al. (2009) ont fait varier le degré d'entraînement chez deux groupes de sujets selon des programmes de renforcement à intervalles variables. Conformément aux données du rongeur, la réponse du sujet n'est sensible à la dévaluation que dans des conditions d'entraînement modéré. De manière remarquable, le contrôle de la réponse par le système habitude coïncide avec l'augmentation de l'activité au sein du putamen postérieur et ventral (i.e. région latérale du striatum).

3. Objectifs du travail de thèse

L'ensemble des données neurocomportementales exposé dans ce document suggère que la réponse est sous le contrôle de deux systèmes dissociables. Les premières propositions théoriques sur ce sujet considéraient que le contrôle par les actions et les habitudes opéraient de manière séquentielle, les habitudes remplaçant peu à peu les actions avec l'entraînement (Dickinson et al., 1985). Un nombre croissant de données récentes suggèrent que cette conception mérite d'être reconsidérée, et que les systèmes actions et habitudes opèrent selon un mode parallèle, et dès les premiers étapes de l'apprentissage (voir par exemple, Coutureau et Killcross, 2003 ; Killcross et Coutureau, 2003 ; Yin et al. 2005). Cette conception pourrait

correspondre à des modes distincts d'apprentissage pour ces deux systèmes, comme le suggèrent les analyses computationnelles de ce problème (voir Daw et al., 2005). Dès lors, la question n'est plus d'étudier les mécanismes par lesquels l'entraînement modifie le mode d'expression de la réponse mais plutôt d'identifier les mécanismes de sélection du système action ou habitudes. Mon travail de thèse a visé à contribuer à cette entreprise. Il s'est déroulé selon 4 étapes, qui sont exposées dans les 4 chapitres suivants.

1. Chapitre 1: Ce chapitre est purement méthodologique. Il vise à mettre au point des procédures comportementales permettant d'explorer une réponse dirigée vers un but.
2. Chapitre 2: Ce chapitre aborde une question peu explorée et vise à comprendre le rôle du contexte dans le contrôle de la réponse.
3. Chapitre 3: Dans ce chapitre, une approche invasive (inactivation réversible) est utilisée pour étudier le décours temporel de l'implication du cortex prélimbique dans les mécanismes de sélection de la réponse.
4. Chapitre 4: Enfin, dans ce chapitre, nous avons mis en œuvre une démarche corrélative visant à une étude des circuits cortico-striataux du contrôle de la réponse.

Chapitre I

Chapitre I

Introduction

Au cours des 40 dernières années, de nombreux travaux ont permis l'élaboration d'un riche corpus méthodologique concernant les paramètres expérimentaux des actions et des habitudes (Adams, 1982 ; Colwill et Rescorla, 1985 ; Balleine et Dickinson, 1998a). Les expérimentations visant à déterminer la nature d'une réponse instrumentale se développent sur 4 phases principales : une phase d'habituation, une phase de conditionnement instrumental, une phase de manipulation post-conditionnement et enfin une phase de test. Pour mieux appréhender les enjeux de chacune des expériences constituant ce chapitre, nous nous proposons en premier lieu d'introduire les notions et paramètres à prendre en compte lors de l'établissement de chacune de ces phases.

Durant la phase d'habituation, les animaux sont placés dans le dispositif où ils sont exposés de manière répétée à la récompense alimentaire afin de diminuer la néophobie (Domjan, 1977) qui caractérise cette espèce. Comme nous l'avons indiqué dans l'introduction générale, l'observation d'une réponse de type dirigé ou de type habitude peut dépendre du choix des programmes de renforcement, les programmes de renforcement à intervalle variable étant supposés favoriser un fonctionnement de type habitude par rapport aux programmes en Ratio (Dickinson, 1985). Néanmoins, une dissociation entre ces deux types de programme ne peut être établie de façon catégorique puisque de nombreuses données récemment publiées font état d'actions dirigées résultant de conditionnements appliquant un programme de renforcement à intervalle variable (Nelson et Killcross, 2006 ; Yin et Knowlton, 2006 ; Coutureau et al., 2009). De même, on observe une certaine disparité quant au nombre de séances de conditionnement nécessaires à l'obtention d'une réponse de type habitude. Ces disparités pourraient provenir du fait que la formation des habitudes dépendrait aussi de l'enchaînement des séances d'entraînement pendant le conditionnement (massées vs distribuées) ainsi que du nombre total des récompenses distribuées (Adams, 1982). Cet aspect est exploré dans les Expériences I.3 et I.4.

Comme nous l'avons souligné dans l'introduction générale, des changements de valeur par dévaluation constituent des outils puissants pour l'étude de la représentation de la récompense. Dans ce chapitre ? nous avons également étudié l'impact sur la réponse opérante de deux méthodes de dévaluation, l'aversion gustative conditionnée (Expériences I.1 et I.2) et la satiété sensorielle sélective (Expérience I.4). Enfin, la phase de test vise à mettre en évidence les effets ou l'absence d'effet de manipulations postérieures au conditionnement sur la performance instrumentale. Le sujet doit donc être mis en présence du levier, qui était absent lors de la dévaluation. Mais il est aussi nécessaire que l'appui sur le levier ne soit suivi d'aucune conséquence. En effet, ce n'est qu'en son absence qu'il est possible d'évaluer l'utilisation d'une représentation mentale de la récompense et son pouvoir incitatif (Balleine et Dickinson, 1998a et b). Enfin, il peut être opportun d'inclure également dans cette phase de test une séance durant laquelle la réponse est de nouveau renforcée, particulièrement dans l'étude des effets d'une dévaluation de la récompense (Adams et Dickinson, 1981 ; Dickinson, 1985 ; Killcross et Coutureau, 2003).

Tirant parti de ces nombreuses observations, ce premier chapitre est consacré à l'étude de 3 protocoles différents de conditionnement opérant. Les deux premières expériences sont consacrées à l'établissement d'une procédure devant aboutir à l'expression d'une action dirigée. Les deux suivantes exploreront les effets potentiels de l'intensité du conditionnement et des programmes de renforcement sur l'expression d'une réponse de type habitude.

Expérience I.1

Action dirigée et aversion gustative conditionnée

Dans cette première expérience, nous avons opté pour un conditionnement opérant de courte durée (6 séances) devant aboutir à une action dirigée vers un but. Pour tester la nature de la récompense nous mettons en œuvre une procédure de dévaluation de la récompense par aversion gustative. Cette phase a lieu dans le même contexte que l'entraînement instrumental.

A. Matériel et méthode

a. Sujets

Les sujets de cette expérience sont 16 rats mâles de la souche Long-Evans, en provenance du centre d'élevage Janvier (Le Genest-Saint-Isle). A leur arrivée, les rats pesant de 275 à 350 g, sont placés par 2 dans des cages en polycarbonate (longueur : 46 cm ; largeur : 26 cm ; hauteur : 20 cm). La température de l'animalerie est maintenue à $22^{\circ}\text{C} \pm 1$ et la phase éclairée du cycle jour/nuit s'étend de 7h à 19h. Durant une phase d'acclimatation d'au moins une semaine, les rats ont accès *ad libitum* à l'eau et à une nourriture standard pour rongeurs (granules compactés à partir d'un mélange de céréales et de solubles de poissons, A04, Scientific Animal Food et Engineering).

Deux jours avant le début des expérimentations, les rats sont pesés et manipulés afin de les familiariser à l'expérimentateur. La restriction alimentaire commence la veille de la première séance d'entraînement et se prolonge jusqu'à la fin de l'expérience. Afin de les maintenir à environ 90% de leur poids d'origine, les animaux reçoivent une dose de leur alimentation habituelle (15 g) tous les soirs après l'entraînement. Ils ont en revanche, dans leur cage d'habitation, un accès libre et non contrôlé à l'eau de boisson.

b. Dispositif expérimental

Le dispositif expérimental est constitué de 8 cages de conditionnement identiques (42.5 x 41 x 29 cm). Elles sont reliées à un ordinateur PC, via une interface électronique (Imétronic, Pessac, France) dont les programmes permettent de contrôler l'ensemble du dispositif et de recueillir les données expérimentales. Chaque cage est disposée dans une enceinte en bois isolante (74 x 46 x 50 cm), munie de quatre diodes pour l'éclairage d'ambiance (10 lux) et d'un ventilateur fournissant un bruit de fond de 55 dB. Ces cages de

conditionnement sont entièrement modulables à l'exception de la porte d'entrée rabattable qui est en plexiglas transparent (figure I.1.1).

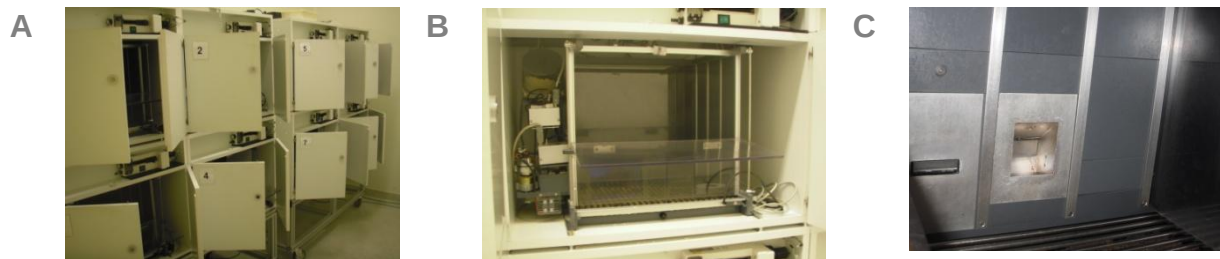


Figure I.1.1 Photographies du dispositif expérimental.

Les 8 cages de conditionnement sont isolées individuellement dans des enceintes en bois (A et B). Entièrement modulables, elles sont configurées pour cette expérience avec une mangeoire et un levier rétractable.

Dans cette expérience, nous utiliserons les cages dans la configuration suivante : Un sol constitué d'une grille de barreaux métalliques (diamètre : 5 mm) est placé au dessus d'un tiroir amovible garni de copeaux de bois. La paroi faisant face à la porte d'entrée est constituée de plexiglas transparent doublé d'une plaque de PVC noir. La paroi latérale droite est entièrement composée de blocs de PVC gris amovibles. Les blocs de la paroi gauche (figure I.1.1, panneau C) sont disposés de façon à présenter à gauche un levier rétractable (4 x 1 x 1,5 cm) et au milieu une mangeoire (6 x 4,5 x 4,5 cm). La mangeoire est reliée à un distributeur de pastilles de nourriture et est équipée d'une cellule à infra-rouge détectant les accès. Pour cette expérience, les pastilles (45 mg) utilisées comme récompenses sont formulées à base de grains et de graisses animales (Rodent Grain-Based Diet, Bio_Serv).

c. Protocole expérimental

Le protocole de cette première expérience comprend 4 phases et se déroule au rythme d'une séance de conditionnement par jour (figure I.1.2).



Figure I.1.2 Représentation schématique du déroulement de l'expérience.

Mag : séance d'entraînement à la mangeoire ; FR1 et VI30 : séances de conditionnement instrumental ; D : séance de dévaluation par AGC ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; FdR : test pendant lequel la réponse est renforcée.

- *Phase 1 : Entraînement à la mangeoire*

En premier lieu, les sujets sont soumis à 2 séances permettant de les habituer à leur cage de conditionnement respective. Ils apprennent alors à aller chercher dans la mangeoire une pastille de nourriture, nouvelle pour eux, qui est distribuée automatiquement. Chaque

séance, d'une durée de 40 minutes, commence par l'allumage de l'éclairage d'ambiance, puis 40 pastilles sont délivrées selon un programme à intervalle de temps variable, à raison d'une pastille délivrée en moyenne toutes les 60 secondes. Aucun levier n'est présenté à ce stade. La séance se termine par l'extinction de l'éclairage. L'animal est alors retiré de la cage de conditionnement et replacé au sein de l'animalerie.

- *Phase 2 : Conditionnement instrumental*

L'entraînement instrumental s'étend sur 6 séances pendant lesquelles les sujets doivent apprendre à associer l'action d'appui sur le levier à la distribution de la pastille de nourriture dans la mangeoire. Une fois l'animal placé dans sa cage de conditionnement, la séance instrumentale commence par l'allumage des lumières et par la présentation du levier. Lorsque 40 appuis sur le levier ont été récompensés ou lorsque 45 minutes se sont écoulées, la séance instrumentale est terminée, le levier est rétracté et les lumières s'éteignent. Le sujet est alors replacé dans sa cage d'habitation au sein de l'animalerie.

Les deux premières séances instrumentales suivent un programme de renforcement en ratio fixe FR1, où une pastille est distribuée dans la mangeoire à chaque appui. Lors des quatre séances suivantes, le programme de renforcement VI30 impose après chaque appui récompensé un intervalle variable d'une durée moyenne de trente secondes pendant lequel le levier n'a plus d'effet.

- *Phase 3 : Dévaluation de la récompense par AGC*

A l'issue de la phase d'entraînement, nous constituons 2 groupes d'animaux équilibrés en ce qui concerne le nombre d'appuis sur le levier et le nombre de visites à la mangeoire: un groupe *Devalued (D)* pour lequel la récompense est dévaluée et un groupe *NonDevalued (ND)* pour lequel elle n'est pas dévaluée. Deux rats occupant une même cage d'habitation sont toujours affectés à un même groupe et reçoivent donc le même traitement.

La phase de dévaluation s'étend sur six séances. Une séance de dévaluation est analogue à une séance d'entraînement à visiter la mangeoire, en absence du levier, avec distribution automatique de 40 pastilles pendant 40 minutes. A la fin de chaque séance, les animaux reçoivent une injection intra-péritonéale de chlorure de lithium ou de chlorure de sodium. Ainsi, les rats du groupe *D* reçoivent une injection de chlorure de lithium (LiCl, 0,15 M, 10 ml/kg, Hydrate Fluka Chemica). Les rats du groupe *ND* reçoivent une injection de chlorure de sodium (NaCl, 0.9 M, 10 mL/kg, Cooper). Le soir, afin d'éviter toute interférence

liée à leur nourriture habituelle, tous les sujets reçoivent leur dose quotidienne d'alimentation dans l'animalerie au moins 2 heures après la fin de la séance de dévaluation.

- *Phase 4 : Tests comportementaux*

- Test RdR

Afin d'évaluer le contrôle de la réponse instrumentale par une représentation mentale de la récompense, les animaux sont soumis en premier lieu à un test RdR (Representation dependent Response). Les appuis sur le levier ne sont jamais récompensés au cours de cette séance. Un test RdR se déroule de la façon suivante : les sujets sont placés dans leur cage de conditionnement respective, les lumières s'allument et le levier est présenté. Les appuis effectués et les visites à la mangeoire sont enregistrés. Au bout de 30 minutes, les lumières s'éteignent et le levier est retiré, l'animal réintègre sa place au sein de l'animalerie.

- Test FdR

Une ultime séance permet de tester l'effet direct de la récompense, présente physiquement, sur la réponse instrumentale (Food dependent Response). Ainsi, lors du test FdR les appuis sur le levier sont de nouveau récompensés. Ce second test est identique à une séance d'entraînement instrumental VI30 (cf. Phase 2).

***d.* Analyses statistiques**

A la fin de chaque séance, nous calculons un taux de visites à la mangeoire et/ou un taux d'appuis sur le levier par unité de temps (minute). Pour chaque individu, 2 taux de référence sont établis en moyennant entre eux les taux d'appuis ou de visites des 2 dernières séances instrumentales. Les performances exprimées subséquentement pourront être rapportées au taux de référence. Des cinétiques d'appuis et de visites peuvent également être analysées. Enfin, un indice de consommation peut être calculé en rapportant le nombre de pastilles consommées au nombre total de pastilles distribuées.

Afin de comparer les performances entre les groupes, tous les résultats sont analysés à l'aide du logiciel StatView® (version 5.0.1). Les différentes variables dépendantes (*taux d'appuis, taux de visites, indice de consommation*) sont comparées par des analyses de variance à facteurs simples ou multiples (ANOVA).

Dans cette expérience les variables indépendantes étudiées sont les suivantes :

- Le facteur **séance** (1-6), correspondant à la répétition des 6 séances de conditionnement (Phases de conditionnement opérant et de dévaluation) ;

- le facteur **temps**, correspondant à l'écoulement du temps (par intervalles de 5 minutes) au sein d'une même séance;
- le facteur **dévaluation** (*Devalued/NonDevalued*), correspondant au traitement (LiCl, NaCl) reçu au cours des séances de dévaluation.

Si besoin, un test post-hoc, le Student-Newman-Keuls (SNK), permet de préciser les différences entre les séances de conditionnement et/ou entre les groupes. Pour tous les tests statistiques, le risque de type 1 est fixé à 5%.

B. Résultats

• *Entraînement à la mangeoire*

Les 16 sujets s'habituent rapidement au contexte des cages de conditionnement. Dès la première séance, ils vont visiter la mangeoire et consomment une partie des pastilles de nourriture. A la fin de la deuxième séance, ils consomment en moyenne $39,5 \pm 0,5$ des 40 pastilles distribuées et ils présentent un taux moyen de $7,3 \pm 0,8$ visites par minute (données non représentées).

• *Conditionnement instrumental*

On peut suivre, sur la figure I.1.3, l'évolution des performances comportementales des sujets au cours du conditionnement instrumental.

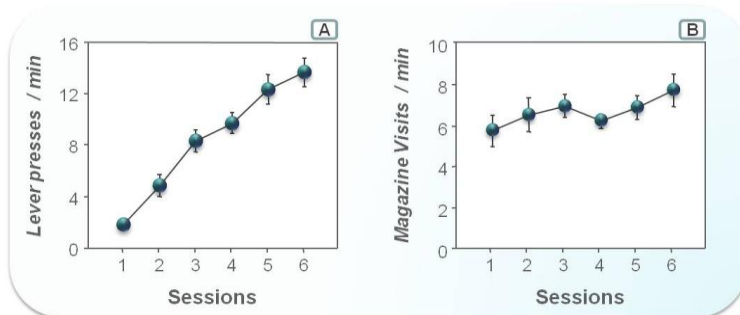


Figure I.1.3 Comportements d'appui (A) et de visite (B) ($\pm$ erreur standard) au cours du conditionnement opérant. Les sujets acquièrent facilement la tâche. Au fil des séances, les taux d'appuis augmentent progressivement. Les taux de visites augmentent légèrement lors de cette phase.

Le panneau A consacré au taux d'appuis moyen, nous montre que les animaux apprennent sans difficulté à appuyer sur le levier. En effet, le niveau de performance augmente progressivement et continue à progresser y compris lors des dernières séances d'entraînement. En comparaison, le taux de visites à la mangeoire évolue faiblement au cours des séances d'apprentissage instrumental (panneau B). À partir de la 2ème séance, le critère maximum

des 40 récompenses distribuées est atteint invariablement par chaque individu, elles sont consommées dans leur totalité.

- * Une ANOVA simple confirme l'effet du facteur **séance** sur le taux d'appuis ($F_{5,15} = 36,708$; $p < 0,001$). Il n'y a pas d'effet du facteur **séance** sur le taux de visites ($F_{5,15} = 1,304$; ns).

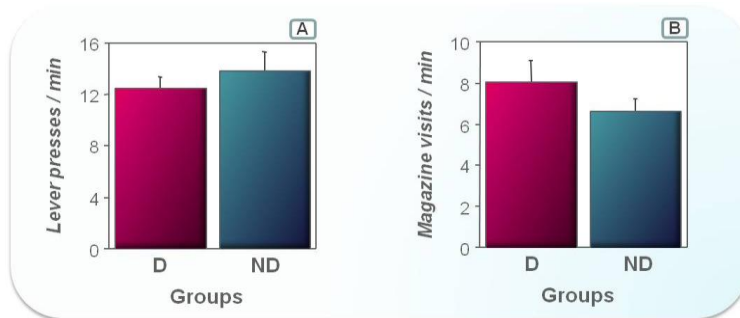


Figure I.1.4 Taux de référence d'appuis (A) et de visites (B) ($\pm$ erreur standard).

Les performances des deux dernières séances du conditionnement opérant sont moyennées et servent de base pour analyser les comportements observés lors de la phase de test.

A l'issue du conditionnement opérant, Les 2 groupes *NonDevalued* ($n=8$) et *Devalued* ($n=8$) sont composés sur la base des performances. Leur taux de référence d'appuis sur le levier et de visites à la mangeoire se situent respectivement aux environs de 13 appuis/min et de 7 visites/min (figure I.1.4).

- * L'analyse statistique confirme l'absence de différence significative entre ces deux groupes aussi bien dans leurs performances instrumentales ($F_{1,14} = 0,632$; ns) que dans leurs comportements de visites ($F_{1,14} = 1,334$; ns). Ces deux groupes présentent donc des performances instrumentales comparables juste avant la phase de dévaluation.

• Dévaluation de la récompense par AGC

Les résultats de cette troisième phase sont représentés sur la figure I.1.5.

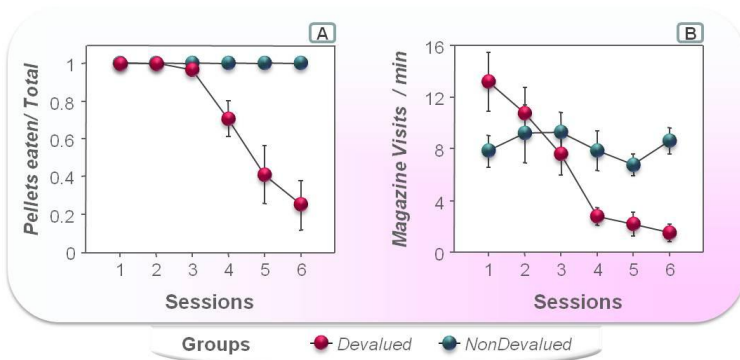


Figure I.1.5 Dévaluation de la récompense.

Evolution (A) des indices de consommation (B) des taux de visites ($\pm$ erreur standard) au cours de la phase de dévaluation. Les animaux recevant du LiCl diminuent leur consommation de la récompense à partir de la 3^{ème} séance. Leur comportement de visite présente une diminution dès la seconde séance de dévaluation.

On peut observer sur le panneau A, l'évolution de l'indice moyen de consommation au cours des séances de dévaluation. Alors que les sujets du groupe *NonDevalued* consomment invariablement l'intégralité des pastilles distribuées, la quantité de nourriture consommée par les rats qui reçoivent du LiCl, diminue progressivement à partir de la troisième séance. A la

fin de la phase de dévaluation, les sujets du groupe *Devalued* ne consomment plus que 25% en moyenne de la totalité des récompenses distribuées.

- * Une ANOVA à 2 facteurs (**dévaluation**, **séance**) révèle des effets significatifs du facteur **dévaluation** ($F_{1,14} = 19,827$; $p < 0,001$), du facteur **séance** ($F_{5,70} = 19,146$; $p < 0,001$) ainsi qu'une interaction entre ces deux facteurs ($F_{5,70} = 19,146$; $p < 0,001$).
- * Le test SNK confirme une différence significative entre la quantité de récompense consommée par le groupe *Devalued* et celle du groupe *NonDevalued* lors de la dernière séance de dévaluation.

Le panneau B représente l'évolution des taux de visites à la mangeoire au cours de la phase de dévaluation. Lors de la première séance de dévaluation en absence du levier, ces deux groupes ne présentent pas les mêmes taux de visites malgré l'établissement de groupes équilibrés en taux de visites pendant le conditionnement opérant. Le taux de visites est plus élevé chez le groupe *Devalued*, mais cet écart en performance tend à s'inverser au fil des séances. En effet, le comportement d'approche du groupe *NonDevalued* reste relativement stable tout au long des séances. Le taux final est de $8,6 \pm 1$ visites/min pour les animaux recevant du NaCl. Au contraire, les rats du groupe *Devalued* qui présentaient un taux de visites nettement supérieur lors de la première séance, visitent de moins en moins la mangeoire au fil des séances de dévaluation. Ils atteignent un taux final de $1,5 \pm 0,7$ visites/min.

- * Une ANOVA révèle une absence d'effet du facteur **dévaluation** ($F_{1,14} = 1,314$; ns), mais confirme un effet significatif du facteur **séance** ($F_{5,70} = 15,211$; $p < 0,001$), ainsi que l'existence d'une interaction **dévaluation x séance** ($F_{5,70} = 11,703$; $p < 0,001$).
- * Le test SNK confirme une différence significative dans le comportement d'approche de la mangeoire entre le groupe *Devalued* et le groupe *NonDevalued* à la fin de cette phase de dévaluation.

- *Tests comportementaux*

- **Test RdR**

Les performances comportementales exprimées par les 2 groupes au cours du test RdR sont représentées sur la figure I.1.6

On peut observer sur le panneau A l'évolution des taux d'appuis au cours des 30 minutes de test. On constate chez les animaux du groupe *NonDevalued* un niveau moyen d'appuis important durant les cinq premières minutes du test qui diminue progressivement avec le déroulement de la séance. Le patron de résultats est semblable pour le groupe *Devalued*, mais

ces animaux présentent des taux d'appuis nettement inférieurs à ceux du groupe *NonDevalued* tout au long de la séance.

- * L'analyse statistique confirme cette description des données. Il y a un effet significatif du facteur **dévaluation** ($F_{1,14} = 16,586$; $p < 0,001$) et un effet significatif du facteur **temps** ($F_{5,70} = 25,992$; $p < 0,001$), mais elles indiquent une absence d'interaction **dévaluation x temps** ($F_{5,70} = 1,929$; ns).

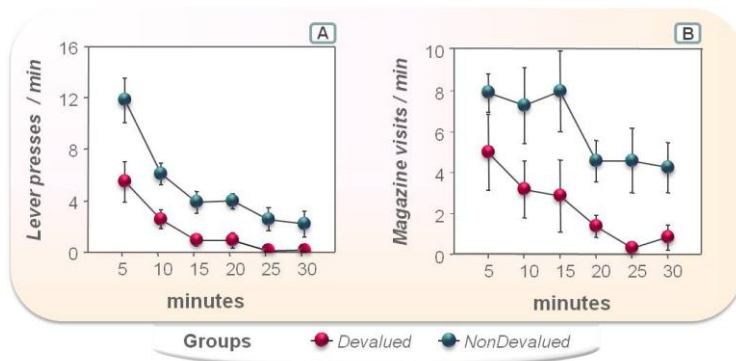


Figure 1.1.6 Test RdR.

Evolution temporelle (A) des taux d'appuis et (B) des taux de visites ($\pm$ erreur standard) au cours des 30 minutes.

Les performances comportementales des 2 groupes diminuent de façon progressive. Les animaux ayant reçu du LiCl expriment des niveaux de performances plus faibles.

L'évolution des taux de visites est représentée sur le panneau B. Comme pour le comportement d'appui, le groupe *Devalued* exprime un niveau de visites réduit en comparaison du groupe *NonDevalued*.

- * Une ANOVA révèle un effet significatif du facteur **dévaluation** ($F_{1,14} = 8,914$; $p < 0,01$) et un effet significatif du facteur **temps** ($F_{5,70} = 4,694$; $p < 0,001$), mais ne montre pas d'interaction significative **dévaluation x temps** ($F_{5,70} = 0,255$; ns).

Ainsi, l'analyse statistique n'indique pas d'effet du traitement par LiCL sur la vitesse d'extinction de la réponse. L'effet de la dévaluation sur la réponse instrumentale devra essentiellement être étudié avant que cette réponse soit éteinte. Nous nous proposons donc d'étudier les résultats comportementaux des 15 premières minutes de test, représentés sur les figures I.1.7 et I.1.8.

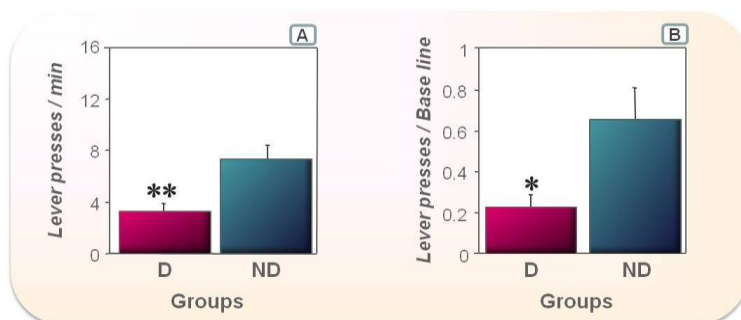


Figure 1.1.7 Performances instrumentales exprimées lors des 15 premières minutes du test RdR.

Les résultats sont exprimés en appuis par minute ($\pm$ erreur standard) (A) ou rapportés aux taux de référence (B). Les sujets du groupe *Devalued* (D) appuient peu sur le levier par rapport au groupe *NonDevalued* (ND). * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$

Les taux moyen d'appuis sur le levier exprimés par chaque groupe sont représentés sur le panneau A de la figure I.1.7. Nous observons un faible taux d'appuis chez le groupe *Devalued*

($3,1 \pm 0,83$ appuis/min) en comparaison au groupe *NonDevalued* ($7,35 \pm 1,1$ appuis/min). La description des résultats est analogue si on rapporte les performances émises lors du test au taux de référence obtenus à la fin du conditionnement opérant (panneau B). Avant la phase de dévaluation, ces taux de référence ne différaient pas entre les 2 groupes. Après la dévaluation, les animaux du groupe *Devalued* appuient à un niveau correspondant à 22% de leur performance de référence alors que les animaux du groupes *NonDevalued* sont à 65%.

* Une ANOVA indique d'un effet de la **dévaluation** sur la performance instrumentale brute ($F_{1,14} = 9,597$; $p < 0,01$) ou rapportée aux taux de référence ($F_{1,14} = 6,560$; $p < 0,05$) .

Les panneaux A et B de la figure I.1.8 représentent respectivement les taux de visites exprimés en valeurs brutes et le taux de visites rapporté aux taux de référence.

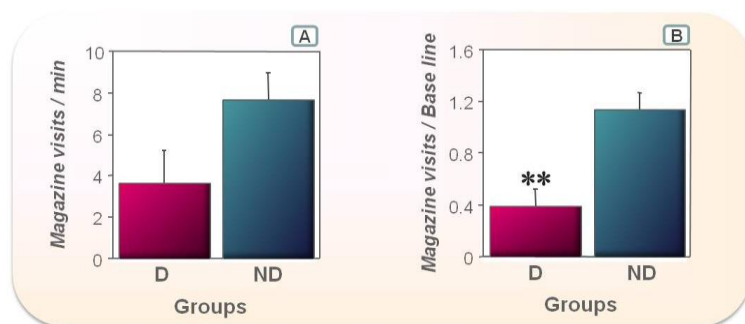


Figure I.1.8 Comportement de visites au cours des 15 premières minutes du test RdR

Les résultats peuvent être exprimés en visites par minute ($\pm$ erreur standard). (A) ou rapportés aux taux de référence (B). Les sujets du groupe *Devalued* (D) visitent peu sur la mangeoire par rapport au groupe *NonDevalued* (ND).
* $p < 0.05$; ** $p < 0.01$; *** $p < 0,001$

Le patron des résultats obtenus pour le comportement de visite est similaire à celui du comportement opérant. Les sujets du groupe *NonDevalued* font $7,9 \pm 1,33$ visites/min à la mangeoire, ce qui correspond à 113% de leur comportement de référence. Au contraire, les sujets du groupe *Devalued* visitent peu la mangeoire, avec un taux de $3,7 \pm 1,6$ visites/min, soit 40% de leur comportement de référence.

* Les analyses de variance révèlent une tendance à un effet de la **dévaluation** pour le comportement d'approche brut ($F_{1,14} = 3,770$; $p = 0,073$) et un effet significatif de la **dévaluation** sur ce comportement quand il est rapporté aux taux de référence ($F_{1,14} = 14,163$; $p < 0,01$) .

- Test FdR

Les résultats obtenus pour ce second test sont représentés sur les figures I.1.9, I.1.10 et I.1.11.

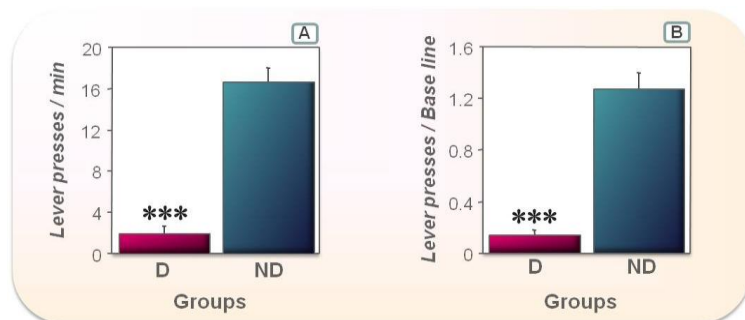


Figure I.1.9 Taux d'appuis lors du test FdR, en appuis/min (A) ou rapportés aux taux de référence (B) ($\pm$ erreur standard).

Quand la réponse est de nouveau récompensée, les sujets du groupe *NonDevalued* (ND) augmentent leurs taux d'appuis, ceux du groupe *Devalued* (D) ont une performance instrumentale très faible.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0,001$

Les panneaux A et B (figure I.1.9) représentent respectivement les taux d'appuis exprimés en valeur brute et les taux d'appuis rapportés aux taux de référence. Les sujets formant le groupe *NonDevalued* appuient activement sur le levier avec un taux moyen de $16,7 \pm 1,4$ appuis/min, ce qui représente 126 % de leur performance instrumentale de base. Au contraire, les sujets du groupe *Devalued* présentent une performance instrumentale très faible de $1,9 \pm 0,73$ appuis/min, soit 13% de leur performance de référence.

- * L'analyse statistique des résultats confirme un effet significatif du facteur **dévaluation** sur les performances instrumentales exprimées en valeur brute ($F_{1,14} = 90,625$; $p < 0,001$) et rapportées à leur taux de référence ($F_{1,14} = 69,841$; $p < 0,001$).

La figure I.1.10 représente les résultats obtenus pour le comportement de visite, en valeur brute (panneau A) et en valeur rapportées aux taux de référence (panneau B).

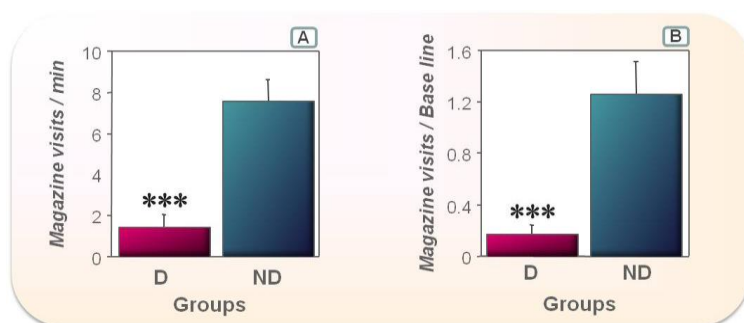


Figure I.1.10 Comportement de visite au cours du test FdR, en visites/min (A) et rapportés aux taux de référence (B) ($\pm$ erreur standard).

Les résultats du groupe *Devalued* (D) sont conformes à ceux attendus, les sujets évitent la mangeoire. (ND, *NonDevalued*).

* $p < 0.05$; ** $p < 0.01$; *** $p < 0,001$

Une fois de plus, les sujets du groupe *NonDevalued* sont plus actifs, ils visitent en moyenne $7,6 \pm 1,02$ fois la mangeoire par minute, exprimant ainsi 126% de leur comportement de référence. Les rats du groupe *Devalued*, n'expriment en revanche que 21% de leur comportement de référence, ils visitent en effet la mangeoire $1,5 \pm 0,6$ fois par minute.

- * Les analyses de variance attestent une fois encore de l'effet significatif de la **dévaluation** sur le comportement d'approche exprimé en valeur brute ($F_{1,14} = 26,425$; $p < 0,001$) ou rapporté aux taux de référence ($F_{1,14} = 16,795$; $p < 0,01$).

Enfin, la figure I.1.11 montre le nombre de récompenses obtenues lors du test FdR et les indices moyens de consommation pour les 2 groupes.

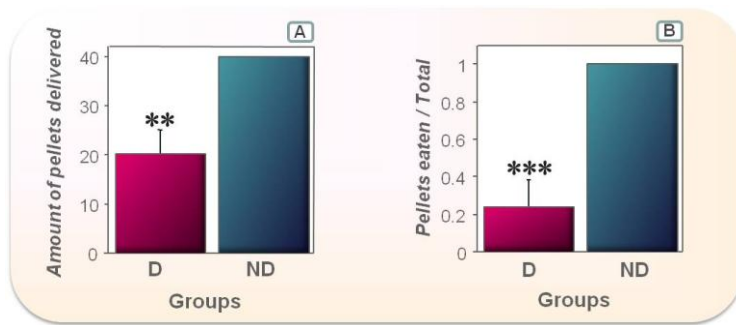


Figure I.1.11 Test FdR.

Nombre de récompenses distribuées (A) et indices de consommation (B) ($\pm$ erreur standard). Les rats du groupe *Devalued* (D) obtiennent moitié moins de récompenses que ceux du groupe *NonDevalued* (ND). Leur indice de consommation atteste de leur aversion gustative pour la pastille de grain. * $p < 0.05$; ** $p < 0.01$; *** $p < 0,001$

Au cours du test FdR, tous les sujets du groupe *NonDevalued* atteignent le critère maximal des 40 récompenses qu'ils consomment dans leur totalité. Au contraire, les animaux du groupe *Devalued* se distribuent en moyenne $20,25 \pm 4,8$ pastilles et n'en consomment que 24%.

* L'ANOVA indique de l'effet significatif du facteur **dévaluation** sur la quantité de récompense obtenues ($F_{1,14} = 16,966$; $p = 0,001$) et sur la quantité consommée ($F_{1,14} = 27,877$; $p < 0,001$).

Conclusion

Après une courte phase d'habituation à la nouvelle pastille de nourriture et au contexte des cages de conditionnement, les sujets apprennent sans difficulté la tâche instrumentale proposée, puis ils sont soumis à la procédure de dévaluation. Elle se caractérise par une acquisition lente, mais permet ainsi un contrôle plus précis des niveaux d'aversion acquis par chaque individu.

Le niveau final d'AGC est suffisant pour produire un effet sur la réponse d'appui. En effet, nous observons chez les animaux traités avec du LiCL une réponse d'appui diminuée donc sensible à la dévaluation de la récompense. Cette sensibilité est observée aussi bien dans les conditions où l'appui n'est pas récompensé (test RdR), que dans les conditions contraires (test FdR).

La procédure de conditionnement appliquée aboutit bien à une réponse de type dirigée vers un but sensible au changement de valeur.

Expérience I.2

Extinction de l'AGC et Action dirigée

Cette expérience concerne les mêmes sujets que l'Expérience I.1, et lui fait immédiatement suite. Elle vise à déterminer la force du conditionnement aversif, par l'étude de son extinction et de ses conséquences sur la réponse instrumentale.

A. Matériel et méthode

«. Protocole expérimental

Ce protocole comporte 3 phases schématisées sur la figure I.2.1.



Figure I.2.1 Déroulement de l'expérience.
E : séance d'extinction de l'AGC ; VI30 : séance de réacquisition, RdR : test pendant lequel la réponse dépend d'une représentation de la récompense.

- *Phase 5 : Extinction de l'aversion gustative*

Les séances d'extinction de l'aversion pour la récompense se déroulent de façon identique aux séances d'entraînement à visiter la mangeoire (cf. Expérience I.1, Phase 1). Elles concernent aussi bien les rats du groupe *Devalued* que ceux du groupe *NonDevalued*. Une fois 40 minutes écoulées et 40 pastilles distribuées automatiquement, le sujet est replacé directement dans sa cage d'habitation au sein de l'animalerie. Les séances prennent fin lorsque tous les sujets consomment la totalité des pastilles qui leurs sont distribuées.

- *Phase 6 : réacquisition*

L'ensemble des sujets est soumis à une séance instrumentale récompensée selon un programme de renforcement en VI30 (cf. Expérience I.1, Phase 2).

- *Phase 7 : Test RdR*

Enfin, les animaux sont soumis à un ultime test RdR dont le déroulement est identique à celui décrit lors de l'Expérience I.1 (cf. Phase 4).

b. Analyses statistiques

Les variables dépendantes et indépendantes étudiées correspondent à celles décrites précédemment en Expérience I.1.

B. Résultats

- *Extinction de l'aversion gustative*

Il a fallu six séances d'exposition à la nourriture dévaluée avant d'obtenir une extinction totale de l'AGC pour cette récompense chez l'ensemble des animaux. On peut suivre sur la figure I.2.2, l'évolution des indices moyens de consommation.

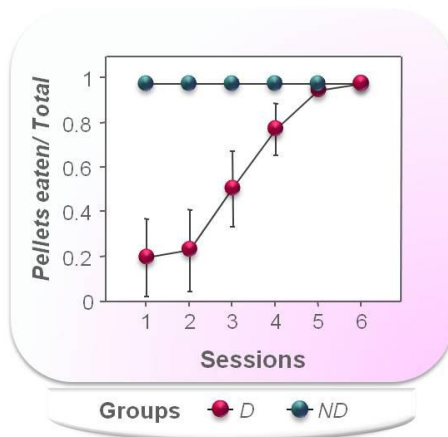


Figure I.2.2 Evolution de la consommation de la récompense au cours des séances d'extinction de l'AGC ($\pm$ erreur standard).

En absence d'appariement avec du LiCl, les sujets du groupe *Devalued* (D) consomment de plus en plus de pastilles. Leur consommation devient maximale et rejoint celle du groupe *NonDevalued* (ND) au bout de 6 séances.

Le groupe *Devalued* présente à la première séance un indice moyen de $0,2 \pm 0,4$, puis l'indice augmente progressivement. Les animaux consomment plus de 50% des récompenses distribuées au bout de la 3ème séance, et finissent par en consommer l'intégralité à la 6ème séance. Comme attendu, l'indice des animaux du groupe *NonDevalued* reste invariablement égal à 1.

* Une ANOVA à 2 facteurs confirme les effets significatifs des facteurs **dévaluation** ($F_{1,14} = 15,876$; $p < 0,01$) et **séances** ($F_{5,70} = 10,893$; $p < 0,001$) ainsi que l'existence d'une interaction **dévaluation x séance** ($F_{5,70} = 10,893$; $p < 0,001$).

- *Séance de réacquisition*

On peut observer sur la figure I.2.3 que lors de cette séance, tous les sujets se comportent de façon similaire, indépendamment du traitement reçu en Expérience I.1.

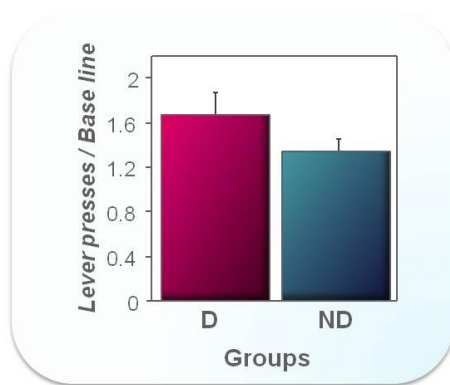


Figure I.2.3 Comportement d'appui rapporté aux taux de référence lors de la séance de réacquisition ($\pm$ erreur standard).

Une fois l'AGC éteinte les animaux du groupe *Devalued* (D) appuient activement sur le levier à des niveaux comparables au groupe *NonDevalued* (ND).

En effet, tout comme les animaux du groupe *NonDevalued*, les sujets du groupe *Devalued* appuient activement sur le levier et se distribuent ainsi 40 récompenses qu'ils consomment dans leur totalité.

* Une ANOVA simple confirme l'absence d'effet du facteur **dévaluation** sur la performance instrumentale ($F_{1,14} = 1,421$; ns).

- *Test RdR*

La figure I.2.4 représente les taux d'appuis sur le levier exprimés lors du test RdR.

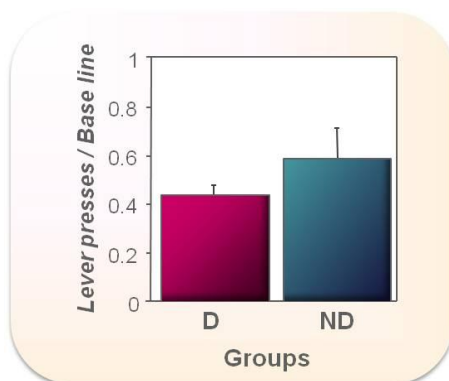


Figure I.2.4 Comportement d'appui rapporté aux taux de référence lors du test RdR. ($\pm$ erreur standard).

Les animaux du groupe *Devalued* (D) appuient activement sur le levier à des niveaux comparables au groupe *NonDevalued* (ND).

Les animaux des groupes *Devalued* et *NonDevalued* expriment des performances instrumentales à des niveaux comparables, respectivement de $5,4 \pm 0,7$ appuis/min et de $6,5 \pm 0,7$ appuis/min.

* L'analyse statistique des résultats indique de l'absence d'effet significatif du facteur **dévaluation** sur la performance instrumentale ($F_{1,14} = 1,169$; ns).

Conclusion

La procédure de dévaluation de la récompense décrite en Expérience I.1 aboutit à une AGC relativement stable. Elle est néanmoins réversible quand on présente à l'animal de façon répétée la récompense alimentaire non appariée à un malaise digestif. Le comportement opérant des animaux n'est pas affecté par l'expérience révolue d'une aversion gustative pour la récompense.

Nous pouvons conclure des résultats de cette expérience que l'effet de la dévaluation sur la performance instrumentale décrite en Expérience I.1 est réversible et dépend d'une valeur actualisée de la récompense.

Expérience I.3

Procédure Habitude, 1^{er} essai

Afin d'étudier l'effet d'un entraînement plus soutenu sur la nature de la récompense instrumentale, nous avons entraîné les animaux avec un nombre de séances de conditionnement doublé. Le programme de renforcement VI30 reste en revanche inchangé.

A. Matériel et méthode

a. Sujets

Dix-huit rats mâles naïfs de la souche Long-Evans, de 350 à 390 g, sont utilisés pour cette expérience. Les conditions d'hébergement et le dispositif expérimental sont identiques à ceux décrits dans l'Expérience I.1. La quantité de nourriture fournie le soir est réduite à 13 g par rat, en raison des deux séances d'entraînement quotidiennes.

b. Protocole expérimental

Comme illustré en figure I.3.1, le protocole de cette expérience est en tout point identique à celui de l'Expérience I.1, à l'exception de la Phase 2.



Figure I.3.1 Schématisation du déroulement de l'expérience.

Mag : séance d'entraînement à la mangeoire ; FR₁ et VI₃₀ : séances de conditionnement instrumental ; D : séance de dévaluation par AGC ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; FdR : test pendant lequel la réponse est renforcée.

L'entraînement à la mangeoire (Phase 1) et la phase de dévaluation de la récompense se déroulent (Phase 3) comme décrites en Expérience I.1.

- *Phase 2 : Conditionnement instrumental*

L'entraînement instrumental commence comme en Expérience I.1 puis est prolongé par 5 séances en VI₃₀. Il se déroule à un rythme de 2 séances par jour.

- *Phase 4 : Tests comportementaux*

L'expérience s'achève par une Test RdR d'une durée de 15 minutes suivi le lendemain par un test FdR (cf. Expérience I.1).

c. Analyses statistiques

Les variables dépendantes et indépendantes étudiées correspondent à celles décrites précédemment en Expérience I.1.

B. Résultats

- *Entraînement à la mangeoire*

L'ensemble des sujets s'acclimatent facilement au dispositif et apprennent à aller chercher la récompense dans la mangeoire (données non représentées).

- *Conditionnement instrumental*

La figure I.3.2 illustre l'évolution des appuis au cours du conditionnement instrumental.

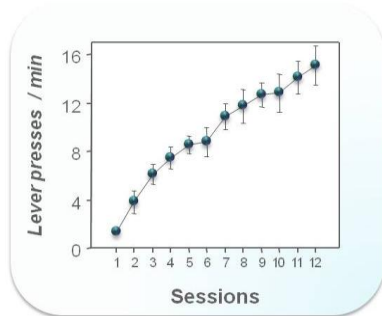


Figure I.3.2 Evolution des taux d'appuis au fil des séances de conditionnement opérant ($\pm$ erreur standard).

La performance instrumentale des sujets est en constante progression tout au long de l'entraînement.

Au fil des séances d'entraînement, la performance instrumentale évolue selon une courbe à pente ascendante. En effet, les taux d'appuis augmentent séance après séance sans qu'aucun plateau ne soit constaté. Concernant le comportement de visite, les taux sont stables jusqu'à la 8^{ème} séance ; en effet les sujets semblent plus actifs à la mangeoire lors des quatre dernières séances (données non représentées). A la fin du conditionnement, deux groupes sont constitués de façon équilibrés en vue de la phase de dévaluation. Leurs taux de référence d'appuis se situent aux environs de 15 appuis/min, ils sont donc proches des valeurs obtenues en Expérience I.1.

* Une ANOVA confirme un effet significatif du facteur **séance** sur la performance instrumentale ($F_{11,187} = 35,149$; $p < 0,001$).

* Il n'apparaît pas de différence entre les taux de référence d'appuis ($F_{1,16} = 0,013$; ns) des deux groupes ainsi constitués.

- *Dévaluation de la récompense par AGC*

Les données concernant le comportement de consommation au cours de la phase de dévaluation sont représentées sur la figure I.3.3.

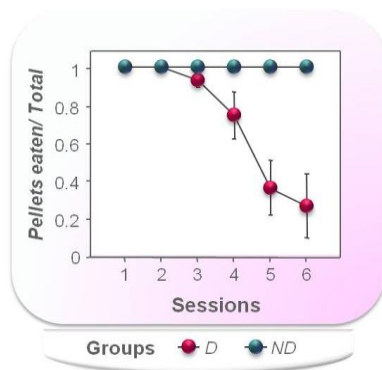


Figure I.3.3 Evolution des indices de consommation au fil des séances de dévaluation ($\pm$ erreur standard). Les sujets acquièrent l'AGC à une vitesse et un niveau comparables à l'expérience précédente.

Le comportement des sujets est globalement comparable à celui observé en Expérience I.1. Les rats du groupe *NonDevalued* consomment invariablement l'intégralité des récompenses distribuées et visitent la mangeoire à des niveaux relativement stables. Au contraire, comme en Expérience I.1, l'indice de consommation du groupe *Devalued* diminue à partir de la 3^{ème} séance. La préexposition à de plus grande quantité de récompense n'a donc pas d'effet sur la vitesse de l'acquisition de l'aversion. Lors de la dernière séance les animaux ayant reçu du LiCl ne consomment en moyenne que 34% des pastilles distribuées.

* L'analyse statistique indique des effets significatifs des facteurs **dévaluation** ($F_{1,16} = 18,659$; $p < 0,001$) et **séance** ($F_{5,80} = 13,257$; $p < 0,001$) et également de l'existence d'une interaction **dévaluation x séance** ($F_{5,80} = 13,257$; $p < 0,001$).

- *Tests comportementaux*

- **Test RdR**

Les résultats du test RdR sont représentés sur la figure I.3.4. Les performances de chaque animal sont rapportées à son taux de référence.

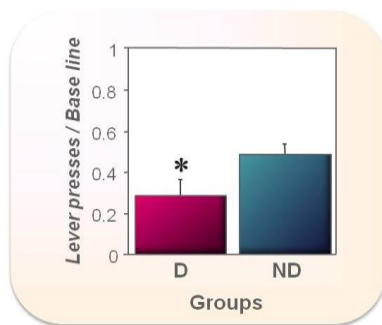


Figure I.3.4 Taux d'appuis au cours du test RdR. ($\pm$ erreur standard).
Le groupe *Devalued* (D) présente un taux d'appuis moyen plus faible que le groupe *NonDevalued* (ND). Néanmoins, l'effet de la dévaluation semble atténué en comparaison de l'effet observé lors de l'expérience précédente.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0,001$

On constate un patron de résultats similaire à celui de l'Expérience I.1. Le groupe *Devalued* exprime des taux d'appuis plus faibles que ceux du groupe *NonDevalued*, leur réponse est donc dirigée par une représentation de la récompense. Néanmoins, il semble que les écarts entre les deux groupes soient amoindris en comparaisons à ceux précédemment observés en Expérience I.1.

* Une ANOVA confirme l'existence d'une différence significative entre les comportements d'appui des deux groupes ($F_{1,16} = 4,707$; $p < 0,05$).

- Test FdR

Les résultats concernant les taux d'appuis et de visites exprimés par les sujets lors du test FdR sont représentés sur la figure I.3.5.

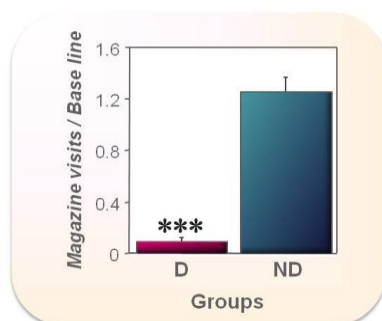


Figure I.3.5 Taux d'appuis au cours du test FdR. ($\pm$ erreur standard).
Lorsque l'appui est de nouveau récompensé, le groupe *Devalued* (D) présente des taux d'appuis proches de zéro. Au contraire les performances du groupe *NonDevalued* (ND) sont élevées.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0,001$

Lorsque l'appui est de nouveau récompensé, les animaux du groupe *Devalued* appuient très peu sur le levier et ne visitent la mangeoire que très rarement. Alors, les écarts entre les performances se creusent et le patron obtenu réplique celui du test FdR de l'Expérience I.1.

* Une ANOVA confirme la différence significative entre les taux d'appuis ($F_{1,16} = 79,161$; $p < 0,001$) et de visites ($F_{1,16} = 12,882$; $p < 0,01$) des deux groupes.

Concernant les niveaux de consommation de la récompense, les sujets du groupe *NonDevalued* consomment la totalité des 40 pastilles. Au contraire, les rats du groupe

Devalued consomment seulement 36% des $17,2 \pm 4,4$ pastilles de grains obtenues (données non représentés).

- * L'analyse statistique confirme que les deux groupes diffèrent significativement quant au nombre de récompenses obtenues ($F_{1,16} = 26,615$; $p < 0,001$) et la quantité de pastilles consommées ($F_{1,16} = 18,322$; $p < 0,001$).

Conclusion

Cette expérience reproduit dans son ensemble les résultats précédemment obtenus avec un conditionnement opérant plus court.

Malgré les 12 séances instrumentales, la performance instrumentale exprimée par les sujets n'atteint pas de plateau. De plus, on note que l'exposition prolongée des sujets à la récompense alimentaire a peu d'effet sur la vitesse d'acquisition de l'aversion gustative. Le niveau d'aversion final exprimé semble moins marqué mais il reste proche de celui précédemment observé et il se trouve être efficace lors de la phase de test.

*Enfin, la diminution de la performance instrumentale chez les animaux *Devalued* indique que la réponse instrumentale reste contrôlée par le système associatif mettant en jeu une représentation actualisée de la récompense.*

Le protocole appliqué au cours de cette expérience (2 FR1 + 10 VI30) n'aboutit pas à l'expression d'une réponse de type habitude. En revanche, les données obtenues attestent de la reproductibilité des résultats précédemment observés au cours de l'Expérience I.1, notamment ceux concernant le comportement d'appui.

Expérience I.4

Procédure Habitude, 2^{ème} essai

Afin de poursuivre notre étude sur les facteurs favorisant l'expression d'une réponse instrumentale de type habitude, nous avons soumis nos sujets à deux conditionnements instrumentaux en parallèle, de composition et de durée différentes. L'un d'entre eux était un conditionnement long imposant une corrélation réponse-récompense plus faible (VI60). D'autre part, la procédure de dévaluation mise en œuvre est la satiété sélective.

A. Matériel et méthode

a. Sujets

Cette expérience a été menée sur 24 rats mâles naïfs de la souche Long-Evans (390 à 490 g). Les conditions d'hébergement sont identiques à celles décrites dans l'Expérience I.1. Concernant la restriction alimentaire, la quantité de nourriture fournie est ajustée aux nombre des séances de conditionnement quotidiennes, soit 15 g par rat pour 1 séance et 13 g pour 2 séances.

b. Dispositif expérimental

Le protocole de cette expérience nécessite l'utilisation de 3 lots de cages, chacun situé dans une salle différente (figure I.4.1):

- Dans la première salle se trouve un lot de 8 cages de conditionnement dans la configuration précédemment décrite en Expérience I.1 (panneau A). Des pastilles de grains constituent la récompense distribuée.
- La deuxième pièce comporte un lot de 8 cages de conditionnement ayant un agencement des parois latérales différent : la paroi droite comporte en son milieu un abreuvoir (7,9 x 5,6 x 9,7 cm) relié à un distributeur de liquide et à son extrémité droite le levier rétractable ; la paroi gauche est entièrement constituée de blocs de PVC (panneau B). Le distributeur délivre des doses de 0.1 mL d'une solution sucrée (saccharose 10%, Sigma).

- Enfin dans la troisième pièce, adjacente aux deux autres, se trouvent 8 cages que nous nommerons cages de consommation. Faites de polycarbonate transparent, ces cages sont comparables aux cages d'habitation de l'animalerie. Cependant, elles sont plus grandes (42 x 28 x 20 cm) et leurs parois sont opacifiées par l'application d'un motif de bandes horizontales noires et blanches. Le sol y est couvert de copeaux de bois au lieu de sciure. Afin de mettre à disposition les récompenses alimentaires, une coupelle de verre (diamètre : 6 cm) peut être fixée au sol et un biberon de solution sucrée peut être posé sur la grille fermant la cage (panneau C).



Figure I.4.1 Photographies du dispositif expérimental.

Il comporte deux lots de 8 cages de conditionnement et 8 cages classiques. Les cages du premier lot distribuent la pastille de grains, le levier et la mangeoire sont situés au sein de la paroi gauche (A). Le second lot comprend des cages avec un levier et un abreuvoir situés sur la paroi latérale droite. Une solution sucrée y est distribuée (B). Les séances de satiété et les tests de consommation ont lieu dans des cages rendues opaques par l'application d'un motif sur leurs parois (C).

c. Protocole expérimental

Le protocole de cette expérience est représenté sur la figure I.4.2. Il comporte 3 phases. L'ensemble des sujets apprend une action d'appui récompensée par une pastille et une action récompensée par la solution sucrée.



Figure I.4.2 Schématisation du déroulement de l'expérience.

Le conditionnement *High* impliquant la pastille de grain est représenté en première ligne. Le conditionnement *Low* impliquant la solution sucrée est représenté en seconde ligne.

Mag : séance d'entraînement à la mangeoire ou à l'abreuvoir ; FR1, VI30 et VI60 : séances de conditionnement instrumental ; S : séance de dévaluation par satiété ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; Tc : test de consommation.

- *Phase 1 : Entraînement à la mangeoire puis à l'abreuvoir*

L'entraînement à la mangeoire distribuant la pastille de grains commence le lendemain de la restriction alimentaire (cf. Expérience I.1). Le conditionnement pour la solution sucrée étant de plus courte durée, les séances d'habituation à l'abreuvoir ont lieu de façon décalée le 9^{ème} jour de l'expérience (cf. figure I.4.2). Les sujets sont alors entraînés le matin dans les cages associées à la pastille de grains, puis l'après-midi dans les cages associées à la solution sucrée (avec un intervalle de temps minimum de 3 heures entre les 2 séances).

- *Phase 2 : Conditionnements instrumentaux*

Le conditionnement instrumental impliquant la solution sucrée se déroule selon le protocole établi dans l'Expérience I.1 : 2 séances en FR1 suivies de 4 séances en VI30. Il est d'une durée courte, nous le nommerons *Low*. L'entraînement à l'action récompensée par la pastille est plus intensif, nous le nommerons *High*. Il comporte 2 séances en FR1, 5 séances en VI30 et 8 séances en VI60.

Les 2 derniers jours d'entraînement instrumental, les animaux sont placés, avec leurs cages d'habitation, au sein de la pièce de consommation afin de les habituer à ce nouvel environnement. Au bout de 15 minutes, ils sont déplacés chacun dans une cage de consommation pour 20 minutes d'habituation. Aucune nourriture ou boisson n'est disponible au cours de ces séances.

- *Phase 3 : Dévaluation par satiété sélective et tests comportementaux*

Cette phase s'étend sur 2 jours de dévaluation et de tests, séparés par une journée de réacquisition. Au cours des deux journées de satiété et de tests, l'ordre de dévaluation, des tests de consommation et de passage dans les deux cages de conditionnement sont contrebalancés. Néanmoins, pour faciliter la compréhension du protocole et des résultats obtenus, nous décrirons uniquement le cas où la première journée est consacrée à la dévaluation de la récompense *solide*, la récompense *liquide* étant dévaluée au cours de la seconde journée de dévaluation.

1^{er} jour

Séance de satiété

Les sujets sont placés dans la pièce de consommation. Une coupelle contenant 30g de pastilles est mise à disposition directement dans la cage d'habitation. Les rats y étant logés par 2, la

quantité de nourriture prise individuellement n'est pas contrôlable. La séance dure 1h afin de laisser le temps aux sujets de consommer la récompense jusqu'à son rejet.

Tests RdR

Après la séance de satiété, les animaux sont soumis consécutivement à deux tests RdR, un test dans chaque cage de conditionnement instrumental (*High* et *Low*). D'une durée de 15 minutes, ils se déroulent comme décrit dans l'Expérience I.1 (Phase 4).

Tests de consommation

Immédiatement après le dernier test RdR, les rats sont placés individuellement dans les cages de consommation pour y effectuer 2 tests successifs. Ils ont d'abord accès à un biberon de 10 mL de solution sucrée. Au bout de 15 minutes, on le retire pour disposer au sol une coupelle contenant 10 g de pastilles, elle est mise à disposition pour une durée égale de 15 minutes.

Enfin, les animaux sont replacés au sein de l'animalerie. Le soir venu, compte-tenu des quantités de récompenses alimentaires auxquelles ils ont eu accès lors de cette journée, aucun supplément de nourriture ne leur est fourni.

2^{ème} jour

Le deuxième jour est consacré à 2 séances de ré-acquisition instrumentale, une séance en VI60 pour la pastille et une séance en VI30 pour la solution sucrée.

3^{ème} jour

Le troisième jour a un déroulement identique au premier. Cependant, lors de la séance de satiété ce sont deux biberons de solution sucrée (2 x 50 mL) auxquels les rats ont accès pendant une heure.

d. Analyses statistiques

Chaque animal sera son propre témoin lors de l'analyse des tests comportementaux. Les variables dépendantes (*taux d'appuis, les taux de visites et les quantités (en g ou mL) de récompenses consommées*) seront analysées à l'aide des variables indépendantes suivantes :

- le facteur **séance** (1- 6) (cf. Expérience I.1) ;

- le facteur **dévaluation** (*NonDevalued/Devalued*), correspondant à la consommation ou non de la récompense lors de la séance de satiété ;
- le facteur **conditionnement** (*High-Low*), correspondant aux deux conditionnements opérants auxquels les sujets ont été soumis.

B. Résultats

• *Entraînement à la mangeoire et à l'abreuvoir*

Les sujets apprennent tous sans difficulté à aller visiter la mangeoire et l'abreuvoir pour y consommer la pastille de grains ou la solution sucrée (données non représentées).

• *Conditionnements instrumentaux*

Les résultats concernant les performances instrumentales exprimées lors des conditionnements *High* et *Low* sont représentés sur la figure I.4.3.

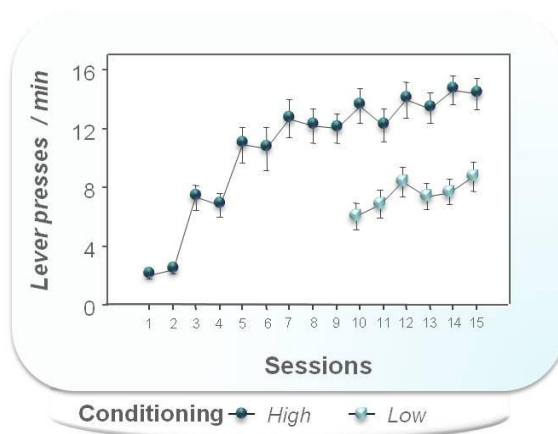


Figure I.4.3 Evolution des taux d'appuis au cours des deux conditionnements ($\pm$ erreur standard).

Les niveaux de performance exprimés au cours du conditionnement *High* atteignent un plateau à partir de la 8^{ème} séance. Le conditionnement *Low* aboutit à une performance instrumentale modérée.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Au cours du conditionnement *High*, le taux d'appuis moyen augmente constamment jusqu'à la 7^{ème} séance, puis les variations observées sont plus faibles. Cette stabilisation aux environs de 15 appuis/min, coïncide avec l'application du programme en VI60 pendant les 8 dernières séances.

Les taux d'appuis effectués lors de la première séance du conditionnement *Low* sont élevés (6 appuis/min) en comparaison à ceux de la toute première séance instrumentale *High* (2 appuis/min). Cependant, ils augmentent faiblement au cours du conditionnement *Low* et atteignent un maximum de 9 appuis/min. Le critère des 40 distributions de la récompense sucrée est néanmoins atteint pour l'ensemble des 6 séances.

* L'analyse de variance indique de l'effet significatif du facteur **séance** sur les performances instrumentales au cours des conditionnements *High* ($F_{14,322} = 43,324$; $p < 0,001$) et *Low* ($F_{5,115} = 3,000$; $p < 0,05$).

- *Dévaluation par satiété sélective, tests comportementaux*

Au cours des deux séances de satiété, visant chacune l'obtention d'une satiété sélective pour l'une des récompenses, un sujet consomme en moyenne 8 g de pastilles de grains ou boit 12 mL de la solution sucrée (données non représentées).

- **Tests de consommation**

Les données des tests de consommation diffèrent en fonction de la récompense, elles sont représentées sur la figure I.4.4.

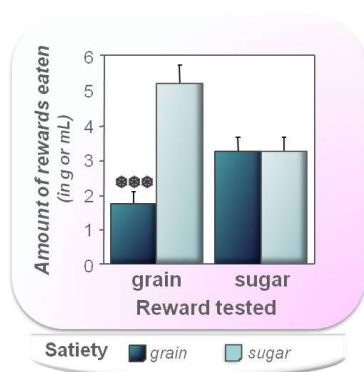


Figure I.4.4 Quantité de récompense ingérée pendant les tests de consommation ($\pm$ erreur standard).

On ne constate pas de différence pour la récompense sucrée. En revanche la séance de satiété est efficace et sélective pour la pastille de grains, la récompense est dévaluée. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Lors de la 1^{ère} journée de dévaluation, les rats ayant absorbé au préalable la pastille en grande quantité, la consomment peu lors du test (1,5g) et absorbent 3mL de la solution sucrée. Lors de la 2^{ème} journée, la séance de satiété met en jeu la solution sucrée, la quantité de pastilles ingérées au cours du test est plus élevée (5g) mais le volume de solution sucrée absorbée reste inchangé. L'absence d'effet de la séance de satiété sur la consommation de la solution ne permet pas de statuer en faveur de l'existence d'une dévaluation pour cette récompense. En revanche, le patron des résultats relatif à la récompense solide témoigne d'une satiété sélective et donc de l'efficacité de la dévaluation.

* L'analyse statistique confirme la description des résultats, les niveaux de consommation diffèrent significativement pour la récompense solide ($F_{1,23} = 57,426$; $p < 0,001$) mais pas pour la récompense liquide ($F_{1,23} < 10^{-30}$; ns).

- **Tests RdR**

Les résultats des tests RdR sont représentés sur la figure I.4.5. Conformément aux résultats des tests de consommation de la solution sucrée, les niveaux de performance sur le levier *Low* ne varient pas. Les résultats concernant l'action apprise au cours du conditionnement *High* sont tout autres. En effet, lorsque les animaux sont dans un état de satiété sélective pour

la pastille, ils appuient à des niveaux très faibles sur le levier. Ils sont donc sensibles à la dévaluation de la récompense, la réponse exprimée est donc de type action dirigée.

- * Une ANOVA révèle les effets significatifs des facteurs **conditionnement** ($F_{1,23} = 19,995$; $p < 0,001$) et **dévaluation** ($F_{1,23} = 38,370$; $p < 0,001$) ainsi qu'une interaction entre ces deux facteurs ($F_{1,23} = 47,660$; $p < 0,001$).
- * Le test SNK localise un comportement significativement différent seulement pour la réponse acquise au cours du conditionnement *High*.

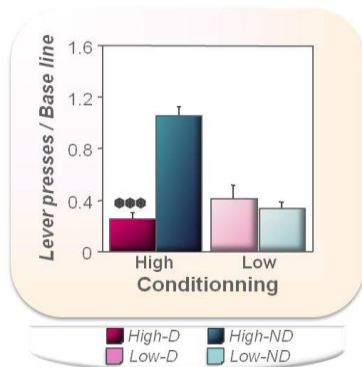


Figure I.4.5 Comportements d'appuis au cours du test RdR ($\pm$ erreur standard).

Les taux d'appuis sur levier *Low* est stable. En revanche, après avoir consommé la pastille de grains, les animaux appuient très peu sur le levier *High*. Leur réponse est dirigée vers un but. * $p < 0.05$; ** $p < 0.01$; *** $p < 0,001$

Conclusion

Les résultats comportementaux relatifs à la solution sucrée soulèvent quelques remarques quant à l'emploi d'une procédure de satiété sélective. En effet, si l'on se basait sur les seuls résultats du test RdR, on en déduirait naturellement que la réponse instrumentale exprimée est de type habitude. Néanmoins, les tests de consommation indiquent que la motivation des individus pour la récompense sucrée est stable, ceci même s'ils ont eu l'occasion d'en consommer à volonté au préalable. En complément, on remarque que les niveaux de performance instrumentale obtenus lors de cette expérience sont assez faibles, l'hypothèse d'une faible motivation de nos sujets pour cette récompense sucrée n'est pas à exclure.

Malgré l'obtention d'un plateau lors des dernières séances du conditionnement intensif, la performance instrumentale exprimée au cours du test RdR reste sensible à la dévaluation de la récompense. Le comportement des animaux est toujours dirigé vers un but.

Chapitre I : Discussion

Les données actuellement publiées dans notre domaine de recherche traitent préférentiellement de la neuroanatomie fonctionnelle des actions et habitudes. De fait l'exploration des paramètres comportementaux permettant le contrôle de la réponse par les actions ou les habitudes sont mal connus. Dans ce premier chapitre, nous nous proposons de tester trois procédures de conditionnement instrumental à travers leur sensibilité à la dévaluation de la récompense. Dans leur ensemble, ces expériences montrent que la réponse instrumentale observée en l'absence de récompense est bien sensible à la dévaluation de la récompense, que celle-ci soit obtenue par aversion conditionnée ou par satiété sélective. Les rats qui refusent de consommer la récompense dévaluée sont également capables de s'abstenir de presser le levier préalablement associé à l'obtention de cette récompense. Cet effet est réversible et dépend d'une représentation actualisée de la valeur de la récompense, ce qui caractérise une réponse dirigée vers un but. Les résultats obtenus soulèvent cependant plusieurs observations.

Conditionnement instrumental et formation des habitudes

La composition de nos trois protocoles de conditionnement visait à évaluer les théories actuellement avancées dans la littérature selon lesquelles les actions se développeraient sous des entraînements courts, mettant en jeu des programmes de renforcement permettant une bonne perception de la corrélation récompense/réponse. Au contraire, la formation des habitudes serait favorisée par la répétition de programmes de renforcement imposant des intervalles de temps longs (Dickinson, 1985). Nous avons donc établi d'une part un programme court (2xFR1+ 4xVI30) au cours de l'Expérience I.1. et d'autre part des protocoles de durée plus longue (Expériences I.3, 2xFR1 + 10xVI30) avec des intervalles de temps étendus (Expérience I.4, 2xFR1 + 5xVI30 + 8xVI60).

Comme nous l'envisagions, la première procédure assez brève (2xFR1, 4xVI30) aboutit à l'expression d'une réponse dirigée vers un but. Ce résultat est également obtenu lors d'un entraînement deux fois plus long (2xFR1, 10xVI30) mais « massé ». Le conditionnement *High* (2xFR1, 5xVI30, 8xVI60) semblait réunir de nombreuses conditions favorables à la formation des habitudes : un intervalle variable allongé (VI60), 15 séances (1 séance/jour), l'obtention de 600 récompenses. De plus, notre protocole aboutit à un faible taux de récompense et permet d'observer une chute drastique de la corrélation récompense-appui dès

la 3^{ème} séance (cf. figure I.α). De fait, nos animaux atteignent un plateau de performance vers la 8^{ème} séance. Toutefois, ceci ne doit pas être interprété comme la preuve d'un surentraînement ni du caractère « habituel » de la réponse instrumentale. En effet, la réponse instrumentale exprimée lors du test RdR reste indéniablement sensible à la valeur actualisée de la récompense.

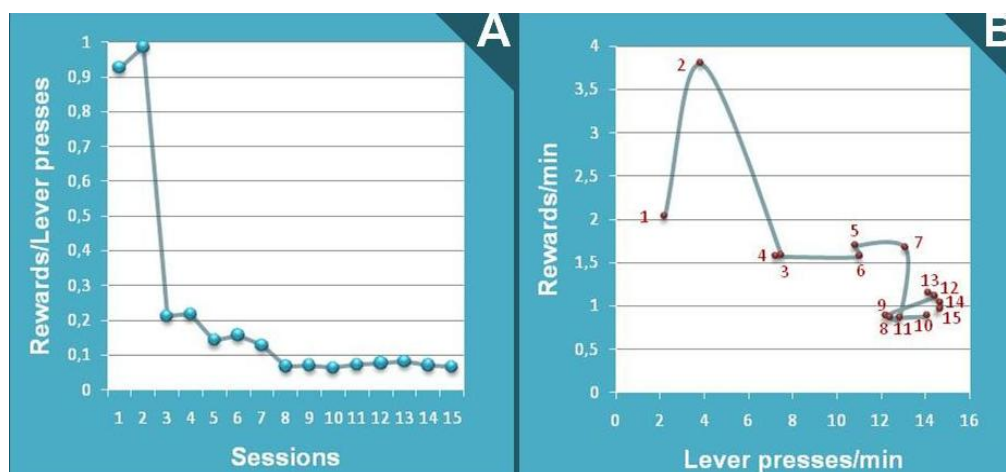


Figure I.α Relation entre taux de récompenses et taux de réponses. (A) Rapport récompense/réponse en fonction des séances au cours du Conditionnement *High*. Initialement proche de 1, Le rapport récompense/réponse chute à 0,2 lors des séances en VI30, puis il se stabilise sous la barre de 0,1 pour les 8 dernières séances. (B) Diagramme de corrélation entre taux de récompenses et taux d'appuis. Au cours des séances en VI (4 à 15), le taux de récompenses obtenues par les sujets ne varie plus avec leur comportement d'appui. (Les chiffres en rouges indiquent le numéro de la séance instrumentale.)

Dans deux articles de recherche récents (Hilário et al., 2007 ; DeRusso et al. ; 2010), le groupe de Rui Costa reproduit l'effet favorable des programmes VI pour l'expression d'habitudes avec des protocoles que nous qualifierons de modérés. Les auteurs attribuent leurs résultats à une différence dans la perception de la contiguïté entre l'action et la récompense. Cette interprétation va pourtant à l'encontre de la définition d'un système A-O sensible à la contingence et d'un système associatif S-R sensible à la contiguïté entre la réponse et la récompense (Balleine et Dickinson, 1998a ; Baleine et O'Doherty, 2010). De plus, chez Adams (1982) l'effet du surentraînement est obtenu même après des séances en FR1. En outre, Yin et al. (2004), décrivent l'apparition d'habitudes chez le rat Long Evans après un nombre de séances réduit. Il reste donc difficile d'interpréter la discordance des résultats de la littérature à moins de les imputer au choix du modèle animal. Nous tenterons d'aborder cette question au cours du chapitre suivant par une analyse des protocoles de dévaluation de la récompense.

Dévaluation de la récompense par satiété sélective

Afin d'évaluer la nature de nos réponses instrumentales, nous avons eu recours aux deux procédures de dévaluation de la récompense classiquement utilisées. De prime abord la procédure de satiété sélective semble plus simple à mettre en place, pourtant elle n'est pas exempte de difficultés (cf. Expérience I.4). En effet, nous avons constaté un niveau de performance instrumentale assez modéré pour la solution sucrée et nous n'avons pas réussi à dévaluer cette récompense. Pourtant, une procédure comparable avait été précédemment utilisée avec succès au sein de notre équipe (cf. Naneix et al., 2009). Toutefois, dans notre expérience, les sujets étaient systématiquement entraînés le matin pour la pastille de grains, l'après-midi pour la solution sucrée. Il est donc possible que le tout premier contact avec la solution sucrée ainsi que les suivants n'aient pas lieu dans des conditions de faim/restriction alimentaires les plus favorables à l'attribution d'une forte valeur hédonique à la récompense. Cette faible motivation pourrait expliquer l'obtention d'une performance instrumentale modérée et l'absence d'une consommation jusqu'à satiété. Le faible taux de réponse observé lors du test, qui avoisine les niveaux émis sur le levier associé à la pastille quand celle-ci est dévaluée conforte également cette interprétation.

En dépit des résultats concernant la solution sucrée, ceux obtenus pour la pastille de grains sont probants : la réponse instrumentale est fortement diminuée suite à la dévaluation de la récompense solide par satiété sélective. On observe également un effet de cette séance sur la quantité de pastilles consommées pendant les tests de consommation.

Dévaluation de la récompense par aversion gustative

Contrairement à une procédure de satiété sélective, le choix d'un conditionnement d'aversion gustative implique de nombreux paramètres expérimentaux. Nous avons sélectionné des paramètres (dose de LiCl injectée : 0,15M, 10mL/kg, quantité de récompense proposée : 40, nombre d'appariements récompense-LiCl : 6) proches de ceux utilisés lors de l'expérience princeps d'Adams et Dickinson (0,15M, 13mL/kg ; 50 récompenses ; 3 appariements ; contexte inchangé, cf. Chapitre II).

Au cours des phases de dévaluation par AGC des Expérience I.1 et I.3, trois séances de dévaluation, donc deux appariements Récompense-LiCl, sont nécessaires avant de pouvoir observer une diminution de la consommation des groupes *Devalued* et nous avons fixé le nombre de nos séances de dévaluation à 6. Comme certains individus acquièrent l'AGC de façon retardée par rapport à l'ensemble du groupe, nous n'obtenons pas un rejet à 100% de la

récompense alimentaire ; il est de 25% pour la première expérience et de 34% pour la troisième.

Il est rapporté dans la littérature que l'AGC est un phénomène rapide qui peut s'obtenir en un seul appariement. Mais cette observation est valable exclusivement dans le cas d'une nourriture nouvelle, de l'emploi d'une dose de LiCl élevée ou d'une quantité de récompenses abondante. En effet, la consommation au préalable de la récompense par le sujet conduit à une inhibition latente (IL) de l'AGC (Domjan et Wilson, 1972 ; Elkins, 1973 ; De La Casa et Lubow, 1995). Ce phénomène correspond au retard de conditionnement du fait de la pré-exposition répétée et non renforcée d'un stimulus avant son association avec le stimulus inconditionnel (Lubow, 1989, Coutureau et al., 1999). Dans sa revue faisant état des résultats actuellement disponibles, Lubow (2009) précise que l'IL concerne en particulier le nombre de séances nécessaires à obtenir les mêmes niveaux de consommation. Nos résultats sont conformes à cette notion puisque nous observons le même retard d'effet (2 appariements) mais un niveau d'aversion plus faible chez les animaux ayant eu une procédure d'entraînement instrumentale plus longue.

Effet de la dévaluation sur la motivation incitative

Nous avons choisi de soumettre nos sujets à 6 séances de dévaluation, ils sont donc réexposés à la récompense de multiples fois avant d'être confrontés au levier. Par le passé, le nombre d'appariements récompense-LiCl nécessaires pour l'observation d'un effet sur la réponse instrumentale a fait l'objet de nombreux débats : Balleine et Dickinson (1991) n'obtenaient un effet de la dévaluation que dans le cas où les animaux étaient réexposés à la récompense après un premier appariement. Par contre, Rescorla (1992) observait une réponse fortement diminuée après un seul appariement Récompense-LiCl. En vue d'expliquer ces différences Balleine et Dickinson ont mené diverses études dont certains résultats sont rapportés dans la figure I.β. Ils ont montré que la dose de LiCl utilisée peut affecter la nature des processus cognitifs mis en jeu. En effet, Rescorla utilise une solution hypertonique (0,6M ; 5mL/kg soit 3mol/kg) et Balleine et Dickinson une solution hypotonique (0,15M ; 5mL/Kg soit 0,75mol/Kg).

Ligne	référence	Anesthésie	LiCl		Injection		Réexposition	Performance	
			[LiCl]	mol	min	Après 6h		appuis	effet
1	a	-	0,15 M	0,75	X		sucré		↘
2	b	-	0,6 M	3	X		sucré		↘
3	b	+	0,6 M	3	X		sucré		↘
4	a	-	0,15 M	0,75		X	sucré		non
5	b	-	0,6 M	3		X	sucré		non
6	a	-	0,15 M	0,75	X		eau		non
7	b	-	0,6 M	3	X		eau		↘
8	b	+	0,6 M	3	X		eau		non

Figure 1.8 Résumé des résultats obtenus par Balleine et Dickinson suite à des injections de LiCl en solution hypotonique (a, 1991) et en solution hypertonique (b, 1992).

Les animaux ont reçu une injection unique de LiCl après une séance instrumentale récompensée par du sucre, dans un état conscient (anesthésie, -) ou sous anesthésie générale (anesthésie, +). Lorsque la solution hypotonique (0,15 M) est injectée, on observe un effet de la dévaluation du sucre seulement si cette injection est immédiate et que l'animal est réexposé au sucre (ligne 1). Dans le cas d'une solution hypertonique la réponse est sensible aux effets du LiCl qu'il y ait réexposition au sucre (ligne 2) ou non (ligne 7). Si cette injection est faite sous anesthésie générale, seule la réexposition à la récompense instrumentale provoquera un effet sur la performance instrumentale (ligne b).

Dans leur ensemble, ces résultats indiquent que l'emploi d'une solution hypertonique provoque immédiatement des troubles gastriques. De ce fait, le contrôle de la réponse s'effectue dans ce cas là par un feedback émotionnel qui pourrait s'apparenter aux processus liés à la peur. Au contraire, le malaise éprouvé suite à l'injection d'une solution moins concentrée engage un processus latent nécessitant une mise à jour de la valeur incitative de la récompense, donc une réexposition à la nourriture dévaluée.

Ces résultats font écho à la théorie de Garcia (1989) qui suggère que lors d'une AGC la nouvelle valeur d'un aliment est encodée seulement pendant la seconde exposition à l'aliment. Lors du premier appariement, le malaise permettrait la formation d'une association entre les propriétés sensorielles de la récompense et le feedback gastrique. Cette association serait évoquée lors de la seconde présentation (appariée ou non) et alors les réactions de dégoûts émises par l'animal permettraient de changer la valeur hédonique de la récompense. Le changement de la motivation incitative ne serait donc observable qu'après cette seconde exposition et un effet sur les performances instrumentales ne serait donc visible qu'après la réexposition.

Compte tenu de l'ensemble de ces données, nous avons choisi une dose modérée de LiCl (0,15 M ; 10 mL/kg soit 1,5 M/kg) qui nous permet de suivre dans le temps l'acquisition par nos sujets de l'aversion gustative (cf. Nelson et Killcross, 2006). En accord avec la théorie

de Garcia, nous observons une diminution de la consommation seulement après la seconde réexposition donc au cours de la 3^{ème} séance aussi bien en Expérience I.1 qu'en Expérience I.2. Ainsi, nous pouvons avancer que la diminution des performances instrumentale observée au cours des phases de tests de ces deux expériences est bien due à un contrôle de l'action par de la valeur incitative de la récompense.

Effets sur les différents types de réponses

Il est intéressant de noter que, contrairement au comportement de consommation, on observe une diminution des visites à la mangeoire dès la 2^{ème} séance de dévaluation chez les sujets recevant du LiCl. L'acquisition de la réponse conditionnée d'évitement semble donc plus rapide que celle du dégoût de la nourriture. Il existe des données dans la littérature qui laisseraient penser que ces deux manifestations seraient deux phénomènes dissociables chez le Rat, par exemple sous l'effet d'un anti-émétique (Parker 2006), ou après lésion du cortex gustatif (Kiefer et Orr, 1992).

Diverses données de la littérature indiquent que la visite de la mangeoire correspond à une réponse conditionnée d'approche, de nature Pavlovienne (Pickens et Holland, 2004). Néanmoins, la nature précise des interactions entre la réponse opérante et la réponse d'approche est à l'heure actuelle toujours sujette à débat. Il semble que ces réponses puissent être acquises de manière indépendante par l'animal et soient dissociables sur les plans comportemental, neurobiologique (Killcross et Coutureau, 2003; Balleine et Ostlund, 2007) et neurocomputationnel (Daw et al., 2005). Un argument fort en faveur de cette dissociation est la persistance de l'évitement de la mangeoire après le surentraînement ou suite à des lésions cérébrales empêchant une action dirigée (Killcross et Coutureau, 2003). Ce travail de thèse portant essentiellement sur les actions et les habitudes, nous avons fait le choix de limiter l'analyse du comportement de visite aux seuls cas où cette analyse apparaissait nécessaire.

En résumé, le travail d'exploration mené au cours de ce premier chapitre nous permet d'affirmer que dans nos conditions expérimentales et chez notre modèle animal, le caractère dirigé d'une action résiste aussi bien à de faibles corrélations entre l'appui et la récompense, qu'à la répétition des séances d'entraînement, ou l'atteinte d'un plateau de performance. Le protocole de renforcement de la première expérience ainsi que la dévaluation par AGC constitueront donc la base des expérimentations menées pour la suite de nos études.

Chapitre II

Chapitre II

Introduction

Les résultats présentés dans le Chapitre I indiquent de la capacité des animaux à modifier leur comportement en fonction des changements de valeur du résultat de leur action, en accord avec un important corpus de données. Lorsque la valeur du résultat est diminuée, les sujets diminuent leur réponse opérante, ce qui suggère d'une part qu'ils encodent durant l'entraînement une relation entre l'action et le résultat et d'autre part qu'ils utilisent une représentation du résultat pour guider leur réponse. Le contenu ainsi que le format de cette représentation sont à l'heure actuelle assez mal connus. Les expériences du Chapitre II visent précisément à documenter les caractéristiques du contenu de la représentation du résultat d'une action, en l'abordant sous différents aspects.

Le premier aspect concerne les mécanismes d'encodage eux-mêmes. Les données de la littérature suggèrent que la réponse de l'animal dépend d'un encodage détaillé du résultat. Nous avons déjà évoqué dans l'introduction générale un nombre important d'études montrant que l'encodage du résultat intégrait les propriétés sensorielles, motivationnelles ainsi que des liens d'association plus élaborés avec la réponse (Balleine et Dickinson, 1998a). De fait, l'Expérience II.1 vise à explorer la spécificité de l'encodage, en mettant en œuvre une procédure discriminative. Dans l'Expérience II.2, nous avons cherché à voir si l'encodage prenait en compte les informations contextuelles acquises lors de la dévaluation. Enfin dans l'Expérience II.3, nous avons cherché à voir, en modulant l'exposition passive à la récompense, dans quelles conditions la représentation prenait en compte une valeur actualisée de l'information.

Le second aspect concerne l'encodage des informations contextuelles dans la représentation du résultat. A l'heure actuelle, il existe peu de données disponibles sur ce thème mais certains auteurs (Colwill et Rescorla, 1990, Rescorla, 1991) ont proposé l'existence d'une structure associative hiérarchique du contrôle de la réponse. Selon cette analyse, en plus des relations binaires (stimulus-résultat, stimulus-réponse et réponse-résultat), il existerait une forme de « contrôle supérieur » sur la production de la performance

pouvant prendre la forme S – (R-O) (voir Rescorla, 1991 pour une description de cette analyse). Ce type de proposition trouve son écho dans des élaborations plus récentes (voir par exemple Balleine et O'Doherty, 2010) qui considèrent que les différentes contingences apprises par le sujet contrôlent différents aspects de la réponse. Ainsi, en se basant sur les expériences de Transfer-Pavlovien instrumental (TPI, voir Holmes et al., 2010), on peut considérer que l'activation des mécanismes de sélection et de production de la réponse opérante résulte de l'évocation d'une représentation du résultat par une contingence Pavlovienne acquise au cours de la même tâche instrumentale (Balleine et O'Doherty, 2010). Ainsi, dans le cadre de ces deux interprétations, les informations contextuelles pourraient jouer un rôle différent dans le contrôle puisqu'elles pourraient soit directement signaler le résultat, soit moduler l'activation de sa représentation et donc la sélection de l'action. Les Expériences II.3 et II.4 visent à explorer ces deux possibilités.

Expérience II.1

Action dirigée et spécificité de l'encodage de la récompense

Cette expérience a pour but de tester la spécificité de l'encodage de l'association Réponse-Résultat. Nous avons voulu vérifier si l'effet d'une dévaluation de la nourriture effectuée dans le contexte de la cage de conditionnement est bien spécifique à une action associée à cette nourriture. Pour cela, nous avons entraîné des animaux avec deux types de récompenses. Lors de séances différentes, les animaux devaient soit appuyer sur un levier pour obtenir une récompense d'un certain type, soit aller collecter une récompense d'un autre type qui était distribuée gratuitement. Dans une phase ultérieure, la valeur de l'une des récompenses était dévaluée par aversion gustative de manière identique aux expériences du chapitre 1. Enfin, nous avons comparé les performances d'appui de deux groupes de rats pour lesquels la nourriture dévaluée était celle associée au levier ou au contraire celle distribuée gratuitement.

A. Matériel et méthode

a. Sujets

Cette expérience concerne 40 rats mâles naïfs de la souche Long-Evans (380 à 450 g). Les conditions d'hébergement sont identiques à celles décrites dans l'Expérience I.1. Concernant la restriction alimentaire, la quantité de nourriture fournie est ajustée au nombre quotidien des séances de conditionnement, 15g pour 1 séance, 13 g pour 2 séances.

b. Dispositif expérimental

Les 8 cages de conditionnement utilisées correspondent à celles décrites en Expérience I.1. Pour les besoins de cette expérimentation, nous avons modifié l'agencement de la paroi gauche afin qu'elle comporte 2 mangeoires séparées au milieu par la présence du module comprenant le levier rétractable (figure II.1.1, panneau A). Ces éléments ainsi disposés, la compétition entre les comportements d'appui et de visite se trouve diminuée du fait de leur proximité. Lors de la phase d'entraînement, l'appui sur le levier est associé à une distribution de nourriture uniquement dans la mangeoire de gauche. Deux récompenses différentes sont

utilisées : la pastille formulée à base de grains (cf. Expérience I.1) et une pastille formulée à base de sucrose et d'huile végétale (45 mg, Rodent Purified Diet, Bio_Serv).

Afin de ne pas biaiser nos résultats, nous avons pris soin de charger les distributeurs de gauche et de droite de façon contrebalancée avec les deux formulations de pastilles. Néanmoins, pour faciliter la compréhension du protocole de conditionnement par le lecteur, nous ne décrivons que le cas suivant : la récompense de type *grains* est distribuée dans la mangeoire de gauche, la récompense de type *sucré* dans la mangeoire de droite.

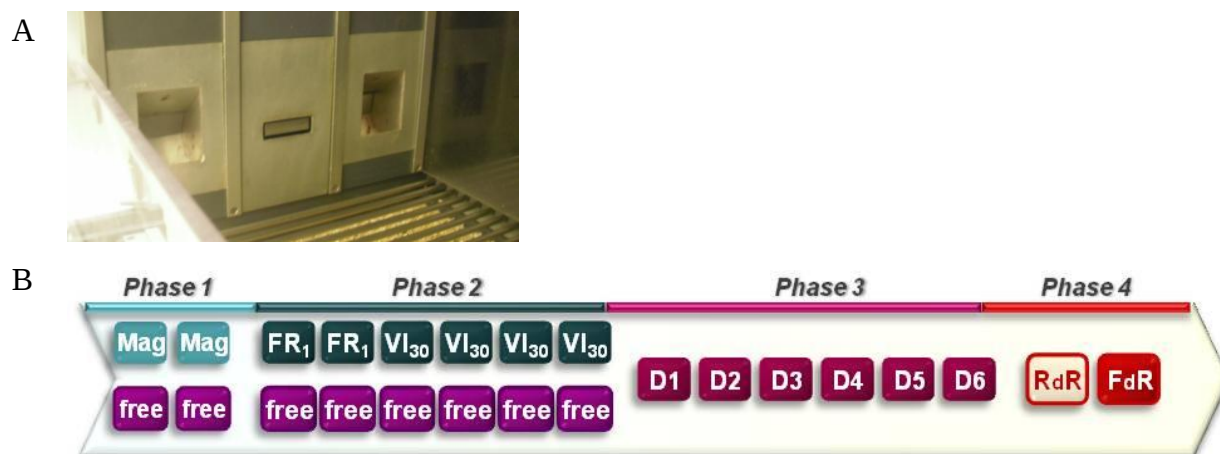


Figure II.1.1 Photographie du dispositif expérimental (A) et schématisation du déroulement de l'expérience (B).

Le module comprenant le levier rétractable sépare les deux mangeoires. Elles sont chacune reliées à un distributeur contenant l'une des deux récompenses.

Les séances Mag et Free sont construites selon le même programme (VI60) de distribution automatique.

Mag : séance d'entraînement à la mangeoire de gauche ; Free : séance d'entraînement à la mangeoire de droite ; FR1 et VI30 : séances de conditionnement instrumental ; D : séance de dévaluation par aversion gustative conditionnée ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; FdR : test pendant lequel la réponse est renforcée.

c. Protocole expérimental

Ce protocole se déroule sur 4 phases (figure II.1.1, panneau B). Les 2 premières phases comportent 2 séances par jour, séparées au minimum par 3 heures de repos au sein de l'animalerie. Les autres phases comportent une seule séance par jour.

- *Phase 1 : Entraînement aux mangeoires*

La phase d'entraînement à la mangeoire comporte 4 séances et se déroule conformément à l'Expérience I.1. Les animaux sont successivement entraînés à aller collecter les récompenses dans les mangeoires de gauche et de droite.

- *Phase 2 : Conditionnement instrumental et Conditionnement non contrôlé*

Cette phase se divise quotidiennement en 2 temps : le matin est consacré à l'une des récompenses (*grains*) et l'après-midi à l'autre type de récompenses (*sucre*).

- Conditionnement instrumental

Au cours des séances matinales, le levier est présent et l'animal contrôle par son appui la distribution des récompenses (*grains*) dans la mangeoire de gauche. La procédure comprend les 6 séances d'entraînement (2 séances en FR1 + 4 séances en VI30).

- Conditionnement non contrôlé

L'après-midi, l'entraînement à la mangeoire de droite initié en Phase 1 se poursuit par 6 séances supplémentaires et identiques. Le levier est donc absent et la distribution de la deuxième récompense (*sucre*) n'est pas contrôlée par le sujet mais entièrement programmée.

- *Phase 3 : Dévaluation de l'une des récompenses*

Au cours des 6 séances de dévaluation (cf. Expérience I.1), une seule des 2 récompenses est distribuée. La moitié des animaux (groupe *Controlled*) reçoivent la récompense (*grains*) qui était précédemment délivrée en résultat de l'activité d'appui. L'autre moitié des sujets forment le groupe *Free* et reçoivent l'autre récompense (*sucre*) dont ils n'ont jamais pu contrôler la distribution. Selon le traitement reçu (LiCl ou NaCl), les sujets sont répartis au sein des 4 groupes suivants :

- ***Controlled-Devalued*** (*ctled-D*) : la récompense associée au levier est dévaluée ;
- ***Controlled-NonDevalued*** (*ctled-ND*) : la récompense associée au levier n'est pas dévaluée ;
- ***Free-Devalued*** (*free-D*) : la récompense non contrôlable est dévaluée ;
- ***Free-NonDevalued*** (*free-ND*) : la récompense non contrôlable n'est pas dévaluée.

- *Phase 4 : Tests comportementaux RdR et FdR*

Le lendemain de la dernière séance de dévaluation, les animaux sont soumis à un test RdR de 15 minutes suivi le surlendemain par un test FdR (cf. Expérience I.1).

d. Analyses statistiques

Les résultats obtenus (*taux d'appuis, taux de visites, indices de consommation*) sont analysés à l'aide des variables indépendantes suivantes :

- le facteur **séance** (1- 6) (cf. Expérience I.1) ;
- le facteur **mangeoire** (*Controlled/Free*), correspondant aux mangeoires de gauche et de droite ;
- le facteur **dévaluation** (*NonDevalued/Devalued*) (cf. Expérience I.1) ;
- le facteur **formule** (*grains/sucre*), correspondant aux 2 formulations des pastilles de nourriture ;
- le facteur **contrôle** (*Controlled/Free*), correspondant au caractère contrôlé ou non-contrôlé de la distribution de la récompense en Phase 2 ;
- le facteur **groupe**, correspondant à la répartition pseudo-aléatoire des sujets au sein de groupes expérimentaux distincts.
-

B. Résultats

Sur les 40 sujets, un seul n'a pas réussi à acquérir la tâche d'appui sur le levier. Il a donc été exclu de l'analyse des résultats. Les effectifs finaux au sein des groupes expérimentaux sont les suivants : *ctled-D*, n= 10 ; *ctled-ND*, n= 9 ; *free-D*, n= 10 ; *free-ND*, n= 10.

- **Entraînement aux mangeoires**

Les animaux ne présentent pas de préférence pour l'une des récompenses. Lors de la 2ème séance, en moyenne 39 des pastilles de grains et 39 des pastilles de sucre sont consommées (données non représentées).

- **Conditionnement instrumental et Conditionnement non contrôlé**

Lors des séances consacrées à la mangeoire de type free, les animaux vont chercher les récompenses distribuées automatiquement. A chaque séance, ils consomment l'intégralité des 40 pastilles distribuées (données non représentées).

Les résultats du conditionnement instrumental sont représentés sur la figure II.1.2.

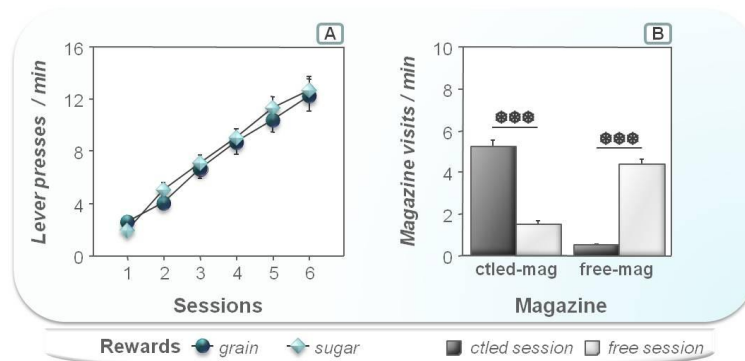


Figure II.1.2 Evolution des taux d'appuis au cours de l'entraînement instrumental (A) et taux des visites aux mangeoires lors de la dernière journée d'entraînement (B) ($\pm$ erreur standard).

Une moitié des sujets sont entraînés à appuyer sur le levier pour obtenir la pastille de grain, l'autre pour la pastille de sucre. La nature de la récompense n'a aucun effet sur la vitesse d'acquisition de la tâche.

A la fin de la phase d'entraînement, les sujets visitent de façon préférentielle la mangeoire *ctled* lorsque la distribution dépend de l'action sur le levier et la mangeoire *free* quand la distribution est automatique.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Le panneau A représente l'évolution des taux d'appuis sur le levier au fil des séances de conditionnement instrumental. Les animaux récompensés par une pastille de grains appuient à des niveaux comparables à ceux des animaux récompensés par des pastilles de sucre. Le type de nourriture distribuée n'influe pas sur la vitesse et le niveau d'apprentissage de la réponse instrumentale.

- * Une ANOVA à deux facteurs confirme l'effet significatif du facteur **séance** ($F_{5,185} = 105,128$; $p < 0,001$) l'absence d'effet de la **formule** ($F_{1,37} = 0,221$; ns) et l'absence d'interaction entre ces deux facteurs ($F_{5,185} = 0,644$; ns).

On peut comparer sur le panneau B, le comportement de visites aux 2 mangeoires en fonction du type de séance à laquelle l'animal est soumis au cours de la dernière journée de la phase d'entraînement. Les animaux visitent en priorité la mangeoire dans laquelle est distribuée la récompense : la mangeoire *ctled* en présence du levier, la mangeoire *free* dans le cas d'une distribution non contrôlée.

- * Une ANOVA révèle l'absence d'effet du **contrôle** ($F_{1,38} = 0,253$; ns) mais confirme un effet significatif du facteur **mangeoire** ($F_{1,38} = 26,844$; $p < 0,001$) et une interaction significative entre ces deux facteurs ($F_{1,38} = 350,106$; $p < 0,001$).
- * Le test SNK indique que chaque mangeoire est significativement plus visitée lors des séances où la récompense y est distribuée.

A la fin de l'entraînement, les quatre groupes expérimentaux, *ctled-D*, *ctled-ND*, *free-D* et *free-ND*, sont constitués de façon équilibrée sur la base de leurs performances comportementales (données non représentées).

- * L'analyse de variance confirme que les 4 groupes ne diffèrent ni dans leur taux de référence d'appuis ($F_{3,35} = 0,413$; ns) ni dans leur taux de référence de visites ($F_{3,35} = 0,777$; ns).

• Dévaluation de l'une des récompenses

On constate, sur la figure II.1.3, que les résultats de la phase de dévaluation sont conformes à ceux obtenus lors de la première expérience du Chapitre 1.

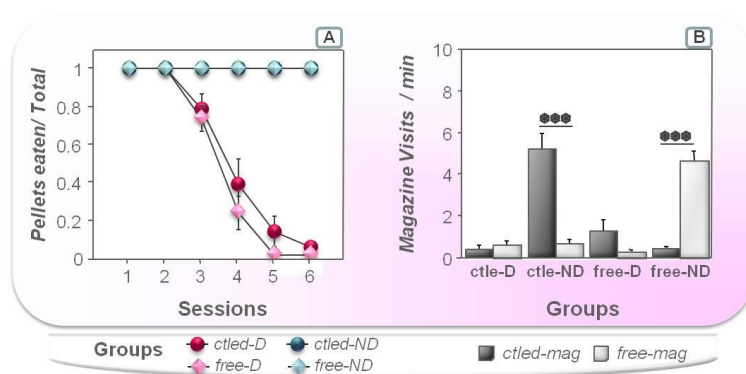


Figure II.1.3 Evolution de l'indice de consommation au fil des séances de dévaluation (A) et taux de visites aux mangeoires lors de la dernière séance (B) ($\pm$ erreur standard).

Les sujets acquièrent une aversion gustative à la même vitesse qu'elle soit conditionnée pour la récompense précédemment associée au levier ou non. De même, contrairement à des sujets qui reçoivent du NaCl, ils visitent peu la mangeoire qui distribue la récompense dévaluée.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

On peut suivre sur le panneau A l'évolution des indices de consommation au cours de la phase de dévaluation. Les animaux des groupes *ctled-ND* et *free-ND* ont un indice de consommation toujours égal à 1. Inversement, à partir de la 3ème séance, les indices moyens des groupes *ctled-D* et *free-D* diminuent progressivement et à la même vitesse. Le fait que l'animal ait pu précédemment contrôler ou non la distribution de la récompense n'influe pas sur le conditionnement ultérieur d'une aversion gustative.

- * Ces résultats sont analysés à l'aide d'une ANOVA aux facteurs multiples (dévaluation, séance, contrôle, formule). Elle confirme les effets significatifs des facteurs **dévaluation** ($F_{1,31} = 314,09$; $p < 0,001$) et **séance** ($F_{5,155} = 125,509$; $p < 0,001$) ainsi qu'une interaction significative **dévaluation x séance** ($F_{5,155} = 125,509$; $p < 0,001$). En revanche, elle rejette tout effet des facteurs **contrôle** ($F_{1,31} = 1,293$; ns) et **formule** ($F_{1,31} = 1,557$; ns) ainsi que l'existence d'interaction supplémentaire entre les facteurs ($F_5 < 1,26$; $p_5 > 0,20$).
- * A partir des données de la dernière séance de dévaluation, un test SNK localise des différences significatives entre tous les groupes, exceptions faites des groupes *ctled-ND* et *free-ND* entre eux et des groupes *ctled-D* et *free-D* entre eux.

Les résultats obtenus concernant le comportement de visite à la mangeoire sont également conformes à ceux de la première expérience. Les animaux des groupes *ctled-ND* et

free-ND ont des taux relativement stables, tandis que les sujets des groupes *ctled-D* et *free-D* visitent de moins en moins la mangeoire (données non représentées).

Le panneau B de la figure II.1.3 représente les taux moyens de visites dans les 2 mangeoires lors de la dernière séance de dévaluation. Les groupes *ctled-ND* et *free-ND* visitent fréquemment la mangeoire qui distribue la pastille non dévaluée, en revanche, ils négligent la mangeoire non récompensée. En comparaison, les groupes *ctled-D* et *free-D* visitent très peu la mangeoire distribuant la récompense dévaluée. Seul le groupe *free-D* semble présenter en contrepartie une augmentation de ses visites dans la mangeoire non récompensée.

- * Une analyse statistique confirme la description de ces résultats. Il y a un effet significatif du facteur **dévaluation** ($F_{1,35} = 65,099$; $p < 0,001$) sur les taux de visites, mais pas d'effet du facteur **contrôle** ($F_{1,35} = 0,06$; ns), ni du facteur **mangeoire** ($F_{1,35} = 1,5$; ns). Les interactions **mangeoire x contrôle** et **mangeoire x dévaluation x contrôle** sont significatives ($F_{1,35} = 47,986$; $p < 0,001$ et $F_{1,35} = 85,475$; $p < 0,001$), mais pas les interactions **dévaluation x contrôle** et **mangeoire x dévaluation** ($F_{1,35} = 2,038$; ns et $F_{1,35} = 0,158$; ns).
- * Le test SNK localise un comportement significativement différent vis-à-vis des 2 mangeoires pour les groupes *ctled-ND* et *free-ND*. En revanche, le comportement de visite du groupe *free-D* n'est pas significativement différent.

• Tests comportementaux

- Test RdR

Les résultats du test RdR sont représentés sur la figure II.1.4.

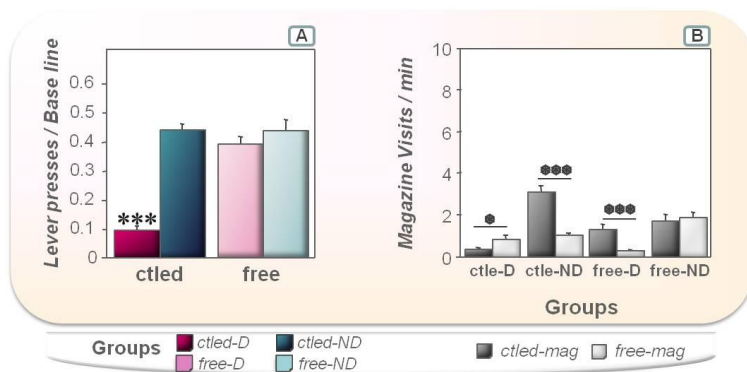


Figure II.1.4 Taux d'appuis (A) et taux de visites aux mangeoires lors du test RdR(B) ($\pm$ erreur standard). Les performances instrumentales sont diminuées seulement pour les sujets ayant acquis une aversion gustative pour la récompense précédemment distribuée par le levier. Les sujets qui ont acquis une aversion gustative pour l'autre type de récompense se comportent comme les animaux des groupes ayant reçu du NaCl. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Les résultats critiques sont représentés dans le panneau A de la figure II.1.4. Comme le montre cette figure, les performances des groupes *ctled* et *free* sont différentes. En effet, seuls les animaux du groupe *ctled-D* diminuent fortement leur activité opérante, ce qui suggère que leur réponse est médiée par une représentation de la valeur, conformément aux résultats de l'expérience 1. De manière intéressante, les animaux du groupe *free-D* ont une activité d'appui tout à fait comparable à celle des animaux contrôles *free-ND* et *Ctled-ND*.

- * Les taux d'appuis sont analysés à l'aide d'une ANOVA à facteurs multiples. Elle révèle des effets significatifs des facteurs **dévaluation** ($F_{1,31} = 22,100$; $p < 0,001$) et **contrôle** ($F_{1,31} = 11,642$; $p < 0,01$) et une interaction significative **dévaluation x contrôle** ($F_{1,31} = 12,714$; $p < 0,01$). Par contre, elle confirme l'absence d'effet du facteur **formule** ($F_{1,31} = 1,557$; ns) et rejette toute interaction entre ce facteur et les deux autres ($F_s < 1,5$; $p_s > 0,20$).
- * Un test SNK confirme que le groupe *ctled-D* diffère significativement des 3 autres.

Les résultats concernant les visites dans les mangeoires sont représentés sur le panneau B, en valeurs brutes. Les animaux se comportent différemment selon la récompense distribuée lors de la phase de dévaluation. Ainsi, le groupe *ctled-ND* présente une fréquence de visite plus haute dans la mangeoire associée au levier et récompensée lors de la phase précédente. Au contraire, les animaux du groupe *ctled-D* visitent peu les deux mangeoires avec une diminution plus marquée pour la mangeoire associée à la récompense dévaluée. Le groupe *free-ND* visite indifféremment les deux mangeoires. Au contraire, le comportement de visite du groupe *free-D* est inégal, la mangeoire associée à la récompense dévaluée est délaissée et la mangeoire associée au levier est visitée.

- * Une ANOVA confirme les effets significatifs des facteurs **dévaluation** ($F_{1,35} = 52,549$; $p < 0,001$) et **mangeoire** ($F_{1,35} = 22,960$; $p < 0,001$) mais il n'y a pas d'effet du facteur **contrôle** ($F_{1,35} = 0,015$; ns). Il y a des interactions significatives **mangeoire x dévaluation** ($F_{1,35} = 7,197$; $p < 0,05$) et **mangeoire x dévaluation x contrôle** ($F_{1,35} = 53,159$; $p < 0,001$), mais pas d'interaction entre les facteurs **contrôle** et **dévaluation** ($F_{1,35} = 1,999$; ns) ou **contrôle** et **mangeoire** ($F_{1,35} = 1,953$; ns).
- * Un test SNK confirme qu'à l'exception du groupe *free-ND*, les 3 autres groupes ont un comportement significativement différent vis-à-vis des 2 mangeoires.

- Test FdR

La figure II.1.5 représente les résultats du test avec appuis récompensés. Les performances de chaque animal sont rapportées à ses taux de référence.

Conformément au test RdR, seuls les animaux du groupe *ctled-D* appuient avec une très faible fréquence sur le levier. Les trois autres groupes, y compris le groupe *free-D* appuient à des niveaux comparables ou supérieurs à leurs taux de référence (panneau A).

- * Une ANOVA à facteurs multiples confirme les effets significatifs des facteurs **dévaluation** ($F_{1,31} = 26,813$; $p < 0,001$) et **contrôle** ($F_{1,31} = 99,154$; $p < 0,001$) et une interaction significative **dévaluation x contrôle** ($F_{1,31} = 69,214$; $p < 0,001$). Il n'y a pas d'effet du facteur **formule** ($F_{1,31} = 1,557$; ns) ni d'interaction entre ce facteur et les deux autres ($F_s < 1$; $p_s > 0,35$).
- * Un test SNK confirme que seuls les groupes *ctled-ND* et *free-ND* ne diffèrent pas entre eux.

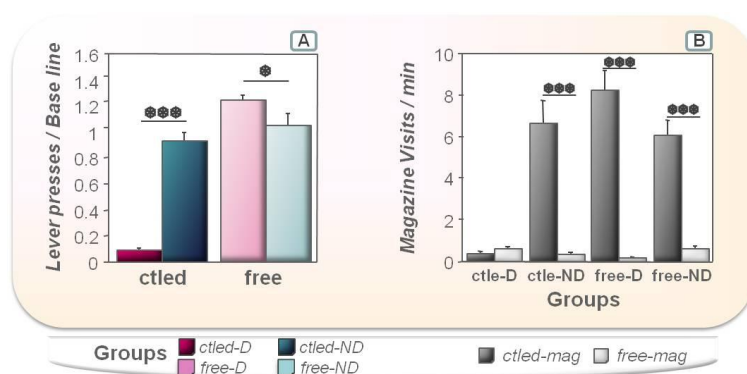


Figure II.1.5 Taux d'appuis (A) et taux de visites aux mangeoires lors du test FdR(B) ($\pm$ erreur standard).

Le comportement des sujets lors du test récompensé est conforme à celui du test non récompensé : seuls les sujets possédant une aversion pour la récompense ctled présentent des taux d'appuis et de visites faibles. La mangeoire free n'étant pas récompensée, aucun groupe ne la visite. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Pour les groupes *ctled-ND*, *free-ND* et *free-D*, le comportement de visites se trouve modifié quand l'appui est de nouveau récompensé (panneau B). En effet, ils visitent tous activement la mangeoire où est distribuée la récompense contrôlée et négligent l'autre mangeoire. Le comportement du groupe *ctled-D* reste inchangé, les sujets ne visitent que très peu les deux mangeoires.

- * L'analyse de variance indique des effets significatifs des facteurs **dévaluation** ($F_{1,35} = 3,845$; $p < 0,05$), **mangeoire** ($F_{1,35} = 141,769$; $p < 0,001$) et **contrôle** ($F_{1,35} = 18,913$; $p < 0,001$). Les interactions entre ces 3 facteurs sont toutes significatives ($F_s > 20$; $p_s < 0,05$).
- * Un test SNK confirme que la mangeoire sous contrôle du levier est significativement plus visitée chez l'ensemble des groupes à l'acceptation du groupe *ctled-D*.

Enfin, les niveaux de consommation de la récompense sont conformes aux comportements d'appuis et de visites. L'ensemble des animaux des groupes *ctled-ND*, *free-ND* et *free-D* consomment la totalité des 40 pastilles obtenues. Les sujets du groupe *ctled-D* possédant une aversion pour la récompense, ne se distribuent en moyenne que $14,2 \pm 2$ unités, qu'ils consomment à hauteur de 20%.

Les quantités de pastilles distribuées et consommées sont analysées par deux ANOVAs à 3 facteurs. Elles confirment les effets des facteurs **dévaluation** et **contrôle** qui interagissent significativement ($F_{s\ 1,31} = 143,912$; $p_s < 0,001$ et $F_{s\ 1,31} = 71,159$; $p_s < 0,001$). Elles rejettent l'effet du facteur **formule** et l'existence de toute interaction entre ce facteur et les 2 autres ($F_s = 0,216$; ns et $F_s = 0,005$; ns).

Conclusion

Cette expérience met en avant l'aptitude de nos sujets à distinguer et à s'adapter aux diverses situations qu'ils rencontrent successivement dans un même contexte. Ainsi, ils modulent leur comportement de visite en fonction de la distribution de la récompense et de la présence du levier au cours de la Phase 2. Si les sujets des groupes ctled-D et free-D acquièrent de façon similaire l'AGC, leurs comportements d'appui lors de la phase de test diffèrent en tout points. L'acquisition d'une AGC pour la récompense non contrôlable dans le contexte de conditionnement opérant ne modifie pas les comportements d'appui et de visite visant à la consommation de l'autre récompense. Ces comportements ne sont affectés que par la dévaluation de la récompense qui leur est associée.

Les données de cette expérience permettent ainsi d'affirmer l'existence d'une spécificité de l'encodage de la récompense contrôlée par la réponse dans le système associatif dédié à l'action dirigée.

Expérience II.2

Action dirigée et spécificité de l'encodage du contexte

Cette expérience a pour but de tester l'importance de l'encodage des informations contextuelles liées à la dévaluation dans la représentation du résultat. Pour cela, nous avons comparé deux conditions. La condition « *Same* » correspond à la condition des expériences précédentes, c'est-à-dire que la dévaluation a lieu dans le contexte de conditionnement. Dans la condition « *Different* » par contre, la dévaluation a lieu dans un autre contexte. Néanmoins, dans tous les cas, les performances des animaux sont évaluées dans le contexte initial d'entraînement.

A. Matériel et méthode

a. Sujets

Les sujets de cette expérience sont 32 rats mâles naïfs de la souche Long-Evans (300 à 375 g). Les conditions d'hébergement et de restriction alimentaire sont identiques à celles décrites précédemment (cf. Expérience I.1).

b. Dispositif expérimental

Le protocole de cette expérience nécessite l'utilisation de 2 lots distincts de cages. Le premier lot est constitué des 8 cages de conditionnement avec une mangeoire au milieu de la paroi gauche et à son extrémité gauche un levier (cf. Expérience I.1). Le deuxième lot se trouve dans une salle d'expérimentation adjacente, il correspond aux 8 cages de consommation précédemment décrites dans l'Expérience I.4. La récompense alimentaire correspond à la pastille de grains utilisée lors des expériences précédentes (cf. Expérience I.1).

c. Protocole expérimental

Cette expérience comporte les 4 phases du protocole établi en Expérience I.1 (figure II.2.1). Seule la phase de dévaluation présente une variation majeure.

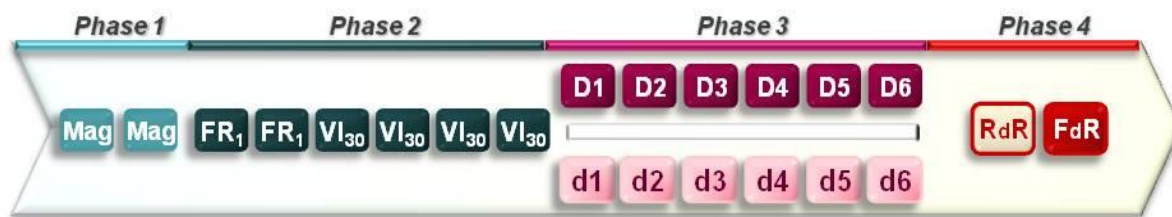


Figure II.2.1 Schématisation du déroulement de l'expérience.

Suite à la phase de conditionnement instrumental, les sujets sont divisés en deux groupes selon le lieu des séances de dévaluation : dans les cages de conditionnement D ou dans des cages de consommation d.

Mag : séance d'entraînement à la mangeoire ; FR1 et VI30 : séances de conditionnement instrumental ; D et d : séances de dévaluation par aversion gustative conditionnée ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; FdR : test pendant lequel la réponse est renforcée.

- *Phase 1 : Entraînement à la mangeoire (cf. Expérience I.1)*
- *Phase 2 : Conditionnement instrumental*

Les 2 derniers jours de l'entraînement instrumental (cf. Expérience I.1), les animaux sont amenés au sein de la salle de consommation et sont placés individuellement dans une cage de consommation pour 20 minutes d'habituation à ce nouvel environnement. Aucune nourriture ou boisson n'est disponible au cours de ces séances.

- *Phase 3 : Dévaluation de la récompense dans 2 contextes différents*

La phase de dévaluation s'étend sur 6 séances et se déroule dans les cages de conditionnement ou dans les cages de consommation.

La moitié des sujets sont placés dans le même contexte que celui du conditionnement opérant, ils forment le groupe *Same*. Le déroulement des séances est alors identique à celui décrit en Expérience I.1.

Pour l'autre moitié des animaux, les séances se déroulent dans les cages de consommation, ils forment le groupe *Different*. Le sujet a alors libre accès à 40 pastilles de nourriture contenues dans la coupelle de verre. Au bout de 15 minutes, il est retiré de la cage de consommation afin de recevoir une injection intra-péritonéale de LiCl ou de NaCl (cf. Expérience I.1), puis il est replacé au sein de l'animalerie.

Les groupes ainsi constitués sont les suivants :

- ***Same-Devalued (same-D)*** : la récompense est dévaluée dans le même contexte que celui du conditionnement opérant ;
- ***Same-NonDevalued (same-ND)*** : la récompense non dévaluée est distribuée dans le même contexte que le conditionnement opérant ;
- ***Different-Devalued (diff-D)*** : la récompense est dévaluée dans un contexte différent du conditionnement opérant ;

- **Different-NonDevalued** (*diff-ND*) : la récompense non dévaluée est distribuée dans un contexte différent du conditionnement opérant.

- *Phase 4 : Tests comportementaux*

L'expérimentation se termine par un test RdR de 15 minutes suivi le lendemain par un test FdR (cf. Expérience I.1).

d. Analyses statistiques

Les résultats obtenus (*taux d'appuis, taux de visites, indices de consommation*) sont analysés à l'aide des variables indépendantes suivantes :

- le facteur **séance** (1-6) (cf. Expérience I.1) ;
- le facteur **groupe** (cf. Expérience II.1) ;
- le facteur **dévaluation** (*NonDevalued/Devalued*) (cf. Expérience I.1) ;
- le facteur **contexte** (*Same/Different*), correspondant aux cages utilisées pour la phase 3 qui différent ou non de celles du conditionnement instrumental.

B. Résultats

Un rat, tombé malade au cours de l'expérimentation, a dû être sacrifié. Les effectifs finaux au sein des groupes sont les suivants : *same-D*, n= 8 ; *same-ND*, n= 8 ; *diff-D*, n= 7 ; *diff-ND*, n= 8.

- *Entraînement à la mangeoire*

L'ensemble des sujets s'habitue bien au dispositif expérimental, ils visitent la mangeoire et consomment la pastille de grains (données non représentées).

- *Conditionnement instrumental*

Tous les animaux apprennent sans difficulté la réponse instrumentale. A la fin du conditionnement, ils sont divisés en quatre groupes équilibrés sur la base de leurs performances (données non représentées).

- * Les analyses de variance ne révèlent aucune différence entre leurs taux d'appuis de référence ($F_{1,27} = 0,077$; ns) ou entre leurs taux de visites de référence ($F_{1,27} = 0,641$; ns).

- *Dévaluation de la récompense dans deux contextes différents*

L'évolution de l'indice moyen de consommation, tout au long des séances de dévaluation, peut être suivie pour chacun des groupes sur la figure II.2.2.

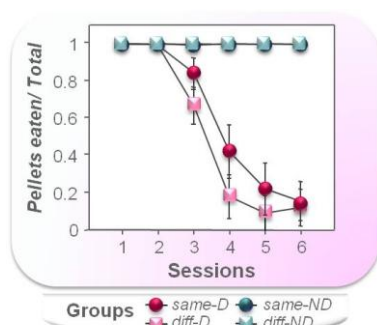


Figure II.2.2 Evolution de l'indice de consommation au fil des séances de dévaluation ($\pm$ erreur standard).

L'aversion gustative conditionnée est acquise à la même vitesse et atteint les mêmes niveaux dans les deux contextes.

Le contexte accueillant les séances de dévaluation influe peu sur le comportement des animaux. En effet, l'ensemble des animaux se comportent conformément aux résultats de l'Expérience I.1. Ainsi, les rats du groupe *same-ND* et *diff-ND* consomment à chaque fois l'intégralité des récompenses. Au contraire, la consommation des groupes *same-D* et *diff-D* ne fait que diminuer à partir de la 3ème séance de dévaluation.

- * Une ANOVA à 3 facteurs confirme cette description des résultats. Elle révèle les effets significatifs des facteurs **dévaluation** ($F_{1,27} = 96,813$; $p < 0,001$) et **séances** ($F_{5,135} = 66,234$; $p < 0,001$) ainsi qu'une interaction significative **dévaluation x séance** ($F_{5,135} = 66,434$; $p < 0,001$). Par contre, elle ne montre pas d'effet significatif du facteur **contexte** ($F_{1,27} = 1,159$; ns) ni d'interaction entre ce facteur et les deux autres ($F_s < 1,5$; $p_s > 0,25$).
- * Un test SNK, réalisé sur les indices de la 6ème séance, localise des différences significatives entre tous les groupes, sauf pour les groupes *same-ND* et *diff-ND* entre eux et pour les groupes *same-D* et *diff-D* entre eux.

Concernant le comportement de visite, seules les données des groupes *same-ND* et *same-D* ont été recueillies. Elles sont conformes aux résultats obtenus lors de l'Expérience I.1 : les rats recevant du NaCl continuent à visiter la mangeoire, les rats recevant du LiCl visitent de moins en moins la mangeoire (données non représentées).

- * Une analyse de variance révèle une tendance d'effet du facteur **dévaluation** ($F_{1,14} = 3,821$; $p = 0,0709$), un effet significatif du facteur **séance** ($F_{5,70} = 5,764$; $p < 0,001$) ainsi qu'une interaction significative entre ces deux facteurs ($F_{5,70} = 11,323$; $p < 0,001$).

- Tests comportementaux

- Test RdR

Les résultats obtenus pour le premier test non récompensé sont observables sur la figure II.2.3 Les performances de chaque animal sont rapportées à ses taux de référence.

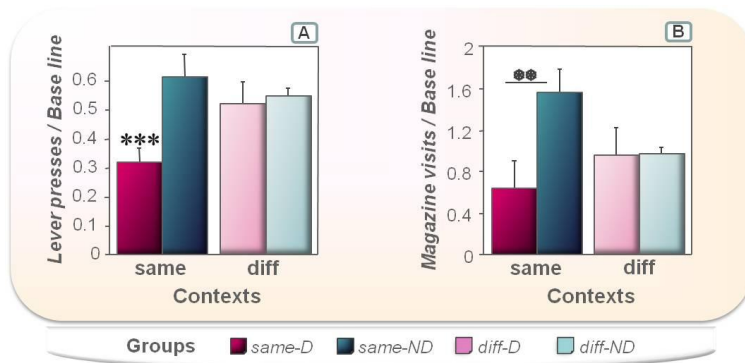


Figure II.2.3 Taux d'appuis (A) et taux de visites au cours du test RdR (B) ($\pm$ erreur standard).

La dévaluation de la récompense a un effet sur le comportement opérant seulement dans le cas où les séances de dévaluation ont eu lieu dans le même contexte que l'apprentissage de la tâche instrumentale. Seuls les animaux du groupe same-ND visitent activement la mangeoire.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Le panneau A est consacré au comportement d'appuis sur le levier. Les résultats obtenus pour les animaux des groupes *same* sont conformes à ceux obtenus en Expérience I.1. On constate bien une sensibilité à la dévaluation de la récompense : les rats du groupe *same-D* appuient moins sur le levier que les rats du groupe *same-ND*. Les résultats sont tout autres chez les animaux des groupes dont la phase de dévaluation a eu lieu dans des cages de consommation. En effet, les rats *diff-D* appuient activement et de façon comparable aux sujets du groupe *diff-ND*, la dévaluation de la récompense n'influe pas sur leurs performances instrumentales.

- * L'analyse de variance indique une absence d'effet du facteur **contexte** ($F_{1,27} = 1,367$; ns) mais un effet significatif du facteur **dévaluation** ($F_{1,27} = 6,614$; $p < 0,05$) ainsi qu'une interaction significative **contexte x dévaluation** ($F_{1,27} = 4,688$; $p < 0,05$).
- * Le test SNK indique que seul le groupe *same-D* diffère significativement des trois autres groupes, qui ne diffèrent pas entre eux.

Le panneau B représente le comportement de visite à la mangeoire. Les groupes *Same* se comportent conformément à l'Expérience I.1 : le groupe *same-ND* est plus actif à la mangeoire que le groupe *same-D*. Au contraire, malgré leur aversion pour la récompense, les animaux du groupe *diff-D* se comportent comme ceux du groupe *diff-ND*. Ils expriment des taux de visites proche de leurs taux de référence, à des niveaux intermédiaires entre ceux des groupes *same-D* et *same-ND*.

- * Une ANOVA indique une absence d'effet du facteur **contexte** ($F_{1,27} = 0,433$; ns) mais un effet significatif du facteur **dévaluation** ($F_{1,27} = 4,556$; $p < 0,05$) ainsi qu'une interaction significative **contexte x dévaluation** ($F_{1,27} = 4,283$; $p < 0,05$).
- * Un test SNK indique que seuls les groupes *same-D* et *same-ND* diffèrent entre eux de façon significative.

- Test FdR

On peut observer les résultats du test FdR sur la figure II.2.4. Les performances de chaque animal sont rapportées à ses taux de référence.

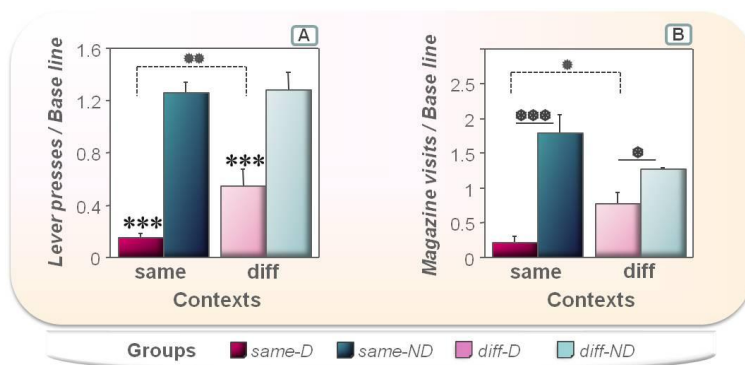


Figure II.2.4 Taux d'appuis (A) et taux de visites au cours du test FdR (B) ($\pm$ erreur standard).

Quand l'appui est de nouveau récompensé, les performances instrumentales sont diminuées chez les deux groupes conditionnés pour une aversion gustative conditionnée pour la récompense distribuée. L'effet de la dévaluation est également observé pour les taux de visites à la mangeoire.

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Les résultats en terme de performance instrumentale sont représentés sur le panneau A. Les groupes *Same* ont un comportement conforme à celui attendu, le groupe *same-D* a un niveau d'appuis beaucoup plus faible que celui du groupe *same-ND*. Lorsque la récompense est de nouveau distribuée, le comportement opérant se trouve changé chez les groupes *Different*. En effet, les sujets du groupe *diff-D* appuient alors sur le levier avec une fréquence nettement plus faible que celle du groupe *diff-ND*.

- * Une ANOVA révèle des effets significatifs des facteurs **contexte** ($F_{1,27} = 4,795$; $p < 0,05$) et **dévaluation** ($F_{1,27} = 85,771$; $p < 0,001$), et une tendance à l'interaction **contexte x dévaluation** ($F_{1,27} = 3,518$; $p = 0,0716$).
- * Un test SNK indique que seuls les groupes *same-ND* et *diff-ND* ne diffèrent pas entre eux. Le comportement du groupe *diff-D*, bien que se rapprochant de celui du groupe *same-D*, en est significativement différent.

Le panneau B représente les résultats obtenus pour le comportement de visite. Comme attendu, les animaux du groupe *same-D* visitent moins la mangeoire que ceux du groupe *same-ND*. De même, lorsque l'appui est récompensé, les sujets du groupe *diff-D* diminuent leurs visites à la mangeoire en comparaison du groupe *diff-ND*.

- * L'analyse de variance indique une absence d'effet du facteur **contexte** ($F_{1,27} = 0,025$; ns) mais un effet significatif du facteur **dévaluation** ($F_{1,27} = 39,951$; $p < 0,001$) ainsi qu'une interaction significative **contexte x dévaluation** ($F_{1,27} = 10,928$; $p < 0,01$).
- * Le test SNK révèle que tous les groupes diffèrent significativement entre eux.

Le nombre moyen de récompenses obtenues lors du test FdR et les indices de consommation moyens sont représentés respectivement sur les panneaux A et B de la figure II.2.5.

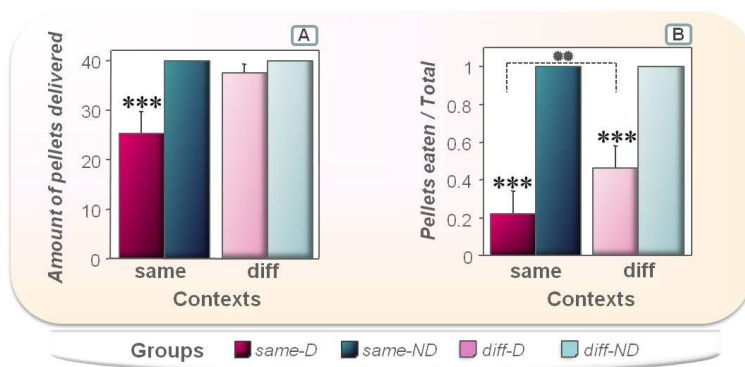


Figure II.2.5 Quantité de récompense obtenue (A) et consommée (B) au cours du test RdR ($\pm$ erreur standard).

La quantité de pastilles distribuée diffère selon le contexte où a eu lieu la dévaluation. Quand elle est effectuée dans un contexte différent, les sujets obtiennent un nombre de récompenses proche du critère maximum. Néanmoins, ils en consomment moins de la moitié. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Les rats des groupes *same-ND* et *diff-ND* consomment la totalité des 40 récompenses qu'ils se sont distribuées. Au contraire, les animaux du groupe *same-D* ne consomment en moyenne que 22% des $25,4 \pm 4,4$ pastilles dont ils ont provoqué la distribution. Les sujets du groupe *diff-D* obtiennent en moyenne $37,6 \pm 1,8$ récompenses. Cette proximité avec le critère maximum, s'explique par la quantité d'appuis qui est certes faible mais assez étalée sur la durée de la séance pour remplir les conditions de récompense du programme en VI30. Les sujets du groupe *diff-D* expriment néanmoins une aversion gustative pour les récompenses ainsi distribuées, ils en consomment moins de la moitié.

- * Concernant le nombre de récompenses distribuées, l'analyse de variance indique des effets significatifs des facteurs **contexte** ($F_{1,27} = 6,422$; $p < 0,05$) et **dévaluation** ($F_{1,27} = 12,556$; $p < 0,01$) ainsi qu'une interaction significative **contexte x dévaluation** ($F_{1,27} = 6,422$; $p < 0,05$).
- * Un test SNK confirme que seul le groupe *same-D* diffère significativement des trois autres groupes.
- * Concernant les indices de consommation, une ANOVA confirme l'effet significatif du **facteur dévaluation** ($F_{1,27} = 66,125$; $p < 0,001$) mais rejette l'existence d'un effet pour le facteur **contexte** ($F_{1,27} = 2,297$; ns) ou d'une interaction entre ces deux facteurs ($F_{1,27} = 2,297$; ns).
- * Le test SNK indique que seuls les groupes *same-ND* et *diff-ND* ne diffèrent pas entre eux.

Conclusion

Dans leur ensemble, les résultats de cette expérience indiquent que le contexte de dévaluation joue un rôle déterminant dans la sensibilité des animaux aux changements de valeur.

Cet ensemble d'expériences démontrent également la robustesse de la procédure, à la fois en terme d'entraînement et de dévaluation. En effet, à l'issue de la dévaluation, les deux groupes dévalués ont un niveau de consommation comparable.

Par contre, il faut noter que lors du test récompensé, le contexte de dévaluation conserve une influence partielle sur les taux d'appui et la consommation des pastilles.

Expérience II.3

Action dirigée et rôle du contexte dans la mise à jour de la valeur de la récompense

Cette troisième expérience se situe dans la continuité de l'expérience précédente. Dans l'Expérience II.2, nous avons montré qu'un changement de contexte entre l'entraînement et la dévaluation altérerait de manière très importante l'adaptation de la réponse à des changements de valeur. Ce résultat pourrait suggérer que la mise à jour de la valeur se fait de manière dépendante du contexte. Par ailleurs, lors du test FdR précédent, les animaux montraient bien les effets de l'aversion, comme si une réexposition, dans le contexte de conditionnement, à la nourriture dévaluée suffisait pour mettre à jour la valeur incitative de la récompense. L'Expérience II.3. a pour but de tester directement cette hypothèse. Elle reprend les mêmes conditions « *Same* » et « *Different* » de l'Expérience II.2. Néanmoins, de manière à permettre la mise à jour de la valeur du résultat, des animaux du groupe « *Different* » sont réexposés à la récompense dans le contexte d'entraînement avant le test RdR.

A. Matériel et méthode

a. Sujets et dispositif expérimental

Trente-deux rats mâles naïfs de la souche Long-Evans (300-370g) participent à cette expérience. Les conditions d'hébergement et de restriction alimentaire sont identiques à celles décrites dans l'Expérience I.1. Le dispositif expérimental correspond à celui des 2 lots de cages décrits en Expérience II.2.

b. Protocole expérimental

Ce protocole comporte les 4 phases de l'expérience précédente. Il diffère uniquement par le traitement au LiCl de tous les sujets et par l'insertion d'une séance dite de réexposition entre la Phase 3 et la Phase 4 (figure II.3.1).

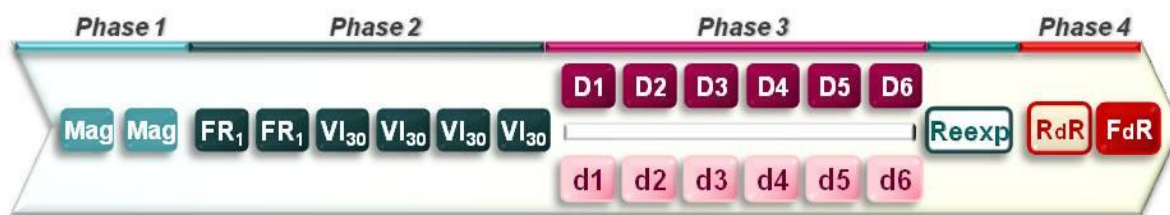


Figure II.3.1 Schématisation du déroulement de l'expérience.

Mag : séance d'entraînement à la mangeoire ; FR1 et VI30 : séances de conditionnement instrumental ; D et d : séances de dévaluation par aversion gustative conditionnée ; Reexp : séance pendant laquelle le sujet est placé dans la cage de conditionnement au contact de la récompense ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; FdR : test pendant lequel la réponse est renforcée.

- *Phase 3 : Dévaluation de la récompense dans 2 contextes différents*

La récompense est dévaluée pour l'ensemble des rats qui forment donc 2 groupes distincts en fonction du contexte de dévaluation : le groupe *Same* et le groupe *Different* (cf. Expérience II.2).

Séance de Réexposition

Elle consiste à réexposer la moitié des sujets à la récompense alimentaire dans le contexte de la cage du conditionnement instrumental. Cette unique séance se déroule comme suivant : 15 pastilles sont préalablement disposées dans la mangeoire, le rat est placé dans sa cage de conditionnement puis les lumières s'allument pour une durée de 10 minutes. A la fin de la séance, les lumières s'éteignent et le sujet est directement replacé dans l'animalerie.

Suite à cette séance les groupes expérimentaux sont les suivants :

- **Same-Controlled** (*same-Ctled*) : la récompense est dévaluée dans le même contexte que celui du conditionnement opérant et le sujet n'est pas réexposé ;
- **Same-Reexposed** (*same-Reexp*) : le sujet est réexposé à la récompense préalablement dévaluée dans le même contexte que celui du conditionnement opérant ;
- **Different-Controlled** (*diff-Ctled*) : la récompense est dévaluée dans un contexte différent de celui du conditionnement et le sujet n'est pas réexposé.
- **Different-Reexposed** (*diff-Reexp*) : le sujet est réexposé dans le contexte du conditionnement à la récompense préalablement dévaluée dans un contexte différent;

c. Analyses statistiques

Les résultats obtenus (*taux d'appuis, taux de visites, indices de consommation*) sont analysés à l'aide des variables indépendantes suivantes :

- le facteur **séance** (1-6) (cf. Expérience I.1) ;

- le facteur **groupe** (cf. Expérience I.1) ;
- le facteur **dévaluation** (*NonDevalued/Devalued*) (cf. Expérience I.1) ;
- le facteur **contexte** (*Same/Different*) (cf. Expérience II.2) ;
- le facteur **réexposition** (*Controlled/Reexposed*), correspond à la réexposition ou non des sujets à la récompense dans le contexte d'entraînement.

B. Résultats

Deux sujets n'ayant pas acquis d'aversion gustative pour la récompense ont été exclus de l'étude. Les effectifs finaux au sein des groupes sont les suivants : *same-Reexp*, $n=8$; *same-Ctled*, $n=6$; *diff-Reexp*, $n=8$; *diff-Ctled*, $n=8$.

• *Entraînement à la mangeoire*

Tous les sujets s'habituent au dispositif expérimental et visitent la mangeoire afin de consommer la pastille de grains (données non représentées).

• *Conditionnement instrumental*

Les sujets retenus pour l'étude apprennent la tâche d'appui sur le levier sans difficulté (données non représentées). A la fin du conditionnement opérant, ils sont répartis en 2 groupes *Same* et *Different*, puis ils sont subdivisés en 4 groupes à la fin de la Phase 3.

* Une ANOVA réalisée à postériori confirme que les 4 groupes, *same-Reexp*, *same-Ctled*, *diff-Reexp* et *diff-Ctled* sont équilibrés. A l'issue de l'entraînement instrumental, leurs taux de référence d'appuis ($F_{3,26} = 0,208$; ns) et de visites ($F_{3,26} = 1,096$; ns) ne différaient pas entre eux.

• *Dévaluation de la récompense dans deux contextes différents*

La figure II.3.2 représente l'évolution des indices de consommation au fil des séances de dévaluation en fonction du contexte où elles ont pris place.

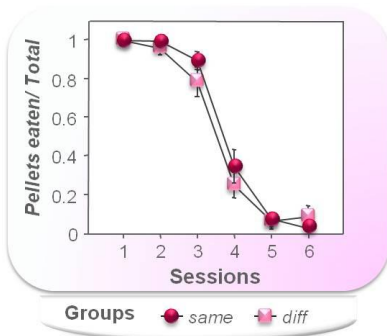


Figure II.3.2 Evolution de l'indice de consommation au fil des séances de dévaluation ($\pm$ erreur standard). L'aversion gustative conditionnée est acquise à la même vitesse et atteint les mêmes niveaux dans les deux contextes.

Conformément aux résultats de l'Expérience II.2, il n'y a pas de différence entre les animaux des groupes *Same* et *Different*, ils acquièrent une aversion gustative pour la récompense au même rythme et atteignent des niveaux finaux équivalents. Ils consomment en moyenne 4% des récompenses lors de l'ultime séance de dévaluation.

- * Une ANOVA confirme l'effet significatif du facteur **séance** ($F_{5,140} = 216,885$; $p < 0,001$) et l'absence d'effet du facteur **contexte** ($F_{1,28} = 0,625$; ns) ou d'interaction entre les deux facteurs ($F_{5,140} = 1,072$; ns). Une ANOVA supplémentaire réalisée sur les résultats de la 6ème séance n'indique pas de différence entre les 4 groupes *same-Reexp*, *same-Ctled*, *diff-Reexp* et *diff-Ctled* ($F_{1,28} = 1,187$; ns).

• Séance de Réexposition

La figure II.3.3 représente les résultats obtenus pour les sujets soumis à la séance de réexposition dans la cage de conditionnement instrumental.

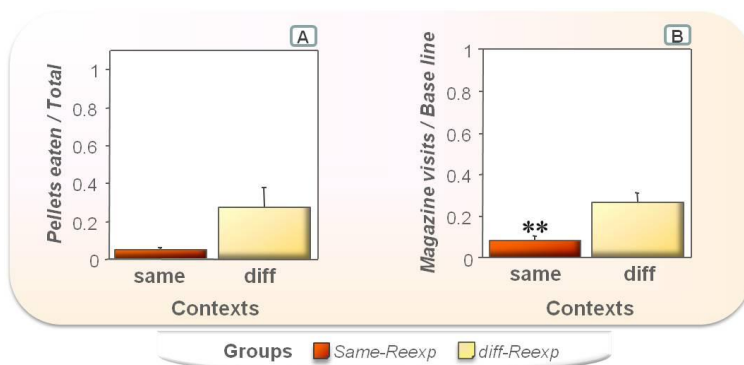


Figure II.3.3 Indices de consommation (A) et taux de visites à la mangeoire (B) lors de la séance de réexposition ($\pm$ erreur standard). Les deux groupes présentent des indices de consommation et des taux de visites caractéristiques d'animaux possédant une aversion gustative conditionnée pour la récompense. Néanmoins, l'effet de la dévaluation semble plus marqué chez les sujets ayant appris la dévaluation dans le contexte du conditionnement opérant. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

On constate sur le panneau A que malgré la faible valeur de leur indice moyen de consommation, les animaux *diff-Reexp* consomment en moyenne plus de pastilles (27%) que les sujets du groupe *same-Ctle* (4%) qui ont été conditionnés à l'AGC dans ce contexte.

- * L'analyse de variance indique effectivement une tendance à un effet du facteur **contexte** ($F_{1,12} = 3,838$; $p = 0,0737$).

Le panneau B représente les taux moyen de visites rapportés aux taux de référence. La description des résultats est analogue : les rats *diff-Reexp* visitent la mangeoire à un niveau faible (26%) mais plus élevé que les rats du groupe *same-Ctled* (6%).

- * Une ANOVA confirme un effet significatif du facteur **contexte** ($F_{1,12} = 9,486$; $p < 0,01$) sur le comportement de visite à la mangeoire.

- *Tests comportementaux*

- **Test RdR**

Les résultats du test RdR sont représentés sur la figure II.3.4. Ils sont exprimés par rapport aux taux de référence.

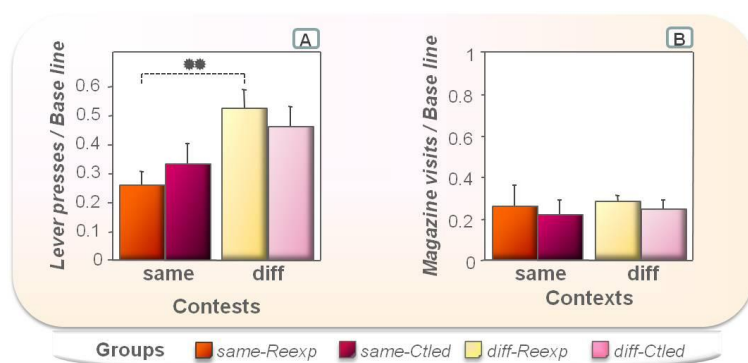


Figure II.3.4 Taux d'appuis (A) et taux de visites au cours du test RdR (B) ($\pm$ erreur standard).

Les sujets ayant acquis l'aversion en dehors de la cage de conditionnement se comportent de façon indifférenciée qu'ils aient été soumis ou non à la séance de réexposition. Ils appuient plus sur le levier que des animaux soumis aux séances de dévaluation dans le même contexte que l'acquisition de la tâche. En revanche, tous les groupes visitent peu la mangeoire. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Le panneau A expose les résultats obtenus pour le comportement d'appui. Leur patron est comparable à celui de l'expérience précédente. Les animaux des groupes *same-Ctled* et *same-Reexp* expriment en moyenne 29% de leur comportement d'appuis de base ; les animaux des groupes *diff-Ctled* et *diff-Reexp* appuient plus fréquemment avec un niveau de 49%. Excepté l'écart entre les niveaux de réponses des groupes *diff-Reexp* et *same-Reexp* qui semble accru, on ne constate pas de différence majeure entre les animaux qui ont été soumis à la séance de réexposition et ceux qui n'y ont pas participé.

- * Une ANOVA confirme l'effet significatif du facteur **contexte** ($F_{1,26} = 9,838$; $p < 0,01$) et rejette l'effet du facteur **réexposition** ($F_{1,26} = 0,005$; ns) ainsi que l'existence d'une interaction entre les deux facteurs ($F_{1,26} = 1,173$; ns).
- * Un test SNK localise une différence significative entre les groupes *diff-Reexp* et *same-Reexp*.

Les résultats du comportement de visite sont représentés sur le panneau B. Les 4 groupes se comportent indépendamment du contexte de dévaluation ou de la réexposition, ils visitent la mangeoire à un niveau diminué de moitié par rapport à leur comportement de référence.

- * Une ANOVA confirme l'absence d'effet des facteurs **contexte** ($F_{1,26} = 0,137$; ns) et **réexposition** ($F_{1,26} = 0,304$; ns) ainsi que l'absence d'interaction entre ces deux facteurs ($F_{1,26} = 0,001$; ns).

- Test FdR

On peut observer les résultats du test FdR sur la figure II.3.5. Les performances de chaque animal sont rapportées à ses performances de référence.

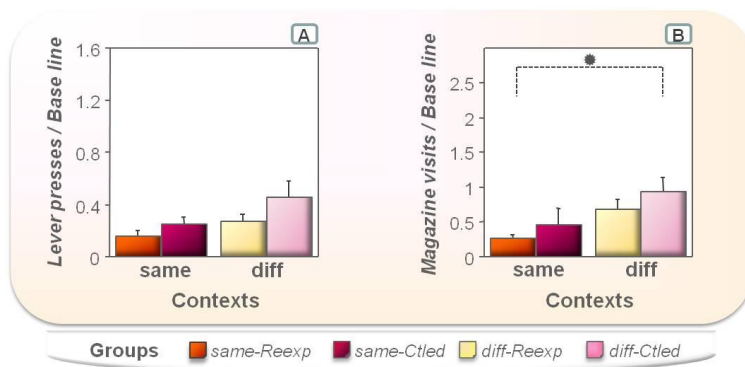


Figure II.3.5 Taux d'appuis (A) et taux de visites au cours du test FdR (B) ($\pm$ erreur standard).

Quand l'appui est de nouveau récompensé, l'ensemble des sujets diminuent leur performance comportementale. L'effet de la dévaluation est moins marqué pour les animaux ayant acquis l'aversion dans un autre contexte et n'ayant pas été soumis à la séance de réexposition.

Le panneau A rapporte les résultats concernant le comportement opérant. Ils sont similaires aux résultats de l'Expérience II.2, lorsque la récompense est distribuée les animaux des groupes *Different*, diminuent leur taux d'appuis à des niveaux plus proches de ceux des groupes *Same*. Cette diminution semble plus marquée pour le groupe *diff-Reexp*.

- * L'analyse statistique révèle une tendance à un effet du facteur **contexte** ($F_{1,26} = 3,737$; $p=0,0642$) mais pas d'effet du facteur **réexposition** ($F_{1,26} = 2,9$; ns) ni d'interaction entre les facteurs ($F_{1,26} = 0,288$; ns).

Les résultats concernant le comportement de visite sont exposés sur le panneau B. L'ensemble des animaux diminuent leurs visites à la mangeoire lorsque la récompense est de nouveau distribuée. Cette diminution est moins marquée chez les groupes *Different*, conformément à l'expérience précédente. Il n'y a pas de différence notable entre les groupes *Reexposed* et les groupes *Controlled*.

- * Une ANOVA révèle un effet significatif du facteur **contexte** ($F_{1,26} = 7,125$; $p<0,05$) mais pas d'effet du facteur **réexposition** ($F_{1,26} = 1,911$; ns) ni d'interaction entre les facteurs ($F_{1,26} = 0,047$; ns).

- * Le test SNK localise une différence significative seulement entre les groupes *diff-Ctle* et *same-Reexp*.

Les résultats concernant le nombre de récompenses obtenues et leur consommation sont conformes aux performances instrumentales et de visites (figure II.3.6, panneaux A et B).

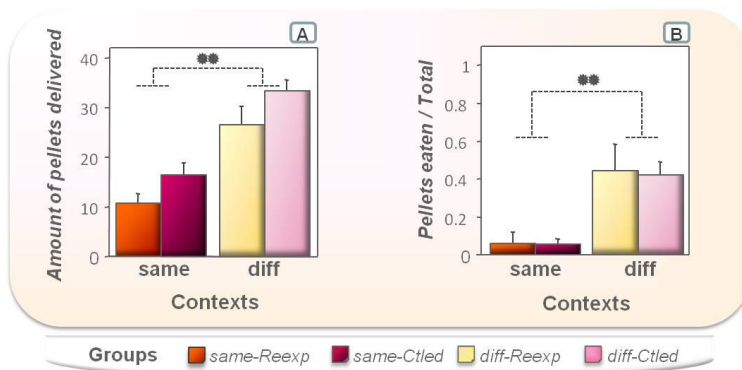


Figure II.3.6 Quantité de récompenses obtenues (A) et consommées au cours du test FdR (B) ($\pm$ erreur standard).

Les sujets soumis aux séances de dévaluation en dehors des cages de conditionnement obtiennent plus de récompense. Néanmoins, leurs indices de consommation indiquent bien une aversion gustative conditionnée pour la récompense. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Les sujets des groupes *same-Ctle* et *same-Reexp* obtiennent peu de récompense et les consomment peu, 5% des $13,1 \pm 1,7$ pastilles distribuées. Au contraire, les animaux des groupes *diff-Ctle* et *diff-Reexp* se distribuent en moyenne $30 \pm 2,4$ récompenses, mais ils en consomment moins de la moitié (43%). Il semblerait que les groupes *same-Reexp* et *diff-Reexp* obtiennent moins de récompenses en comparaison à leur groupe *Controlled* respectif.

- * Concernant le nombre de récompenses distribuées, l'analyse de variance indique des effets significatifs des facteurs **contexte** ($F_{1,26} = 31,872$; $p < 0,001$) et **réexposition** ($F_{1,26} = 4,905$; $p < 0,05$) mais pas d'interaction entre les facteurs ($F_{1,27} = 0,030$; ns).
- * Le test SNK indique que tous les groupes diffèrent entre eux, sauf les 2 groupes *Same* entre eux et sauf les 2 groupes *Different* entre eux.
- * Concernant les indices de consommation, une ANOVA confirme l'effet significatif du facteur **contexte** ($F_{1,26} = 16,878$; $p < 0,001$) mais rejette l'existence d'un effet pour le facteur **réexposition** ($F_{1,26} = 0,03$; ns) ou d'une interaction entre ces deux facteurs ($F_{1,26} = 0,009$; ns).
- * Le test SNK indique que tous les groupes diffèrent entre eux, sauf les 2 groupes *Same* entre eux et sauf les 2 groupes *Different* entre eux.

Conclusion

Dans leur ensemble, les résultats de cette expérience montrent que la réexposition, dans le contexte d'entraînement, à la récompense dévaluée dans un autre contexte est inefficace pour restaurer la sensibilité de l'action instrumentale aux changements de valeur.

Néanmoins, ces résultats de cette expérience sont conformes aux résultats de l'expérience précédente, en cela qu'ils montrent que le changement de contexte lors de la phase de dévaluation affecte drastiquement la sensibilité à la dévaluation du comportement d'appui, mais également le comportement de consommation.

Expérience II.4

Action dirigée et rôle du contexte dans la signalisation de la valeur de la récompense

Les résultats de l'Expérience II.2 ont montré que le contexte de dévaluation est un élément déterminant de l'ajustement de la réponse à des changements de valeur. Les résultats de l'Expérience II.3 suggèrent que la mise à jour de la valeur du résultat n'est pas en elle-même contrôlée par le contexte, puisqu'une réexposition à la récompense dans le contexte d'entraînement est sans effet. Ainsi, l'Expérience II.4 a pour but de poursuivre cette étude en explorant la capacité du contexte à contrôler la prise en compte de la valeur du résultat. Pour cela, nous avons mis en œuvre une procédure d'entraînement dans laquelle les animaux étaient entraînés de manière alternative dans des contextes différents pour obtenir la même récompense. Dans une étape ultérieure, la récompense a été dévaluée soit dans un contexte, soit dans un autre. Les résultats précédents suggèrent que le rat pourrait considérer la récompense comme dévaluée ou non selon le contexte, et en tenir compte pour adapter sa réponse. Nous avons donc testé les mêmes animaux dans le contexte de dévaluation (condition *same*) et dans le contexte alternatif (condition *different*). Si le contexte joue bien un rôle dans la prise en compte de la valeur du résultat lors du test RdR, on doit s'attendre à ce que la performance des animaux soit réduite par la dévaluation en condition *Same* mais pas en condition *Different*.

A. Sujets

Cette expérience concerne 40 rats mâles naïfs de la souche Long-Evans (360 à 420 g). Les conditions d'hébergement sont identiques à celles décrites dans l'Expérience I.1. Le soir, la quantité de nourriture fournie est ajustée au nombre quotidien de séances de conditionnement (cf. Expérience II.1).

a. Dispositif expérimental

Le dispositif comporte les 8 cages de conditionnement. Pour les besoins du protocole, la moitié des cages sont conservées sous la configuration classiquement utilisée jusqu'ici (contexte *classic*, cf. Expérience I.1), tandis que les 4 autres cages sont totalement ré-agencées (figure II.4.1). Afin de créer un contexte modifié (contexte *pattern*), des motifs floraux et étoilés noirs et blancs ont été apposés sur toutes leurs faces, à l'exception du plafond et du

fond de cage. Ce dernier est constitué d'une plaque de PVC noire qui réduit la profondeur de la cage. Le plancher est fait d'une plaque de PVC de surface dépolie présentant un motif en damier. La paroi latérale droite est en outre protégée par l'apposition d'une plaque de PVC transparent de 3 mm d'épaisseur dans laquelle ont été percés des orifices pour permettre l'accès à la mangeoire et au levier respectivement au milieu et à gauche de la paroi. La distribution de la récompense est accompagnée du claquement d'un relais. Une même récompense alimentaire est utilisée pour les deux cages, la pastille de grains (cf. Expérience I.1).



Figure II.4.1 Photographie de la deuxième cage de conditionnement.

Les sujets apprennent la tâche instrumentale dans deux contextes différents. Dans le deuxième contexte des motifs ont été accolés à chaque paroi (A et B), le levier et la mangeoire se trouvent sur la paroi latérale droite (C).

b. Protocole expérimental

L'ensemble des sujets est soumis à 4 phases d'expérimentation (figure II.4.2).

Les deux premières phases sont effectuées au rythme de 2 séances quotidiennes. Ainsi, chaque animal est entraîné dans les 2 types de cages de conditionnement. Tout au long du protocole, nous avons pris soin de contrebalancer l'ordre de passage dans les deux contextes. Néanmoins, pour faciliter la compréhension du protocole par le lecteur, nous limiterons la description au cas suivant : la séance du matin se déroule dans la cage ayant la configuration habituelle *classic*, la séance de l'après midi dans la cage avec motifs *pattern*.



Figure II.4.2 Schématisation du déroulement de l'expérience (B).

Mag : séance d'entraînement à la mangeoire ; FR1 et VI30: séances de conditionnement instrumental ; D : séance de dévaluation par aversion gustative conditionnée ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; FdR : test pendant lequel la réponse est renforcée.

- *Phase 1 : Entraînement aux mangeoires*

Chaque individu est entraîné à aller chercher la pastille de grains le matin dans le contexte *classic* et l'après-midi dans le contexte *pattern* (cf. Expérience I.1).

- *Phase 2 : Conditionnements instrumentaux dans deux contextes*

Au cours de cette phase, l'animal apprend à appuyer sur le levier pour obtenir la récompense, le matin dans le contexte *classic* puis l'après-midi dans le contexte *pattern*. Cette phase s'étend sur 7 jours et se compose pour chaque type de cage de 3 séances FR1 suivies par 4 séances VI30 (cf. Expérience I.1).

- *Phase 3 : Dévaluation de la récompense dans l'un des deux contextes*

La phase de dévaluation se déroule pareillement aux séances décrites en Expérience I.1 et pour sa totalité dans une seule des cages. Ainsi, selon le traitement reçu (LiCl ou NaCl) les sujets sont divisés en 2 groupes *Devalued* et *NonDevalued*, et en 4 sous-groupes selon le contexte de la cage : *classic-Devalued*, *classic-NonDevalued*, *pattern-Devalued* et *pattern-NonDevalued*.

- *Phase 4 : Tests comportementaux*

Le premier jour de tests, l'ensemble des sujets sont soumis à 2 tests RdR (15 minutes chacun) puis le lendemain à 2 tests FdR (cf. Expérience I.1, séparation minimum de 3 heures entre 2 séances de test). En effet, les animaux sont testés, de façon contrebalancée, dans chacun des 2 contextes des phases d'entraînement. Nous qualifierons de *same*, la condition expérimentale où le contexte correspond à celui de la dévaluation. Au contraire, nous nommerons *different*, la condition où l'animal est testé dans le deuxième contexte qui n'est pas celui de la dévaluation.

Par exemple, un animal du groupe *classic-Devalued* testé dans le contexte *classic* est en condition *same*, puis lors du test dans le contexte *pattern* il est en condition *different*. L'ordre des tests est contrebalancé au sein des groupes.

Ainsi, nos sujets peuvent se trouver, en fonction du traitement reçu en Phase 3 et du contexte du test, dans les 4 conditions expérimentales suivantes :

- ***Devalued-Same (D-same)*** : la récompense est dévaluée et le test a lieu dans le contexte de la phase de dévaluation ;

- **Devalued-Different (D-diff)** : la récompense est dévaluée et le test a lieu dans le deuxième contexte qui diffère de celui de la phase de dévaluation ;
- **NonDevalued-Same (ND-same)** : la récompense n'est pas dévaluée et le test a lieu dans le contexte de la phase de dévaluation ;
- **NonDevalued-Different (ND-diff)** : la récompense n'est pas dévaluée et le test a lieu dans le deuxième contexte qui diffère de celui de la phase de dévaluation.

C. Analyses statistiques

Dans cette expérience, chaque animal est son propre témoin. Les résultats obtenus (*taux d'appuis, taux de visites, indices de consommation*) sont analysés à l'aide des variables indépendantes suivantes :

- le facteur **séance** (1- 6) (cf. Expérience I.1) ;
- le facteur **groupe** (cf. Expérience II.1) ;
- le facteur **cage** (*classic/pattern*), correspondant aux 2 types de cages de conditionnement ;
- le facteur **dévaluation** (*NonDevalued/Devalued*) (cf. Expérience I.1) ;
- le facteur **contexte** (*same/different*), correspondant à la différence entre le contexte de dévaluation et le contexte de la séance de test.

B. Résultats

Trois animaux qui n'ont pas réussi à acquérir une aversion gustative pour la récompense ont été exclus de l'étude. Le groupe *Devalued* a donc un effectif de 17 sujets, et le groupe de *NonDevalued* de 20 sujets.

• *Entraînement aux mangeoires*

Les sujets s'habituent sans difficulté aux cages de conditionnement et vont chercher la récompense dans les mangeoires. Le 2ème jour, ils consomment les pastilles de grains à hauteur de 98% dans le contexte *classic* et de 98% dans le contexte *pattern* (données non représentées).

- *Conditionnements instrumentaux dans deux contextes*

Les résultats obtenus au cours du conditionnement opérant sont représentés sur la figure II.4.3.

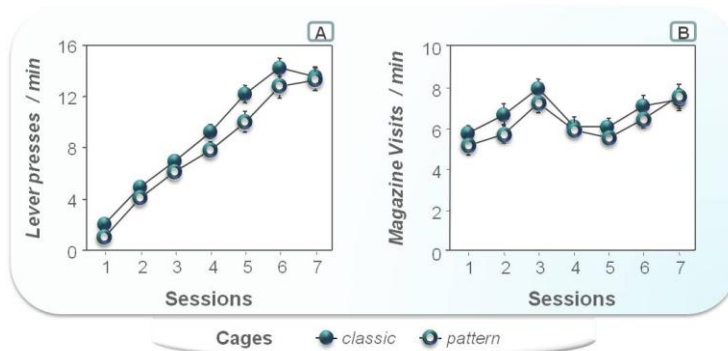


Figure II.4.3 Evolution des taux d'appuis (A) et de visites (B) au cours des deux conditionnements instrumentaux ($\pm$ erreur standard). Les sujets se comportent de façon semblable dans les deux contextes. Ils y acquièrent les réponses instrumentales à des vitesses et niveaux semblables.

On peut suivre sur le panneau A, les comportements d'appui de l'ensemble des sujets dans chacun des contextes. Au fil du conditionnement, les sujets appuient de plus en plus sur les leviers aussi bien dans le contexte *classic* que dans le contexte *pattern*. Ils semblent néanmoins distinguer les deux contextes, en cela qu'ils appuient à des fréquences légèrement plus soutenues quand ils sont placés dans le contexte *classic*. Toutefois, les taux d'appuis exprimés dans les deux contextes se rejoignent lors de la 7ème et dernière journée de conditionnement.

- * Une ANOVA indique des effets significatifs des facteurs **cage** ($F_{1,36} = 14,828$; $p < 0,001$) et **séance** ($F_{6,216} = 124,272$; $p < 0,001$) ainsi qu'une interaction significative entre ces deux facteurs ($F_{6,216} = 2,772$; $p < 0,05$).
- * Un test SNK localise des différences significatives entre les deux contextes pour la première et la cinquième séances.

Comparés à la progression des taux d'appuis, le taux de visites aux mangeoires est relativement stable (panneau B). On retrouve des taux légèrement plus faibles dans le contexte *pattern*.

- * Une ANOVA indique un effet significatif du facteur **séance** ($F_{6,216} = 7,504$; $p < 0,001$) mais pas d'effet du facteur **cage** ($F_{1,36} = 2,801$; ns) ni d'interaction entre les deux facteurs ($F_{6,216} = 0,659$; ns).

A l'issue du conditionnement instrumental, les sujets ont été divisés en 2 groupes *Devalued* et *NonDevalued* en vue de la phase de dévaluation.

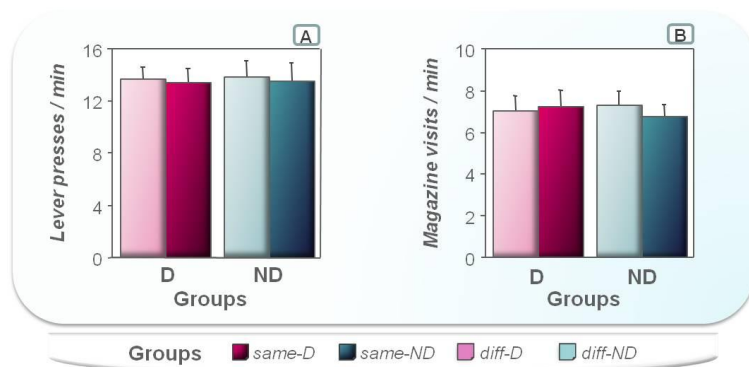


Figure II.4.4 Taux d'appuis (A) et de visites (B) référence des groupes constitués à la fin du conditionnement opérant. ($\pm$ erreur standard).

Les groupes constitués ne diffèrent pas entre eux à la fin de la phase d'entraînement.

Les panneaux A et B de la figure II.4.4 représentent les taux de référence d'appuis et de visites de ces deux groupes. A la fin de l'entraînement, les sujets des deux groupes se comportent dans le futur contexte (*same*) de la phase de dévaluation de façon similaire à l'autre contexte (*different*).

- * Concernant l'évolution du comportement d'appuis, une ANOVA confirme que seul le facteur **séance** a un effet significatif ($F_{6,210} = 120,97$; $p < 0,001$), que les facteurs **groupe** et **contexte** n'ont pas d'effet ($F_{1,35} = 0,217$; ns et $F_{1,35} = 1,521$; ns) et qu'il n'existe aucune interaction entre les 3 facteurs ($F_s < 1,6$; $p_s > 0,20$).
- * En complément, l'analyse statistique des taux de référence d'appuis et de visites rejette tout effet des facteurs **groupe** ($F_{1,35} = 0,009$; ns et $F_{1,35} = 0,010$; ns) et **contexte** ($F_{1,35} = 0,307$; ns et $F_{1,35} = 0,099$; ns) et l'existence d'une interaction entre ces facteurs ($F_{1,35} = 0,006$; ns et $F_{1,35} = 0,641$; ns).

• Dévaluation de la récompense dans un des deux contextes

On peut suivre l'évolution des indices moyens de consommation sur la figure II.4.5.

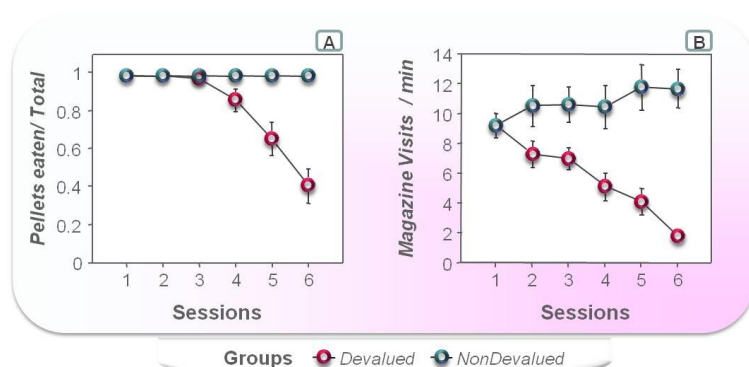


Figure II.4.5 Evolution des taux d'appuis (A) et de visites (B) au cours des deux conditionnements instrumentaux ($\pm$ erreur standard). L'indice de consommation du groupe de sujets recevant du LiCl évolue comme dans les expériences précédentes. Le niveau de consommation final atteint est moins faible que précédemment constaté, mais le comportement de visite indique que ces sujets ont bien acquis une aversion gustative conditionnée pour la récompense.

Les animaux du groupe *NonDevalued* consomment à chaque séance la totalité des pastilles de grains, alors que ceux du groupe *Devalued* diminuent leur consommation à partir de la 3ème séance. Ils atteignent un niveau final d'aversion gustative exprimé par la consommation d'environ 41% des récompenses distribuées. Ce niveau est relativement haut mais peut être

expliqué par le fait que lors de la phase 2, les sujets ont été deux fois plus exposés à la récompense par comparaison aux sujets des expériences précédentes.

- * Une ANOVA à 3 facteurs confirme les effets significatifs des facteurs **dévaluation** ($F_{1,33} = 28,209$; $p < 0,001$) et **séance** ($F_{5,165} = 28,806$; $p < 0,001$) ainsi qu'une interaction significative entre ces deux facteurs ($F_{5,165} = 28,806$; $p < 0,001$). Par contre, elle réfute l'existence d'un effet du facteur **cage** ($F_{1,33} = 1,475$; ns) et toute interaction entre ce facteur et les deux autres ($F_s < 1,5$; $p_s > 0,20$).

Concernant le comportement de visite à la mangeoire, les résultats obtenus sont conformes à ceux attendus (données non représentées). Si les sujets du groupe *NonDevalued* expriment une certaine stabilité comportementale, les sujets du groupe *Devalued* ont des taux de visites qui diminuent régulièrement au fil des séances de dévaluation.

- * Une ANOVA à 3 facteurs confirme les effets significatifs des facteurs **dévaluation** ($F_{1,33} = 13,065$; $p < 0,01$) et **séance** ($F_{5,165} = 4,349$; $p < 0,01$) ainsi qu'une interaction significative entre ces deux facteurs ($F_{5,165} = 15,792$; $p < 0,001$). Par contre, elle réfute l'existence d'un effet du facteur **cage** ($F_{1,33} = 0,213$; ns) et toute interaction entre ce facteur et les deux autres ($F_s < 1,7$; $p_s > 0,14$).

• Tests comportementaux

- Tests RdR

Les résultats obtenus pour les tests non récompensés sont reportés sur la figure II.4.6. Les performances de chaque animal sont rapportées à son taux de référence.

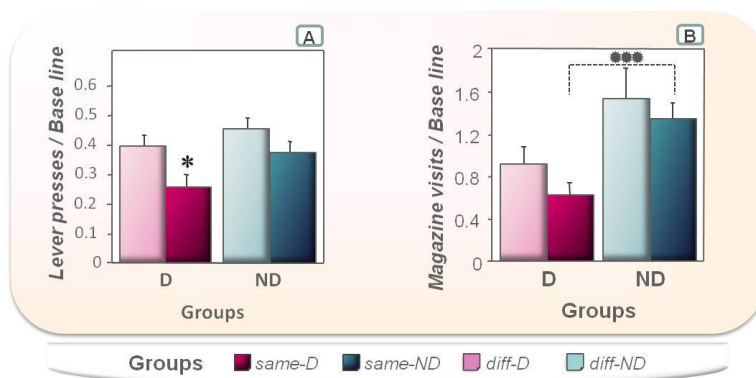


Figure II.4.6 Taux d'appuis (A) et de visites (B) au cours des tests RdR ($\pm$ erreur standard).

Les sujets pour lesquels la récompense est dévaluée se comportent différemment selon le contexte du test. On observe un effet de la dévaluation seulement dans le contexte où se sont tenues les séances de dévaluation. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Le panneau A est consacré au comportement d'appuis sur le levier. Concernant les sujets du groupe *NonDevalued*, ils ne possèdent pas d'aversion pour la récompense et appuient sur le levier à des niveaux élevés aussi bien dans le contexte de dévaluation que dans l'autre contexte. Les résultats sont différents pour les animaux possédant une aversion gustative pour la récompense. En effet, leur comportement d'appui est sensible à l'AGC dans les conditions

contextuelles *same*, mais n'y semble pas sensible quand le même rat est testé en conditions *different*. Autrement dit, les sujets du groupe *Devalued* se comportent à la fois comme dans les Expériences II.2 et II.3 : selon que le contexte de dévaluation est identique ou non au contexte de test opérant, les animaux expriment ou non une sensibilité à la dévaluation de la récompense.

- * L'analyse de variance indique des effets significatifs des facteurs **dévaluation** ($F_{1,33} = 5,02$; $p < 0,05$) et **contexte** ($F_{1,33} = 6,463$; $p < 0,05$). Il n'y a pas d'effet du facteur **cage** ($F_{1,33} = 0,020$; ns) et les 3 facteurs n'interagissent pas entre eux ($F_s < 1$; $p_s > 0,35$).
- * Le test SNK indique que le comportement d'appui du groupe *Devalued* diffère significativement selon le contexte du test. De plus seul le taux d'appuis moyen exprimé dans la condition *D-diff* diffère significativement des taux des trois autres conditions *D-same*, *ND-same* et *ND-diff*.

Le panneau B représente le comportement de visite à la mangeoire. Le patron des résultats obtenus est similaire à ceux décrit pour l'appui sur le levier. Le groupe *NonDevalued* exprime un niveau de visites élevé dans les deux contextes. Les animaux du groupe *Devalued* visitent la mangeoire à des niveaux plus faibles avec une activité qui semble plus haute dans les conditions *Different*.

- * Une ANOVA indique que seul le facteur **dévaluation** ($F_{1,33} = 10,109$; $p < 0,01$) agit de façon significative, les facteurs **contexte** ($F_{1,33} = 2,003$; ns) et **cage** ($F_{1,33} = 0,023$; ns) n'ayant pas d'effet. Il y a une tendance à une interaction **cage x contexte** ($F_{1,33} = 3,256$; $p = 0,0803$), il n'existe pas d'autre interaction ($F_s < 2$; $p_s > 0,15$).

- Tests FdR

On peut observer les résultats des tests FdR sur la figure II.4.7. Les performances de chaque animal sont rapportées à ses taux de référence.

Les résultats en terme de performance instrumentale sont représentés sur le panneau A. Comme précédemment, le groupe *NonDevalued* présente de hauts niveaux d'appuis dans les deux contextes. La distribution de la récompense modifie le comportement des sujets du groupe *Devalued*, ils diminuent nettement leur niveau d'appuis quel que soit le contexte dans lequel ils sont testés.

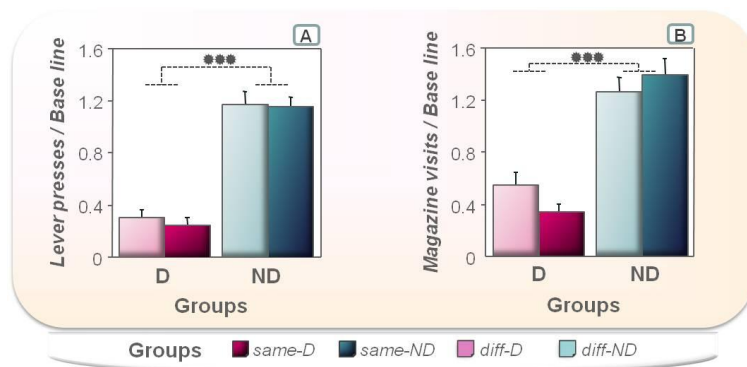


Figure II.4.7 Taux d'appuis (A) et de visites (B) au cours des tests FdR ($\pm$ erreur standard).

L'effet du contexte sur la performance instrumentale s'atténue lorsque l'appui est de nouveau récompensé, les taux d'appuis exprimés dans les deux contextes atteignent des niveaux comparables. * $p < 0.05$; ** $p < 0.01$; *** $p < 0,001$.

- * Une ANOVA confirme le seul effet significatif du facteur **dévaluation** ($F_{1,33} = 74,146$; $p < 0,001$) et l'absence d'effet des facteurs **cage** ($F_{1,33} = 1,056$; ns) et **contexte** ($F_{1,33} = 0,585$; ns) ou de toute interactions entre les 3 facteurs ($F_s < 0,6$; $p_s > 0,30$).
- * Un test SNK confirme l'absence de différence intragroupe.

Le panneau B représente les résultats obtenus pour le comportement de visite. Ils sont conformes aux résultats instrumentaux. Quel que soit le contexte du test, les animaux du groupe *NonDevalued* visitent fréquemment la mangeoire tandis que les niveaux de visites du groupe *Devalued* sont bas. La diminution des visites est un peu moins marquée pour le groupe *Devalued* lorsque ses sujets sont testés dans les conditions *different*.

- * Une ANOVA confirme l'unique effet du facteur **dévaluation** ($F_{1,33} = 46,290$; $p < 0,001$). Elle révèle une tendance à l'interaction entre les facteurs **dévaluation** et **contexte** ($F_{1,33} = 4,071$; $p = 0,0518$), mais pas d'effet propre du facteur **contexte** ($F_{1,33} = 0,295$; ns), ni du facteur **cage** ($F_{1,33} = 0,739$; ns), ni toute autres interactions entre les 3 facteurs ($F_s < 0,9$; $p_s > 0,35$).
- * Un test SNK indique d'une différence significative entre les conditions *same* et *different* pour le groupe *Devalued*.

Le nombre moyen de récompenses obtenues lors du test FdR et les indices de consommation moyen sont représentés respectivement sur les panneaux A et B de la figure II.4.8.

Les animaux du groupe *NonDevalued* obtiennent les 40 récompenses maximales dans les deux contextes et ne possédant pas d'AGC ils en consomment la totalité. Si les sujets du groupe *Devalued* semblent en moyenne se distribuer plus de récompenses dans la condition *different* (33 ± 2) que dans la condition *same* (27 ± 3), ils expriment néanmoins dans les deux contextes des niveaux de consommation comparables aux environs de 58% des pastilles. Le niveaux d'aversion exprimé ici est plus faible que dans les expériences précédentes, mais concorde avec l'indice de consommation observé lors de la 6ème séance de dévaluation.

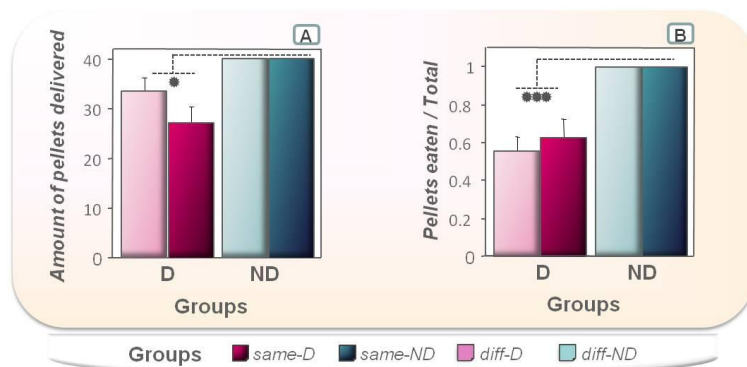


Figure II.4.8 Quantité de récompense obtenue (A) et consommée (B) au cours des tests RdR ($\pm$ erreur standard).

La quantité de pastilles distribuées diffère selon le contexte où a eu lieu le test récompensé. Elle est plus faible quand il correspond au lieu des séances de dévaluation. En revanche, les sujets consomment la récompense à des niveaux comparables quel que soit le contexte.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

- * Concernant le nombre de récompenses distribuées, l'analyse de variance indique des effets significatifs des facteurs **dévaluation** ($F_{1,33} = 17,136$; $p < 0,001$) et **contexte** ($F_{1,33} = 5,030$; $p < 0,05$) mais pas d'effet du facteur **cage** ($F_{1,33} = 0,069$; ns). Si le facteur **contexte** interagit significativement avec les deux autres ($F_s > 5$; $p_s < 0,05$), il n'y a pas d'interaction entre les facteurs **dévaluation** et **cage** ($F_{1,33} = 0,069$; ns).
 - * Un test SNK montre des différences entre groupes *Devalued* et *NonDevalued*.
 - * Concernant les indices de consommation, une ANOVA confirme le seul effet significatif du facteur **dévaluation** ($F_{1,33} = 37,910$; $p < 0,001$) et l'absence d'effet des facteurs **contexte** ($F_{1,33} = 0,444$; ns) et **cage** ($F_{1,33} = 0,009$; ns) ou de toute interactions entre les 3 facteurs ($F_s < 0,7$; $p_s > 0,45$).
- Un test SNK montre des différences entre groupes *Devalued* et *NonDevalued*.

Discussion

Lors de la phase d'acquisition, la légère différence entre les taux d'appuis exprimés dans les deux contextes n'a pas d'incidence pour les suites de l'expérimentation, les groupes sont constitués de façon équilibrée. Néanmoins, cette différence permet d'affirmer que nos sujets sont capables de différencier les deux contextes.

Les résultats de la phase de test concordent avec ceux des Expériences II.2 et II.3. Nous retrouvons en effet les différences de comportement non pas entre des groupes mais pour un même individu selon les paramètres de la séance de test.

Ainsi, il apparaît qu'un individu possédant une AGC pour la récompense peut avoir deux comportements divergents : il sera sensible à la dévaluation dans le contexte où s'est tenue l'AGC, au contraire son comportement d'appui sera augmenté dans un contexte différent.

Chapitre II : Discussion

Les expériences, purement comportementales, décrites dans ce chapitre visaient à contribuer au courant actuel qui considère que la réponse opérante résulte de l'encodage, lors de l'entraînement, d'un lien entre l'action et le résultat (Rescorla, 1991 ; Balleine et Dickinson, 1998a). Les résultats de ces expériences confirment cette assertion, ils présentent d'importantes implications qui sont discutées ci-dessous.

Les résultats de l'Expérience I.1 démontrent la spécificité de l'encodage en cela que les animaux changent uniquement la réponse correspondant à une représentation d'un résultat dont la valeur a été modifiée, conformément aux propositions théoriques liées à cette question (Rescorla, 1991 ; Balleine et O'Doherty, 2010). Des travaux plus anciens de la littérature ont utilisé des procédures discriminatives comparables à la nôtre (ex, Adams et Dickinson, 1981). En plus de la spécificité de l'encodage, ce type de procédure permet également de s'assurer de la spécificité de la réponse. En effet, dans notre procédure, les réponses d'appui sont en compétition avec les réponses de visite. Ainsi, l'utilisation de deux sources de récompense nous a permis de vérifier que la diminution de réponse opérante ne résulte pas d'une augmentation parallèle de l'activité de visite.

Dans notre procédure, la dévaluation s'effectue dans les cages de conditionnement, comme c'est le cas dans un nombre important d'études (Adams et Dickinson, 1981 ; Adams, 1982, Colwill et Rescorla, 1985 ; Colwill, 2006 ; Nelson et Killcross, 2006). Une interprétation simpliste de nos résultats pourrait consister à penser que la diminution de réponse résulte d'une aversion pour le contexte acquise lors de la phase de dévaluation (Melcer et Timberlake, 1985 ; Boakes et al., 1997 ; Rosas et Callejas-Aguilera, 2007). Les données recueillies suggèrent qu'une telle interprétation n'est pas valide, puisque les animaux des conditions contrôles expriment des comportements incompatibles avec une aversion acquise pour le contexte, en effet ils visitent la mangeoire associée à la récompense non dévaluée à des taux comparables aux animaux *NonDevalued*.

Les résultats de l'Expérience II.2 démontrent que les informations relatives au contexte de dévaluation sont intégrées dans la représentation. En effet, lorsque la dévaluation a lieu dans un contexte alternatif, les animaux n'ajustent plus leur réponse aux changements de valeur. Ce résultat est conforme aux données anciennes (Holman, 1975 ; Wilson et al., 1981) mais aussi plus récentes (Yin, 2004 ; Jonkman, 2010). Toutefois, cet effet n'avait pas fait l'objet jusqu'ici d'une investigation systématique.

Cette description des résultats demandait d'être approfondie quant aux mécanismes psychologiques qui sont à l'œuvre, et nous avons testé plusieurs hypothèses (voir introduction).

La première consiste à penser que le contexte joue un rôle déterminant dans l'apprentissage incitatif, c'est-à-dire dans le mécanisme par lequel la valeur de la représentation est mise à jour par contact direct avec la récompense. Les données de l'Expérience II.3 ne vont pas dans ce sens puisqu'ils indiquent que la réexposition est sans effet.

Bien que tout à fait clairs, ces résultats mériteraient d'être complétés par d'autres données. En effet, notre procédure de réexposition présente deux limitations. La première correspond à la méthode de réexposition puisque les animaux doivent eux-mêmes effectuer le comportement pour être exposés à la récompense. La seconde concerne la quantité d'exposition, relativement faible (Balleine et Dickinson, 1991 ; Lopez et al., 1992). Des procédures d'ingestion forcée (Fouquet et al., 2001) permettraient de contourner ces problèmes et d'accumuler d'importantes données sur ce thème.

La seconde hypothèse que nous avons testée correspond au rôle du contexte dans la prise en compte de la représentation (voir également Jonkman et al., 2010) et les résultats de l'Expérience II.4 valident cette hypothèse. Ce rôle peut être de deux types. En effet, le contexte pourrait activer une représentation du résultat, et c'est cette représentation qui serait utilisée pour la sélection et l'activation du système de réponse (voir DeWit et Dickinson, 2009 ; Holmes et al., 2010). De manière alternative, le contexte pourrait avoir un rôle de stimulus discriminatif d'ordre supérieur permettant la sélection de la réponse adaptée (Rescorla, 1991 ; Haddon et al., 2008).

Nos données favorisent plutôt la seconde option. En effet, les résultats des Expériences II.1 et II.2 suggèrent que l'acquisition de la contingence Cx-Résultat ne joue pas ou peu de rôle dans le contrôle de la réponse.

Par contre, l'observation selon laquelle le contexte permet à l'animal d'adapter sa réponse suggère que l'encodage est enrichi de deux dimensions. La première concerne l'encodage du contexte dans la représentation. La seconde, plus inattendue, suggère que la valeur du résultat de l'action peut être encodée simultanément sous deux formes distinctes : une forme où la valeur du résultat est positive (comme lors de l'entraînement), et une forme où la valeur est

ajustée après la dévaluation. Nos résultats indiquent en outre que les informations contextuelles servent à sélectionner la représentation appropriée.

Ces résultats originaux ont d'importantes implications aux plans méthodologiques et théoriques. En effet, un examen approfondi de la littérature indique que la majorité des études portant sur les circuits neuronaux des habitudes mettent en œuvre des procédures proches de celles de notre Expérience II.2 (voir par exemple Yin et al., 2004), c'est-à-dire que la phase de dévaluation est effectuée dans un environnement différent de celui de l'entraînement (en général la cage d'habitation des animaux). A la lumière de nos propres résultats, il apparaît que le système de contrôle en jeu dans ces études n'est pas, à proprement parler, sous tendu par le système habitude. Au contraire, puisque nos données suggèrent qu'il existe deux représentations parallèles de la valeur du résultat de l'action, il est tentant de proposer que le système de contrôle de la réponse privilégié dans ces études est celui qui utilise une représentation de la valeur de la récompense non dévaluée. L'utilisation privilégiée de ce système pourrait être favorisée, nous venons de le voir lorsque la dévaluation est effectuée dans un autre environnement, mais pourrait également être privilégiée dans le cas où un même résultat est obtenu par entraînement intensif dans un contexte ou par entraînement modéré dans un autre contexte. C'est précisément l'observation qui a été faite dans le passé (voir Killcross et Coutureau, 2003).

Ces expériences font l'objet d'une publication, en cours de rédaction.

Chapitre III

Chapitre III

Introduction

Sur le plan anatomique, diverses approches indiquent un rôle du cortex préfrontal médian (mPFC) dans l'action dirigée vers un but (Dalley et al., 2004; Matsumoto et Tanaka, 2004). Des études électrophysiologiques chez le Primate (Matsumoto et al., 2003) et le Rat (Mulder et al., 2003) montrent notamment des activités dans le mPFC en lien avec des associations spécifiques entre une action et ses conséquences (Tanaka et al., 2008). Le cortex préfrontal médian, et particulièrement sa région prélimbique (PL) semble situé au sein d'un réseau incluant l'hippocampe, le cortex insulaire, le noyau accumbens, l'amygdale basolatérale, certains noyaux thalamiques et l'aire tegmentale ventrale, et il a été proposé qu'il intègre des informations passées et présentes relatives à l'action dirigée vers un but, et notamment leurs aspects affectifs (Vertes 2006).

Le PL semble indispensable à une organisation normale de l'action dirigée puisque des lésions sélectives du PL effectuées avant l'apprentissage abolissent la sensibilité de l'action instrumentale vis-à-vis d'une dévaluation de la récompense, ou encore d'une rupture du lien causal avec celle-ci (Balleine et Dickinson, 1998a; Corbit et Balleine, 2003; Killcross et Coutureau, 2003 ; Coutureau et al., 2009). Il est important toutefois de noter que ces lésions, ou plus généralement des lésions du mPFC, ne semblent pas empêcher l'acquisition des actions instrumentales et préservent certains indicateurs de l'action dirigée vers un but comme la facilitation sélective de l'action par la présentation de la récompense ou de stimuli qui y sont associés (Corbit et Balleine, 2003).

Ceci a amené certains auteurs à proposer qu'en l'absence du mPFC la performance instrumentale serait avant tout contrôlée par des stimuli extérieurs (Balleine et Dickinson, 1998a ; Corbit et Balleine, 2003). Sur le plan théorique, cette possibilité d'acquérir une réponse instrumentale uniquement sur la base d'associations S-R va à l'encontre de la conception classique selon laquelle les habitudes ne pourraient contrôler l'action qu'après un apprentissage prolongé (Dickinson, 1985). Une conception plus appropriée pourrait être de considérer que deux systèmes fonctionnent en parallèle dès le début de l'apprentissage, en compétition l'un avec l'autre (Balleine, 2005 ; Daw et al. 2005). L'un d'entre eux (A-O) serait caractéristique de l'action dirigée et sensible à la valeur de la récompense et à la contingence, l'autre (S-R) serait caractéristique d'actions automatiques.

Toutefois, le rôle du PL demande encore à être clarifié, notamment en ce qui concerne son intervention tout au long de l'apprentissage ou du test. En effet, on attribue souvent au cortex préfrontal un rôle dans le contrôle « en ligne » du comportement (Miller et Cohen, 2001; Dalley et al., 2004). La plupart des études chez le rat mettent en jeu des lésions effectuées avant l'apprentissage, ce qui ne permet pas de déterminer si elles sont spécifiquement impliquées dans l'encodage, le stockage ou le rappel des associations A-O. Une étude d'Ostlund et Balleine (2005) indique toutefois que des lésions du mPFC effectuées après l'apprentissage sont sans effet sur la sensibilité de la réponse instrumentale à la dévaluation. Cette étude demande néanmoins à être complétée pour deux raisons : d'une part, le mPFC est une région hétérogène (voir la revue de Heidbreder et Groenewegen, 2003) puisqu'elle inclut (dans le sens dorso-ventral) le cortex cingulaire, le cortex prélimbique et le cortex infralimbique, ces deux derniers étant touchés par la lésion dans l'étude d'Ostlund et Balleine (2005). Or, la région infralimbique joue également un rôle important dans le contrôle de l'action instrumentale puisque sa lésion ou son inactivation est capable de maintenir ou de promouvoir un fonctionnement de type A-O (Killcross et Coutureau, 2003; Coutureau et Killcross, 2003). D'autre part, des lésions effectuées avant l'apprentissage, si elles éliminent bien la région d'intérêt au moment de l'acquisition, ne permettent pas de tester les animaux avec un fonctionnement cérébral intact. Il importait donc de compléter ces expériences par des inactivations sélectives et réversibles lors de phases distinctes de l'expérience. C'est l'objectif des expériences présentées dans ce chapitre.

Expérience III.1

Implication du cortex prélimbique dans l'apprentissage d'une action dirigée

Cette première expérience vise à déterminer l'implication du cortex prélimbique dans l'apprentissage d'une tâche instrumentale et plus spécifiquement dans l'acquisition du caractère dirigé de l'action. Afin d'obtenir une plus grande spécificité temporelle que celle fournie par les lésions excitotoxiques, nous utilisons ici des micro-injections de muscimol, agoniste des récepteurs GABA_A qui a pour effet d'inactiver le cortex de façon réversible et transitoire (Majchrzak et Di Scala, 2000). Ainsi, nos sujets disposeront d'un cortex prélimbique fonctionnel pendant la phase finale de test.

A. Matériel et méthode

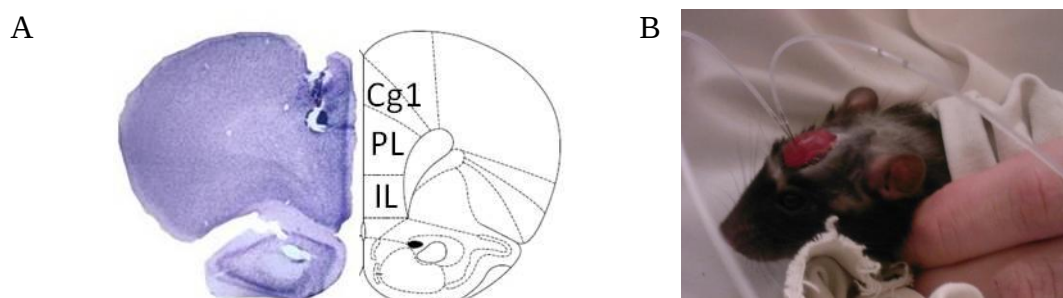
a. Sujets

Les sujets de cette expérience sont 48 rats mâles naïfs de la souche Long-Evans (310 à 400 g). Les conditions d'hébergement et de restriction alimentaire sont identiques à celles décrites dans l'Expérience I.1.

b. Chirurgie

Le rat est prémédiqué par du diazépam (Valium®, Roche ; 5 mg/kg) et de la xylazine (Rompan®, Bayer ; 10 mg/kg) en injections intra-péritonéales, avant d'être anesthésié par injection intra-musculaire de kétamine (Virbac ; 90mg/kg). L'animal a ensuite le crâne rasé puis est fixé sur un cadre stéréotaxique (Kopf instruments, Tujunga, Californie). Après une anesthésie locale par lidocaïne (Xylocaïne®, AstraZeneca) et une incision cutanée, l'os crânien est mis en évidence et nettoyé. Les coordonnées du Bregma et du Lambda sont relevées et la position du crâne est ajustée à l'horizontale. Les coordonnées d'implantation dans le cortex prélimbique des 2 guide-canules (longueur : 8 mm ; diamètre interne : 0,39 mm ; LeGuellec, France) sont calculées par rapport au Bregma. A l'aide d'une fraise dentaire, des orifices sont percés au dessus des régions à atteindre et les guide-canules sont implantés aux coordonnées suivantes : A-P +3,2 mm (antéro-postérieur), L $\pm$ 0,6mm (latéral), V -2,5 mm (ventral) (cf. figure III.1., panneau A). Des mandrins amovibles sont placés à l'intérieur

des guides afin d'éviter qu'ils ne se bouchent. L'ensemble est maintenu en place par un bloc de ciment dentaire (Palavit®G) renforcé à sa base par trois vis placées dans le crâne. A la suite de l'opération, l'animal est placé dans une salle de réveil chauffée. Le lendemain, il réintègre l'animalerie pour une phase de récupération postopératoire de 7 jours pendant laquelle il a accès *ad libitum* à l'eau et à son régime alimentaire habituel.



Localisation de la micro-injection (A) et photographie d'un rat maintenu en contention lors d'une micro-injection (B).

Microphotographie d'une coupe frontale du cerveau d'un rat implanté illustrant la localisation du site d'injection au sein du cortex prélimbique (à gauche) et la trace du guide-canule. Distance au Bregma A-P +3.2 mm. Schéma de droite adapté de Paxinos et Watson (1998).

c. Dispositif expérimental

Les cages de conditionnement sont celles précédemment décrites dans l'Expérience I.1. Elles sont donc équipées d'une mangeoire et d'un levier. La récompense alimentaire est une pastille de grains (45 mg). Les micro-injections intracérébrales ont lieu dans une salle adjacente à la salle de conditionnement instrumental. Le dispositif comprend un pousse-seringue double (Syringe infusion pump 2.2, Harvard Apparatus) sur lequel sont posées des seringues Hamilton reliées via des cathéters à des canules de micro-injection (longueur : 9,5 mm ; diamètre interne : 0,23 mm ; LeGuellec, France) qui seront insérées dans les guides implantés dans le crâne du sujet (figure III.1.1, panneau B).

d. Protocole expérimental

Ce protocole reprend les quatre phases décrites pour l'Expérience I.1. Seul le déroulement de la phase de conditionnement instrumental se trouve modifié (Figure III.1.2).



Figure III.1.2 Représentation schématique du déroulement de l'expérience.

Les éclairs oranges représentent les micro-injections cérébrales.

Mag : séance d'entraînement à la mangeoire ; FR1 et VI30 : séances de conditionnement instrumental ;

D : séance de dévaluation par AGC ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; FdR : test pendant lequel la réponse est renforcée.

- *Phase 1 : Entraînement à la mangeoire et préparation aux micro-injections*

Les deux séances d'entraînement à la mangeoire (cf. Expérience I.1) sont suivies chacune par une séance de préparation aux micro-injections. Les rats sont amenés dans la nouvelle pièce et sont habitués à être maintenus en contention dans une pièce de tissu. Le 2^{ème} jour, les mandrins sont retirés, les guide-canules sont nettoyés et débouchés à l'aide de tire-nerfs dentaires de 8 mm de longueur (Micro-Mega, France). De nouveaux mandrins sont alors mis en place.

- *Phase 2 : Inactivation du PL et conditionnement instrumental*

L'entraînement instrumental comporte 2 séances en FR1 et 4 séances en VI30 (cf. Expérience I.1). Sur toute la durée de cette phase, les sujets sont soumis quotidiennement à une micro-injection intracérébrale selon le protocole suivant.

Avant chaque séance instrumentale, le sujet est amené dans la salle de micro-injection. Il est maintenu immobile dans le tissu par l'expérimentateur et une fois les mandrins ôtés, les guides sont de nouveau nettoyés. Les canules de micro-injection sont insérées dans les guides. Une solution est alors injectée à raison de 0,2 $\mu\text{L}/\text{min}$ pendant une durée de 1 min 15. Les canules sont maintenues en place durant 2 min après la fin de l'injection (cf figure III.1.1, panneau B), puis l'animal est remis dans sa cage d'habitation pour une durée de 5 min. Enfin, chaque animal est placé dans sa cage de conditionnement pour effectuer sa séance d'entraînement instrumental.

Pour la moitié des animaux, constituant le groupe *Muscimol*, la solution injectée est du muscimol à une concentration de 1 $\mu\text{g}/\mu\text{L}$ dans du liquide cébrospinal artificiel (muscimol hydrobromide, Sigma-Aldrich ; CMA Microdialysis, USA). L'autre moitié des sujets constitue le groupe *Vehicle*. Ils reçoivent du liquide cébrospinal artificiel seul.

- *Phase 3 : Dévaluation de la récompense par aversion gustative*

A l'issue de la phase d'entraînement instrumental, la dévaluation de la récompense par aversion gustative conditionnée se déroule de façon analogue à l'Expérience I.1. Nous subdivisons alors chaque groupe de micro-injection en deux sous-groupes selon la nature des injections (LiCl, NaCl) auxquelles ils sont soumis.

Nous obtenons ainsi à l'issue de cette phase quatre groupes de sujets :

- ***Vehicle-Devalued (V-D)*** : le PL n'a pas été inactivé pendant l'acquisition instrumentale et la récompense est dévaluée ;
- ***Vehicle-NonDevalued (V-ND)*** : le PL n'a pas été inactivé pendant l'acquisition instrumentale et la récompense n'est pas dévaluée ;
- ***Muscimol-Devalued (M-D)*** : le PL a été inactivé pendant l'acquisition instrumentale et la récompense est dévaluée ;
- ***Muscimol-NonDevalued (M-ND)*** : le PL a été inactivé pendant l'acquisition instrumentale et la récompense n'est pas dévaluée.

- *Phase 4 : Tests comportementaux*

Le lendemain de la dernière séance de dévaluation, les sujets sont soumis à un test RdR d'une durée de 15 minutes, suivi 1 heure après par un Test FdR (cf. Expérience I.1).

e. Histologie

Les sujets reçoivent une dose létale de Pentobarbital sodique (Ceva Santé Animal ; 1 mL/300 g). A l'aide d'un cathéter intra-cardiaque, les animaux sont perfusés avec 60 mL de NaCl (0,9%) suivis de 120 mL d'une solution de formaldéhyde (10%). Ce procédé permet de fixer les tissus cérébraux. Une fois le cerveau extrait de la boîte crânienne, il est mis sous agitation lente pendant 72 heures dans une solution de formaldéhyde (10%) et de saccharose (30%). A l'aide d'un microtome à congélation (Leica SM 2400), les cerveaux sont sectionnés en coupes de 60 µm d'épaisseur qui sont montées sur des lames gélatinées. Après 48 heures de séchage, les coupes sur lames subissent une délipidation, une coloration des corps cellulaires à la thionine et enfin une déshydratation. Pour finir, les coupes incluses dans de la résine (Eukitt®) sont montées sous lamelles. L'examen des sites d'injection au microscope optique à fond clair permet de reconstituer la position des sites sur un schéma adapté de l'Atlas de

Paxinos et Watson (1998). (Examen effectué par D.T.T.Y et E.C aveugles pour les conditions.)

f. Analyses statistiques

Les résultats obtenus (*taux d'appuis, taux de visites, indices de consommation*) sont analysés à l'aide des variables indépendantes suivantes :

- le facteur **séance** (1- 6) (cf. Expérience I.1) ;
- le facteur **groupe** (cf. Expérience II.1) ;
- le facteur **dévaluation** (*NonDevalued/Devalued*) (cf. Expérience I.1) ;
- le facteur **traitement** (*Vehicle/Muscimol*), selon la nature de l'injection dans le PL.

B. Résultats

• *Histologie*

Tous les sujets ont correctement récupéré des suites de l'implantation des guide-canules et ont été soumis à la procédure comportementale. Cinq animaux ont dû être écartés au cours de l'expérience, leurs canules d'injection s'étant bouchées. Après examen des coupes histologiques, six animaux présentant des sites d'injection trop ventraux ou situés en dehors du cortex prélimbique ont été exclus de l'étude. Les sujets retenus présentent en majorité des sites situés juste en dessous du cortex cingulaire, favorisant ainsi une diffusion dans le cortex prélimbique (figure III.1.3).

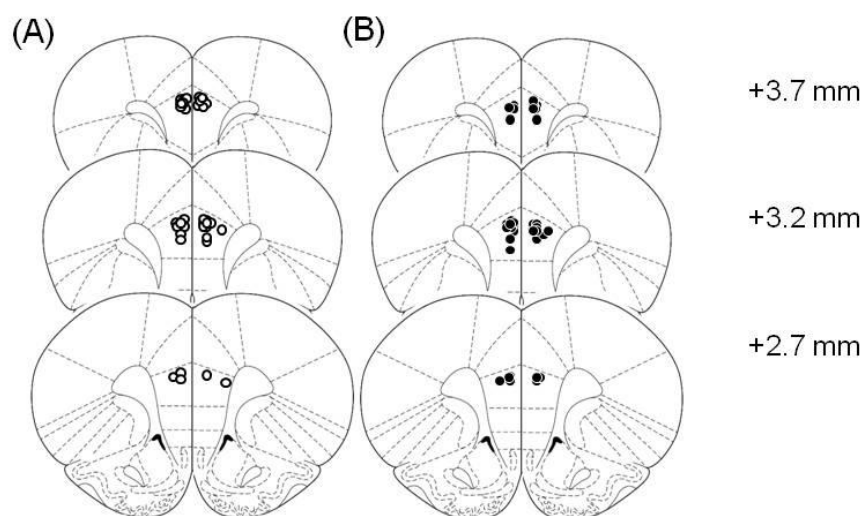


Figure III.1.3 Représentation schématique des sites d'injection dans le cortex prélimbique.

Disques blancs : groupes *Vehicle*. Disques pleins : groupes *Muscimol*. Les schémas sont adaptés à partir de l'atlas de Paxinos et Watson (1998). Sections disposées en fonction de leur position plus ou moins antérieure au Bregma.

Les effectifs finaux pour chaque groupe sont les suivants : *M-D*, n=9 ; *M-ND*, n=9 ; *V-D*, n=8 ; *V-ND*, n=11.

- *Entraînement à la mangeoire*

Tous les sujets s'habituent au dispositif expérimental et visitent la mangeoire afin de consommer les pastilles de grains (données non représentées).

- *Inactivation du PL et conditionnement instrumental*

Tous les sujets soumis aux micro-injections acquièrent sans difficulté la tâche instrumentale. Les résultats de la phase de conditionnement opérant sont représentés sur la figure III.1.4.

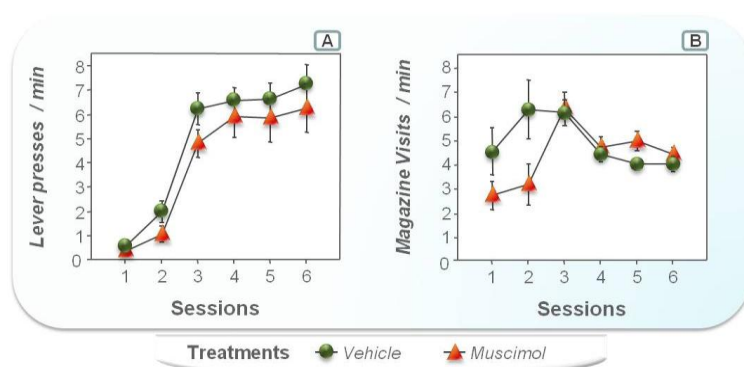


Figure III.1.4 Evolution des taux d'appuis (A) et de visites (B) au cours du conditionnement instrumental ($\pm$ erreur standard).

L'injection de muscimol dans le cortex prélimbique juste avant chaque séance d'entraînement n'empêche pas l'acquisition de la réponse instrumentale. Les taux d'appuis et de visites sont comparables entre les deux groupes.

Le panneau A représente l'évolution des taux d'appuis au fil de l'entraînement instrumental. Si les niveaux d'appuis des animaux *Muscimol* semblent légèrement plus faibles que ceux des sujets *Vehicle*, cette différence reste minime et les deux groupes acquièrent la tâche à la même vitesse. En revanche, nous remarquons que tous les sujets de cette expérience présentent des taux d'appuis diminués de moitié en comparaison des taux précédemment observés chez des animaux non implantés et non injectés. En effet, ils se situent aux environs de 6 appuis/min, alors que celui des sujets de l'Expérience I.1 avoisinait la valeur de 13. Néanmoins, l'ensemble des animaux micro-injectés acquiert correctement la réponse instrumentale : la performance d'appuis sur le levier augmente progressivement au fil des séances et le critère maximum des 40 récompenses distribuées est atteint à chaque séance.

* Une ANOVA à 2 facteurs confirme l'effet significatif du facteur **séance** ($F_{5,175} = 57,279$; $p < 0,001$), l'absence d'effet du **traitement** ($F_{1,35} = 1,628$; ns) et l'absence d'interaction entre ces deux facteurs ($F_{5,175} = 0,310$; ns).

Le panneau B représente l'évolution des taux de visites au fil des séances d'entraînement. Lors des deux premières séances, les animaux du groupe *Muscimol* visitent moins fréquemment la mangeoire que les animaux ayant reçu du solvant. Toutefois, cette différence de comportement entre les deux groupes tend à s'annuler à partir de la 3^{ème} séance. Nous constatons aussi une baisse d'activité chez l'ensemble des sujets micro-injectés. Les taux de référence se situent aux environs de 4 visites/min, alors que celui des sujets de l'Expérience I.1 avoisinait la valeur de 8.

- * Une ANOVA à 2 facteurs confirme l'effet significatif du facteur **séance** ($F_{5,175} = 4,781$; $p < 0,001$), l'absence d'effet du **traitement** ($F_{1,35} = 0,854$; ns) mais l'existence d'une interaction **séance x traitement** significative ($F_{5,175} = 3,980$; $p < 0,01$).

A l'issue de cette phase, les deux groupes de micro-injection sont subdivisés pour former 4 groupes équilibrés (données non représentées).

- * L'ANOVA confirme que les quatre groupes de dévaluation ont été constitués de façon équilibrée. Le facteur **groupe** est sans effet sur le comportement d'appui ($F_{3,33} = 0,579$; ns) ou de visite ($F_{3,33} = 2,275$; ns).

- *Dévaluation de la récompense par aversion gustative*

On peut suivre sur la figure III.1.5 l'évolution des indices de consommation au fil des séances de dévaluation.

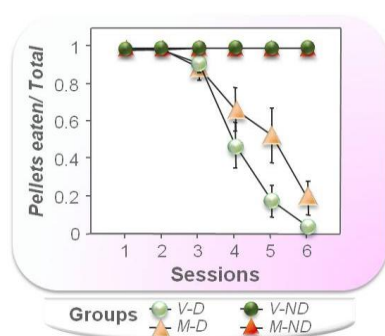


Figure III.1.5 Evolution des indices de consommation au cours des séances de dévaluation ($\pm$ erreur standard).

Le fait d'avoir reçu du muscimol pendant la phase d'entraînement, n'interfère pas avec l'acquisition de l'aversion gustative conditionnée. Les animaux des groupes *NonDevalued* consomment l'intégralité de la récompense à chaque séance. Les animaux des groupes *Devalued* diminuent graduellement leur consommation au fil des séances.

Conformément aux expériences précédentes, les animaux des groupes *NonDevalued*, *V-ND* et *M-ND*, consomment l'intégralité de la nourriture proposée (indice égal à 1). Au contraire, la consommation des animaux des groupes *Devalued*, *V-D* et *M-D*, diminue progressivement à partir de la 3^{ème} séance. Si cette diminution semble plus marquée chez les animaux traités avec du solvant, les deux indices finaux obtenus se situent dans la gamme des valeurs efficaces rapportées au cours des précédents chapitres.

- * Une ANOVA à 3 facteurs (**séance**, **dévaluation**, **traitement**) confirme les effets significatifs des facteurs **séance** et **dévaluation** ($F_{5,165} = 73,345$; $p < 0,001$ et $F_{1,33} = 99,516$; $p < 0,001$) ainsi qu'une interaction significative **séance x dévaluation** ($F_{5,165} = 73,679$; $p < 0,001$). Concernant la micro-injection cérébrale, il n'y a pas d'effet du facteur **traitement** ($F_{1,33} = 2,669$, ns). Toutefois, l'analyse statistique indique une interaction significative **séance x traitement** ($F_{5,165} = 2,997$; $p < 0,05$) et une triple interaction **séance x dévaluation x traitement** ($F_{5,165} = 3,049$; $p < 0,05$).
- * Un test SNK confirme une différence significative entre les indices de consommation finaux des groupes *M-D* et *V-D*.

Les résultats obtenus pour le comportement de visite à la mangeoire sont conformes à ceux attendus (données non représentées). Quel que soit le type d'injection reçue, les sujets des groupes *NonDevalued* ont un comportement relativement stable tout au long de la phase de dévaluation ; les sujets des groupes *Devalued* diminuent leur fréquence de visites dès la 2^{ème} séance.

- * Une ANOVA confirme les effets significatifs des facteurs **séance** et **dévaluation** ($F_{5,165} = 5,849$; $p < 0,001$ et $F_{1,33} = 22,816$; $p < 0,001$) ainsi qu'une interaction significative **séance x dévaluation** ($F_{5,165} = 4,558$; $p < 0,001$). Concernant la micro-injection cérébrale, il n'y a pas d'effet du facteur **traitement** ($F_{1,33} = 0,925$; ns) ni d'interaction entre ce facteur et les deux autres ($F_s < 2,997$; $p_s > 0,17$).
- * Un test SNK, réalisé sur les données de la dernière séance, confirme que les groupes *M-D* et *V-D* ne diffèrent pas entre eux mais diffèrent significativement des deux groupes *V-ND* et *M-ND*.

• Tests comportementaux

- Test RdR

Les résultats du test RdR sont représentés sur la figure III.1.6.

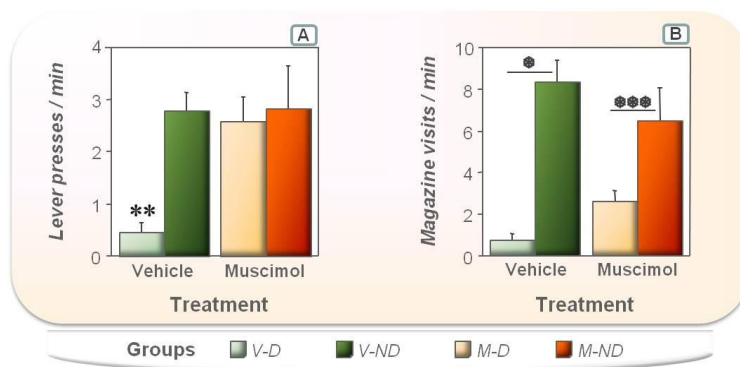


Figure III.1.6 Taux d'appuis (A) et de visites (B) au cours du test RdR ($\pm$ erreur standard).

Chez les animaux ayant reçu le traitement muscimol pendant la phase d'entraînement, on n'observe pas d'effet de la dévaluation sur la performance instrumentale. Ils se comportent vis-à-vis du levier comme des animaux n'ayant pas d'aversion pour la récompense. En revanche, ils visitent peu la mangeoire.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

On peut observer sur le panneau A, les résultats concernant les performances instrumentales. Lors de ce test non récompensé, les animaux des groupes *Vehicle* se comportent conformément aux résultats décrits pour des rats normaux. Les rats du groupe *V-D* appuient moins sur le levier que ceux du groupe *V-ND*, ils expriment alors une réponse instrumentale sensible à la dévaluation et donc dirigée vers un but. Les animaux des groupes *Muscimol* montrent un patron différent de résultats: les rats du groupe *M-D* appuient autant sur le levier que ceux du groupe *M-ND*. Les animaux dont le PL a été inactivé pendant toute la phase d'entraînement, expriment donc une réponse instrumentale insensible à la dévaluation de la récompense.

- * Une ANOVA à 2 facteurs révèle des effets significatifs des facteurs **traitement** ($F_{1,33} = 4,382$; $p < 0,05$) et **dévaluation** ($F_{1,33} = 6,09$; $p < 0,05$) et une tendance à une interaction **traitement x dévaluation** ($F_{1,33} = 4,382$; $p = 0,0554$).
- * Une analyse par un test SNK ne montre pas de différence entre les groupes *M-D* et *M-ND*. Ce test indique également que les performances instrumentales des groupes *V-ND*, *M-ND* et *M-D* sont significativement supérieures à celle du groupe *V-D*.

Concernant le comportement de visite à la mangeoire, il n'y a pas de différence notable entre les groupes *Muscimol* et *Vehicle* (Panneau B). En effet, les animaux des groupes *Devalued* visitent peu la mangeoire en comparaison aux animaux des groupes *NonDevalued*.

- * Une ANOVA à 2 facteurs confirme un effet significatif de la **dévaluation** ($F_{1,33} = 28,791$; $p < 0,001$), l'absence d'effet du **traitement** ($F_{1,33} < 0,01$; ns) et l'absence d'une interaction **traitement x dévaluation** ($F_{1,33} = 2,960$; ns).
- * Une analyse par un test SNK confirme que les groupes *M-D* et *V-D* ne diffèrent pas entre eux mais diffèrent significativement des deux groupes *V-ND* et *M-ND*.

- **Test FdR**

Les résultats concernant les comportements d'appui et de visite exprimés lors du test FdR sont représentés sur la figure III.1.7.

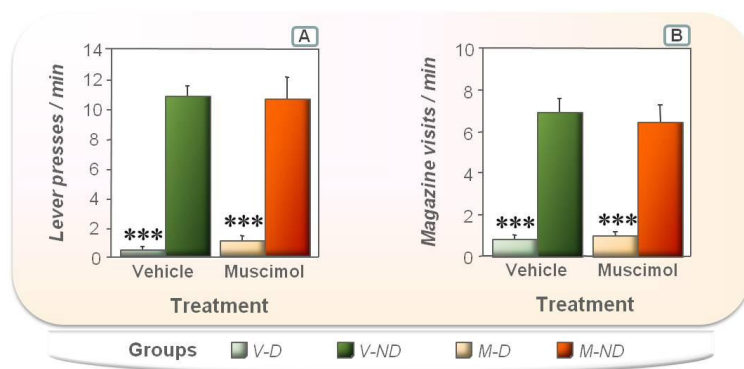


Figure III.1.7 Taux d'appuis (A) et de visites (B) au cours du test FdR ($\pm$ erreur standard).

Quand l'appui est de nouveau récompensé les sujets du groupe *Muscimol-Devalued* adaptent leur comportement et cessent d'appuyer sur le levier. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Lors de ce test, tous les animaux des groupes *NonDevalued* se comportent comme précédemment et montrent des taux d'appuis sur le levier importants. Les animaux du groupe *V-D* montrent un comportement semblable à celui exprimé lors du test RdR, avec des taux d'appuis très faibles (panneau A). Cependant, lorsque l'appui sur le levier est de nouveau récompensé, les animaux du groupe *V-D* modifient leur comportement et cessent d'appuyer sur le levier. Les animaux des groupes *Muscimol* expriment alors une sensibilité à la récompense dévaluée comme ceux des groupes *Vehicle*.

- * Une ANOVA à 2 facteurs, indique l'absence d'effet du **traitement** ($F_{1,33} = 0,75$; ns) et confirme l'effet significatif de la **dévaluation** ($F_{1,33} = 133,566$; $p < 0,001$). Elle ne montre aucune interaction **traitement x dévaluation** ($F_{1,33} = 0,165$; ns).
- * Un test SNK indique que les groupes *V-D* et *M-D* ne diffèrent pas entre eux mais diffèrent significativement des groupes *V-ND* et *M-ND*.

Les comportements de visite (panneau B) sont analogues aux comportements d'appui : les animaux des groupes *Devalued* ne visitent pratiquement pas la mangeoire en comparaison aux sujets des groupes *NonDevalued*.

- * L'analyse statistique confirme un effet significatif de la **dévaluation** ($F_{1,33} = 133,566$; $p < 0,001$) et n'indique aucun effet du **traitement** ($F_{1,33} = 0,056$; ns) ni d'interaction entre les deux facteurs ($F_{1,33} = 0,314$; ns).
- * Un test SNK indique que les groupes *V-D* et *M-D* ne diffèrent pas entre eux mais diffèrent significativement des groupes *V-ND* et *M-ND*.

Nous avons également examiné si les rats consomment les récompenses obtenues. Le nombre de récompenses obtenues au cours du test FdR et les indices moyens de consommation de chaque groupe sont représentés sur la figure III.1.8.

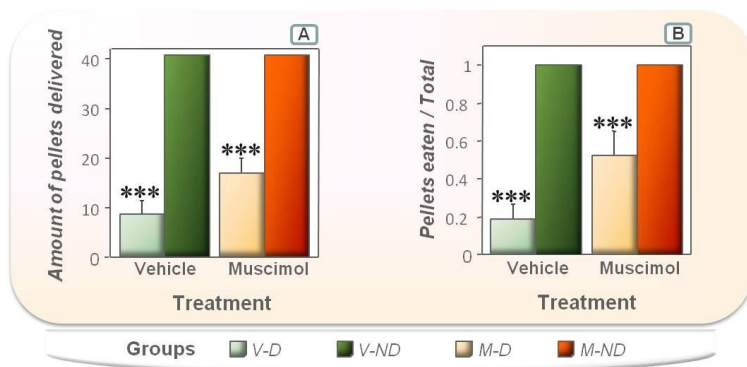


Figure III.1.8 Quantité de récompenses distribuées (A) et consommées (B) au cours du test FdR ($\pm$ erreur standard).

Les sujets des deux groupes *Devalued* reçoivent peu de récompense et en consomment moins de la moitié.

Les sujets des deux groupes *NonDevalued* consomment la totalité des 40 récompenses distribuées. Les animaux des groupes *Devalued* se comportent en accord avec leur comportement d'appui et ne consomment qu'en partie le peu de récompenses obtenues. Nous observons toutefois de légères différences entre ces deux groupes. En effet, le groupe *D-V* obtient en moyenne $8,5 \pm 2,5$ pastilles qu'ils consomment à hauteur de 18%. Les sujets du groupe *D-M* obtiennent en moyenne plus de récompenses $16,5 \pm 3$ et ils en consomment la moitié. Cette différence peut s'expliquer par un étalement des appuis dans le temps chez le groupe *D-M* et par un niveau d'aversion initialement moins fort à l'issue de la phase 3. Néanmoins, ces résultats indiquent bien que les groupes *D-V* et *M-D* possèdent une AGC pour la récompense.

- * Concernant le nombre de récompenses distribuées, l'analyse de variance indique des effets significatifs des facteurs **traitement** ($F_{1,33} = 4,734$; $p < 0,05$) et **dévaluation** ($F_{1,33} = 227,343$; $p < 0,001$) ainsi qu'une interaction significative **traitement x dévaluation** ($F_{1,33} = 4,734$; $p < 0,05$).
- * Un test SNK confirme que seuls les deux groupes *NonDevalued* ne diffèrent pas entre eux.
- * Concernant les indices de consommation, une ANOVA confirme les effets significatifs des facteurs **traitement** ($F_{1,33} = 5,730$; $p < 0,05$) et **dévaluation** ($F_{1,33} = 82,824$; $p < 0,001$) et une interaction significative entre ces deux facteurs ($F_{1,33} = 5,730$; $p < 0,05$).
- * Le test SNK indique que seuls les deux groupes *NonDevalued* ne diffèrent pas entre eux.

Conclusion

Les résultats de cette première expérience nous indiquent que l'absence d'un cortex prélimbique fonctionnel lors de l'acquisition d'une tâche instrumentale n'empêche pas cette acquisition, mais affecte la sensibilité de la réponse acquise vis-à-vis d'une dévaluation de la récompense. Un animal avec un PL inactivé apprend à appuyer sur le levier à une vitesse et un niveau de performance comparable à un rat possédant toutes ses facultés corticales. La diminution générale du comportement moteur observé (appuis et visites) est probablement imputable au stress provoqué par les micro-injections, ou encore à de possibles séquelles postopératoires.

Les données de la phase de dévaluation et du test récompensé attestent de l'existence d'une aversion gustative pour la récompense chez les animaux ayant eu leur PL inactivé pendant le conditionnement opérant. Néanmoins, ces derniers n'expriment pas de sensibilité à la dévaluation lors du test non récompensé. Ce résultat conduit à l'idée que les rats M-D n'utilisent pas de représentation de la récompense absente pour contrôler leur comportement d'appui.

Au vue de ces résultats, nous pouvons émettre l'hypothèse que la réponse instrumentale acquise en l'absence de fonctionnalité du cortex prélimbique dépend essentiellement d'une structure associative de type stimulus-réponse.

Expérience III.2

Implication du cortex prélimbique dans l'expression d'une action dirigée déjà acquise

Les premiers résultats de ce chapitre confirmant l'implication du cortex PL dans l'acquisition du caractère dirigé d'une action, nous nous proposons d'étudier l'implication de ce même cortex dans l'expression d'une action déjà acquise. Ainsi, nos sujets disposeront d'un cortex prélimbique fonctionnel pendant la phase d'acquisition et l'inactivation sera pratiquée seulement lors de la phase de test.

A. **Matériel et méthode**

a. Sujets, chirurgie, dispositif expérimental et histologie

Vingt-huit rats mâles naïfs de la souche Long-Evans (325 à 375 g) implantés dans le prélimbique avec deux guide-canules sont utilisés dans cette expérience. Les conditions d'hébergement, de chirurgie, de restriction alimentaire, le dispositif expérimental et l'histologie sont identiques à ceux décrits dans l'expérience précédente.

b. Protocole expérimental

Le protocole utilisé comprend les quatre phases décrites pour l'Expérience I.1. Contrairement au précédent protocole où des micro-injections répétées étaient effectuées lors de la phase 2, une seule micro-injection est effectuée en phase 4 juste avant les tests comportementaux (Figure III.2.1). Les phases 1 et 2 sont identiques à celles de l'Expérience I.1.



Figure III.2.1 Représentation schématique du déroulement de l'expérience.

La forme en éclair orange représente le moment où est effectuée l'unique micro-injection cérébrale.

Mag : séance d'entraînement à la mangeoire ; FR1 et VI30 : séances de conditionnement instrumental ; D : séance de dévaluation par AGC ; RdR : test pendant lequel la réponse dépend d'une représentation de la récompense ; FdR : test pendant lequel la réponse est renforcée.

- *Phase 3 : Dévaluation de la récompense et préparation aux micro-injections*

Les séances de dévaluation se déroulent comme établies à l'Expérience I.1. Les animaux sont répartis de façon égale dans un groupe *NonDevalued* et un groupe *Devalued*. Quatre heures après chacune des deux dernières séances de dévaluation, les sujets sont soumis à une séance d'habituation aux micro-injections dont le déroulement est identique à celui décrit dans la Phase 1 de l'Expérience III.1.

- *Phase 4 : Innactivation du PL et tests comportementaux*

La phase de tests se déroule sur une seule journée. Les animaux sont amenés dans la salle de micro-injection pour recevoir du muscimol ou du solvant (cf. Expérience III.1), puis sont soumis à un test RdR de 15 minutes, suivi 1 heure après par un test FdR (cf. Expérience I.1).

Les groupes d'animaux constitués selon la solution micro-injectée sont les suivants :

- ***Devalued-Vehicle (D-V)*** : la récompense a été dévaluée et le PL n'est pas inactivé en test ;
- ***NonDevalued-Vehicle (ND-V)*** : la récompense n'a pas été dévaluée et le PL n'est pas inactivé ;
- ***Devalued-Muscimol (M-D)*** : la récompense a été dévaluée et le PL est inactivé en test ;
- ***NonDevalued-Muscimol (M-ND)*** : la récompense n'a pas été dévaluée et le PL est inactivé.

c. Analyses statistiques

Les variables dépendantes et indépendantes analysées sont identiques à celles décrites dans l'expérience précédente.

B. Résultats

- *Histologie*

Tous les sujets montrent une bonne récupération postopératoire et sont soumis à la procédure de conditionnement. Les animaux retenus pour l'étude des résultats ont été sélectionnés, après examen des coupes histologiques, selon les critères définis en Expérience III.1 (figure III.2.2).

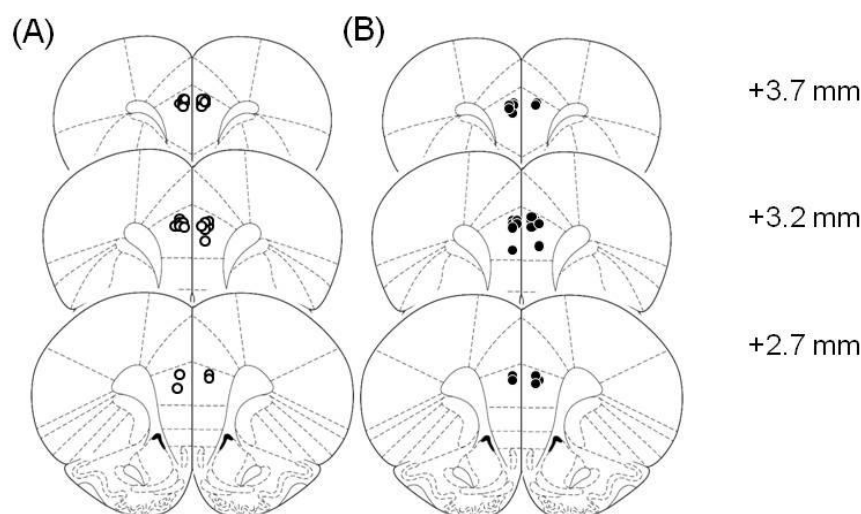


Figure III.2.2 Représentation schématique des sites d'injection dans le cortex prélimbique.

Disques blancs : groupes *Vehicle*. Disques pleins : groupes *Muscimol*. Les sections sont adaptées à partir de Paxinos et Watson (1998), elles sont disposées en fonction de leur position plus ou moins antérieure au bregma.

Les effectifs finaux pour chaque groupe sont les suivants : *D-M*, $n=7$; *ND-M*, $n=6$; *D-V*, $n=7$; *ND-V*, $n=7$.

- *Entraînement à la mangeoire*

Les sujets s'habituent sans difficulté aux cages de conditionnement et vont chercher la récompense dans la mangeoire (données non représentées).

- *Conditionnement instrumental*

Au cours de l'entraînement, les sujets implantés apprennent sans difficulté à appuyer sur le levier. Leurs taux d'appuis et de visites évoluent de manière comparable aux résultats précédemment obtenus chez des animaux non opérés (données non représentées).

* L'analyse statistique confirme l'effet significatif du facteur **séance** sur l'évolution des taux d'appuis ($F_{5,130} = 114,241$; $p < 0,001$) et des taux de visites ($F_{5,130} = 4,891$; $p < 0,001$).

A l'issue de cette 2^{ème} phase les sujets sont divisés en un groupe *NonDevalued* et un groupe *Devalued*, puis en 4 groupes *ND-V*, *ND-M*, *D-V*, *D-M*, à l'issue de la phase 3. Leurs taux de référence se situent aux environs de 11 appuis/min et 7 visites/min (données non représentées).

* Une ANOVA réalisée à posteriori confirme que les quatre groupes constitués pour la phase de test présentaient des performances instrumentales équilibrées avant la phase de dévaluation. En

effet, le facteur **groupe** n'a pas d'effet sur les taux de référence d'appuis ($F_{3,23} = 0,7016$; ns) ou de visites ($F_{3,23} = 2,482$; ns).

- *Dévaluation de la récompense par aversion gustative*

La figure III.2.3 présente les indices de consommation obtenus au cours de la phase de dévaluation et lors de la dernière séance de dévaluation.

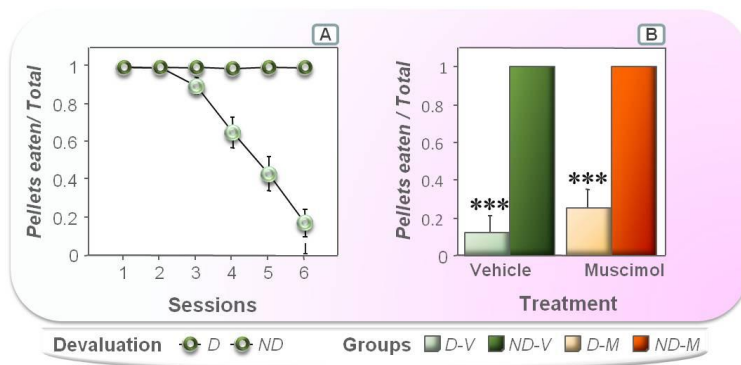


Figure III.2.3 Evolution des indices de consommation au cours des séances de dévaluation (A) et indices finaux moyens lors de la dernière ($\pm$ erreur standard). Les sujets des deux groupes Devalued acquièrent l'aversion gustative conditionnée pour la récompense à la même vitesse et à des niveaux comparables.

On peut suivre sur le panneau A l'évolution des indices de consommation au fil des séances. Comme attendu, les animaux du groupe *NonDevalued* consomment invariablement l'intégralité de la nourriture proposée (indice égal à 1), alors que la consommation du groupe *Devalued* diminue progressivement à partir de la 3^{ème} séance. Le comportement de visite évolue de manière identique (données non représentées).

- * Concernant le comportement de consommation, une ANOVA à 2 facteurs, indique des effets significatifs des facteurs **séance** ($F_{5,25} = 49,761$; $p < 0,001$) et **dévaluation** ($F_{5,125} = 46,371$; $p < 0,001$) et révèle aussi une interaction significative **séance x dévaluation** ($F_{5,125} = 46,346$; $p < 0,001$).
- * Concernant le comportement de visite, l'ANOVA confirme les effets significatifs des facteurs **séance** ($F_{5,25} = 8,282$; $p < 0,001$) et **dévaluation** ($F_{5,125} = 11,649$; $p < 0,01$) et révèle aussi une interaction significative **séance x dévaluation** ($F_{5,125} = 7,524$; $p < 0,001$).

A l'issue de cette phase les sujets sont répartis en quatre groupes pour les micro-injections, leurs indices finaux de consommations sont représentés sur le panneau B. On observe que les groupes *D-V* et *M-D* présentent des niveaux d'aversion proches.

- * Le test SNK confirme que les groupes planifiés pour la phase de test sont homogènes : les groupes *D-V* et *D-M* ne diffèrent pas entre eux.

- *Inactivation du PL et tests comportementaux*

- **Test RdR**

La figure III.2.4 représente les résultats obtenus lors du premier test après micro-injection cérébrale, en condition non récompensée.

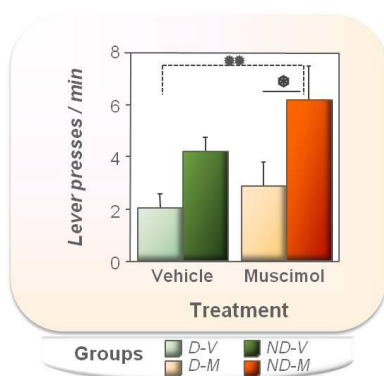


Figure III.2.4 Taux d'appuis au cours du test RdR (± erreur standard).

L'inactivation du cortex prélimbique n'a pas d'effet sur la sensibilité à la dévaluation : les sujets des deux groupes Devalued appuient peu sur le levier.

Chez les groupes *Vehicle*, de même que lors de l'expérience précédente, la performance instrumentale des animaux *Devalued* est plus faible que celle des sujets *NonDevalued*. Mais, contrairement à l'expérience précédente, les animaux des groupes *Muscimol* montrent aussi cette différence. En effet, le groupe *D-M* exprime un taux d'appuis plus faible que celui du groupe *ND-M*. Ainsi, l'inactivation prélimbique effectuée après la phase de conditionnement instrumental n'a pas d'effet sur la sensibilité à la dévaluation de la récompense.

- * Une ANOVA à 2 facteurs confirme l'effet significatif du facteurs **dévaluation** ($F_{1,23} = 2,781$; $p < 0,01$), l'absence d'effet du **traitement** ($F_{1,23} = 2,781$; ns) et l'absence d'interaction entre ces 2 facteurs ($F_{1,23} = 0,441$; ns).
- * Le test SNK indique des différences significatives seulement entre les groupes *ND-M* et *D-M* d'une part et les groupes *ND-M* et *D-V* d'autre part.

Les résultats obtenus pour le comportement de visite sont analogues : indépendamment de l'inactivation du PL, les individus des groupes *Devalued* visitent moins la mangeoire que les sujets des groupes *NonDevalued* (données non représentées).

- * Une ANOVA à 2 facteurs indique d'un effet significatif du facteurs **dévaluation** ($F_{1,23} = 14,977$; $p < 0,001$) mais ne montre aucun effet du **traitement** ($F_{1,23} = 0,487$; ns) ni d'interaction entre ces 2 facteurs ($F_{1,23} = 0,199$; ns).

- *Test FdR*

On peut observer sur la figure III.2.5 les résultats concernant les taux d'appuis obtenus lors du test FdR, le muscimol micro-injecté étant encore actif.

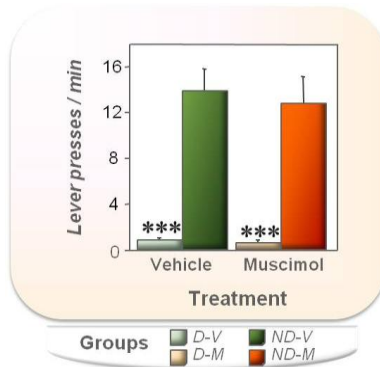


Figure III.2.5 Taux d'appuis au cours du test FdR (± erreur standard).

L'inactivation du cortex prélimbique n'a pas d'effet sur la sensibilité à la dévaluation : les sujets des deux groupes *Devalued* appuient peu sur le levier.

On retrouve lors du test FdR des résultats analogues à ceux du test RdR. Les sujets des groupes *NonDevalued* appuient activement sur le levier contrairement aux sujets des groupes *Devalued* dont les performances instrumentales sont très faibles.

- * Une ANOVA à 2 facteurs, confirme l'effet significatif de la **dévaluation** ($F_{1,23} = 78,702$; $p < 0,001$), l'absence d'effet du **traitement** ($F_{1,23} = 0,227$, ns) et indique que ces 2 facteurs n'interagissent pas ($F_{1,23} = 0,092$; ns).
- * Le test SNK confirme que les deux groupes *Devalued* diffèrent significativement des 2 groupes *NonDevalued*.

Les résultats relatifs au nombre de récompenses obtenues et aux comportements de visite ou de consommation concordent avec les résultats d'appuis sur le levier. Indépendamment de la solution micro-injectée, les deux groupes *Devalued* obtiennent peu de récompenses ($10,14 \pm 2$). Ils visitent peu la mangeoire (0,6 visites/min) et ne consomment que la moitié des pastilles obtenues (47%). Les sujets des groupes *NonDevalued* visitent fréquemment la mangeoire (10 visites/min) pour y trouver les 40 récompenses qu'ils consomment dans leur totalité (données non représentées).

- * Les analyses statistiques de ces résultats confirment des effets significatifs du facteur **dévaluation** ($F_{S\ 1,23} > 17$; $p_S < 0,001$), une absence d'effet du **traitement** ($F_{S\ 1,23} < 1$, $p_S > 0,50$) et indique que ces 2 facteurs n'interagissent pas ($F_{S\ 1,23} < 1$; $p_S > 0,50$).
- * Les tests SNK indiquent que les deux groupes *Devalued* diffèrent significativement des 2 groupes *NonDevalued*.

Conclusion

Dans cette 2^{ème} expérience, l'analyse des résultats des tests comportementaux, montre que suite à l'acquisition de la réponse instrumentale en conditions « normales », l'inactivation du cortex prélimbique est sans effet sur la sensibilité à la dévaluation de la récompense.

Par ailleurs, les résultats de la phase de conditionnement instrumental semblent indiquer l'absence de séquelles postopératoires. Nous n'observons pas la diminution des comportements d'appuis et de visite de l'expérience III.1, qui serait donc attribuable aux micro-injections et mises en contention répétées.

Le cortex prélimbique n'est donc pas indispensable à l'évocation de l'outcome et à la performance d'une action dirigée acquise au préalable.

Chapitre III : Discussion

Ce chapitre avait pour objectif d'étudier le rôle du cortex prélimbique dans le contrôle d'une action dirigée. Dans ce but, nous avons utilisé une approche d'inactivation cérébrale ciblée. Elle nous a permis d'identifier clairement la cinétique d'implication temporelle du PL. L'inactivation du PL pendant l'apprentissage n'empêche pas l'acquisition d'une réponse instrumentale, cependant elle modifie la nature de la réponse qui devient insensible aux variations de la valeur appétitive de la récompense. Au contraire, si le PL est inactivé après l'apprentissage, la réponse instrumentale conserve sa dépendance vis-à-vis de la valeur de la récompense.

Considérations méthodologiques

Micro-injections de muscimol

Nous avons opté pour l'implantation au préalable de guide-canules au niveau du PL ce qui nous a permis d'une part d'inactiver notre cible de manière temporaire lors de la phase choisie, et d'autre part d'éviter de devoir interrompre nos protocoles pour de la chirurgie (Killcross et Coutureau, 2003). L'agent utilisé au cours de nos expériences est le muscimol, couramment utilisé dans la littérature pour ses effets inhibiteurs sur l'activité neuronale. C'est un agoniste des récepteurs GABA_A dont la fixation hyperpolarise la membrane plasmique du corps cellulaire et diminue ainsi ses capacités de dépolarisation. Un tel effet pourrait perturber durablement l'activité neuronale, mais la durée d'action du muscimol n'excède pas 8 h. Ainsi, il a été montré qu'une fois le muscimol éliminé, la zone concernée par la micro-injection redevient fonctionnelle, aussi bien sur un plan physiologique que comportemental (pour revue voir Majchrzak et Di Scala, 2000 ; Edeline et al., 2002).

Au cours de l'Expérience III.1, les micro-injections répétées auraient pu provoquer une inflammation locale du tissu neural, mais nous n'avons pas constaté de gliose lors de l'observation de nos sites d'injection. De même, après avoir « levé l'anonymat » des coupes observées, nous n'avons pas constaté de dissemblance entre les groupes *Muscimol* et *Vehicle*. On constate néanmoins une diminution globale des performances lors de la phase de conditionnement opérant de l'Expérience III.1 et également lors du test FdR de l'Expérience

III.2. Cette baisse de comportement moteur est vraisemblablement imputable au stress provoqué par la mise sous contention, malgré toutes les précautions prises au préalable pour habituer les animaux à cette manipulation. Enfin, le muscimol étant une molécule à faible poids moléculaire diffusant facilement, il n'est pas exclu qu'injecté avec un grand volume sa sphère de diffusion n'atteigne des zones non ciblées (Edeline et al., 2002). Afin de limiter cette diffusion et d'épargner le cortex infralimbique, nous avons choisi une concentration de 1 $\mu\text{g}/\mu\text{L}$, un volume moyen de 0.25 μL (*i.e.* 0,25 μg de muscimol) et un site d'injection situé dans la partie dorsale du cortex prélimbique.

Dévaluation de la récompense

Afin de bien séparer dans le temps la phase de dévaluation de la récompense de celle de tests, nous avons choisi de dévaluer la récompense par aversion conditionnée plutôt que par satiété sélective. Compte tenu des résultats exposés dans le précédent chapitre, cette phase de dévaluation avait lieu dans le même contexte que la phase de conditionnement opérant. Dans l'expérience III.1, le niveau d'AGC est moins marqué dans le groupe *Muscimol* que dans le groupe *Vehicle*. Une étude des performances individuelles indique que cette différence est due à quelques animaux ($n=3$) du groupe *Vehicle* acquérant l'AGC plus rapidement qu'habituellement. En effet, l'indice de consommation du groupe *Vehicle-Devalued* est exceptionnellement bas par rapport à ceux précédemment observés dans les expériences des chapitres I et II. Au contraire, la consommation des animaux *Muscimol-Devalued* est dans la norme pour des animaux ayant subi une dévaluation. Il est donc peu probable que le fait d'avoir reçu du muscimol pendant l'entraînement ait rendu nos sujets moins sensibles au conditionnement aversif. De plus, ce niveau d'aversion pour la récompense est suffisant pour que la performance instrumentale se trouve fortement diminuée lors du test récompensé. Ainsi, cette différence apparente dans les niveaux d'aversion ne peut pas rendre compte de l'absence de sensibilité du groupe *Muscimol-Devalued* lors du test mettant en jeu une représentation de l'outcome.

Les résultats exposés dans ce troisième chapitre sont cohérents avec l'ensemble des résultats obtenus par l'utilisation d'autres méthodes exploratoires. Nous obtenons un effet de l'inactivation du PL pendant le conditionnement instrumental qui fait écho aux effets des lésions précédant l'apprentissage chez Corbit et Balleine (2003) ou chez Killcross et Coutureau (2003). L'absence d'effet de nos inactivations lorsqu'elles ont lieu pendant la phase de test confirme les résultats

obtenus chez Ostlund et Balleine (2005). Nos données nous permettent d'attribuer le déficit à la partie prélimbique du cortex préfrontal médian alors que ces auteurs avaient effectué des lésions plus étendues touchant également la région infralimbique. De plus, nos résultats obtenus avec le conditionnement d'un seul levier complètent l'étude d'Ostlund et Balleine dont les sujets étaient testés en présence des deux leviers. Or, il a été montré que la présentation simultanée de deux choix de réponses pouvait être source de conflit et que dans cette situation le sujet serait amené à évaluer plus complètement sa réponse ; il serait alors difficile d'obtenir des réponses de type habitude (Holland, 2004).

Rôle du cortex prélimbique dans les actions dirigées

L'activité au sein du PL n'est pas nécessaire pour l'acquisition d'une tâche instrumentale (Expérience III.1). Simultanément à la publication de nos résultats, Jonkman et al. (2009) ont publié des résultats allant également dans ce sens : si l'on empêche l'activité neuronale dans le prélimbique juste après les séances de conditionnement (i.e. les mécanismes de consolidation), la cinétique d'apprentissage n'est pas modifiée par rapport à un groupe contrôle. Cependant, contrairement à nos expériences, leur protocole n'inclut pas de phase visant à établir la nature de la réponse acquise dans de telles conditions. Nos observations, en accord avec la littérature (Corbit et Balleine, 2003 ; Killcross et Coutureau, 2003), suggèrent fortement que la réponse qu'ils auraient observée dans ces conditions aurait été insensible à la dévaluation de la récompense.

En effet, la région prélimbique est considérée comme essentielle au contrôle de l'action dirigée (Killcross et Coutureau, 2003). C'est cette région qui était visée par nos injections, et les résultats de ces expériences nous amènent à la conclusion que l'encodage, mais non l'expression, des associations entre la réponse et la récompense (A-O) est sous contrôle du PL. Dès 2003, Corbit et Balleine proposaient que la lésion du PL empêche l'acquisition d'une association entre la réponse et la récompense (A-O) mais sans affecter d'autres composantes de la réponse, telle que la réponse de type habitude (i.e. l'association S-R). En particulier, la répétition des actions au cours de la tâche permettrait une association entre la récompense en tant que stimulus et l'émission de la réponse selon un schéma (O-R). C'est pourquoi nous proposons comme explication à nos résultats que la réponse acquise sans le PL soit sous le contrôle exclusif du système habitude. Cette proposition inclut l'association O-R qui ferait alors partie du système habitude et pour laquelle seules les propriétés sensorielles de la récompense seraient prises en compte. Ainsi, la simple évocation de la

récompense dans la situation expérimentale suffirait à induire l'action à laquelle elle a été associée.

Cependant, le fait que la réponse instrumentale puisse être acquise dès le début selon un mode habitude chez les animaux dont le PL est inactivé va directement à l'encontre de la théorie séquentielle selon laquelle un fonctionnement en mode habitude remplacerait graduellement le système d'action dirigée A-O. De plus, une lésion ou une inactivation du cortex infralimbique permet d'observer à nouveau une sensibilité à la dévaluation chez des sujets surentraînés (Coutureau et Killcross, 2003 ; Killcross et Coutureau, 2003). Ainsi il semble que les deux systèmes associatifs soient activables à tout moment et acquis de façon parallèle. Selon Dickinson, c'est le degré de corrélation entre le taux de réponse et le taux de récompense qui déterminerait le contrôle par le système A-O plutôt que S-R (Dickinson et al., 1983, Yin et Knowlton, 2006). Une possibilité serait alors que l'inactivation du PL pendant l'apprentissage biaise la perception de cette corrélation et favorise l'intervention du système S-R. Nos résultats viennent ainsi compléter le riche corpus de données concernant le contrôle de la réponse instrumentale chez le Rat et sont en accord avec une vision plus moderne des implications temporelles des deux systèmes associatifs impliqués dans le contrôle de la réponse instrumentale.

Suite à l'apprentissage nous n'observons aucun effet de l'inactivation du PL. Ces données sont en accord avec les données obtenues précédemment par Ostlund et Balleine (2005). Nous pouvons donc affirmer que le cortex prélimbique n'est pas essentiel à l'évocation d'une représentation abstraite de la récompense si elle a été acquise alors qu'il était fonctionnel. Ce patron de résultats concernant le cortex prélimbique est notablement différent de celui du striatum dorso-médian postérieur (pDMS). En effet, la sensibilité à la dévaluation est affectée aussi bien par des lésions ou des inactivation du pDMS avant qu'après l'apprentissage (Yin et al., 2005a et b ; cf. introduction du Chapitre IV), ce qui suggère que cette région pourrait être le support de l'expression des associations A-O. Si le PL ne participe pas à l'expression des associations A-O, il pourrait jouer un rôle essentiel dans leur acquisition, notamment en maintenant en mémoire de travail une représentation conjointe de l'action et de la récompense (Corbit et Balleine, 2003).

En conclusion, les expérimentations menées au cours de ce chapitre nous ont permis de mettre clairement en évidence le moment d'implication du cortex prélimbique dans le contrôle de la réponse instrumentale : il est nécessaire à l'acquisition de son caractère dirigé, mais n'est pas indispensable à son expression. Ces résultats ont fait l'objet d'une publication dans *European Journal of Neuroscience* (2009) qui est jointe au présent mémoire.

Chapitre IV

Chapitre IV

Introduction

Les résultats présentés dans le chapitre III démontrent que le fonctionnement de la partie prélimbique du cortex préfrontal médian est indispensable pour l'acquisition d'une réponse dépendante du résultat. De manière remarquable, un ensemble de données suggère également que les régions striatales pourraient fonctionner en relation étroite avec le cortex prélimbique. L'aspect temporel doit néanmoins être pris en compte. En effet, dans un ensemble d'expériences d'exclusion temporaire ou permanente, Yin et collaborateurs (2005b) ont pu montrer que l'intégrité de la partie postérieure du cortex dorso-médian était requise à la fois pour l'acquisition et pour l'expression de la réponse médiée par le résultat (voir également Corbit et Janak, 2010 pour des résultats récents contradictoires). Dans des élaborations ultérieures, ces mêmes auteurs ont proposé que des changements de plasticité synaptique au sein de ces territoires striataux pourraient rendre compte de l'implication permanente du striatum dorsomédian. De fait, des phénomènes de plasticité de type Potentialisation à Long Terme (PLT) ont été décrits dans ces régions (Calabresi et al., 2007). De manière intéressante, des traitements connus pour altérer la PLT striatale altèrent également la performance et l'acquisition d'une réponse liée au résultat (Faure et al., 2005 ; Yin et al., 2005a ; Dang et al., 2006). Une hypothèse séduisante résultant de cette analyse consiste à proposer que l'activité au sein du cortex prélimbique liée aux premières étapes de l'apprentissage entraîne des phénomènes de plasticité au sein du striatum dorso médian.

Les expériences du chapitre IV visent à contribuer à cette thématique, en cela qu'elles visent à étudier les activations différentielles des étages striataux et préfrontaux durant les premières étapes de l'apprentissage. Pour cela, nous avons mis en œuvre une approche d'imagerie cérébrale *ex vivo* de révélation d'expression de la protéine Fos. Fos est un gène précoce (Immediate Early Gene : IEG), dont l'expression ne nécessite pas la synthèse de nouvelles protéines. Les produits de ces gènes sont communément utilisés comme marqueur d'activité et leur détection permet d'établir une cartographie fonctionnelle de l'effet d'une stimulation

de l'environnement sur l'activité cérébrale. Cette approche permet d'appréhender les processus de plasticité liés à un apprentissage à un moment donné et ainsi de déterminer les régions activées. Elle permet également de s'affranchir des problèmes liés au phénomène de compensation, ou de réorganisation cérébrale liés aux approches interventionnelles.

Dans ce cadre, les expériences de ce chapitre constituent une première étape dans l'étude des interactions cortico-striatales durant l'acquisition d'une réponse liée au résultat.

Expérience IV

Circuits neuronaux de l'action

Dans cette expérience, nous avons comparé l'activation cérébrale dans différentes conditions expérimentales. En effet, nous avons fait varier le nombre de séances d'entraînement. De manière à pouvoir contrôler que les animaux étaient exposés à la même quantité de récompense, nous avons opté pour une approche en animaux liés (yoked). Dans cette approche, lorsqu'une pastille était délivrée à un animal en conséquence d'un appui, une pastille était délivrée gratuitement au même moment à un autre animal.

A. Mathériel et méthode

a. Sujets et dispositif expérimental

Les sujets de cette expérience sont 72 rats mâles naïfs de la souche Long Evans (326 à 462g). Les conditions d'hébergement, de restriction alimentaire et le dispositif expérimental sont identiques à ceux décrits dans l'Expérience I.1.

b. Protocole expérimental

Comme illustré en figure IV.1 ce protocole comprend trois procédures de conditionnement différentes.

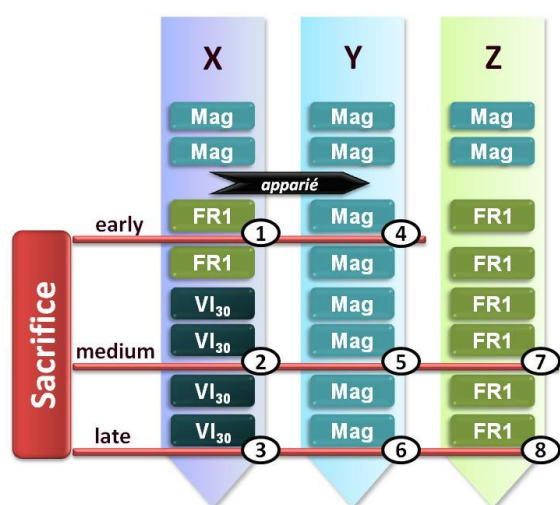


Figure IV.1 Schématisation des trois procédures X, Y, Z et des temps de sacrifice.

Mag : séance d'entraînement à la mangeoire ; FR1 et VI30 : séances de conditionnement instrumental. 8 groupes de sujets sont constitués :

- 1, groupe X.early/Z.early
- 2, groupe X.medium
- 3, groupe X.late
- 4, groupe Y.early
- 5, groupe Y.medium
- 6, groupe Y.late
- 7, groupe Z.medium
- 8, groupe Z.late.

Les deux premières procédures impliquent des couples de rats, placés dans deux cages de conditionnement connectées l'une à l'autre.

- **Procédure X** : les sujets sont soumis à un conditionnement instrumental identique à celui utilisé au cours des précédents chapitres (cf. Expériences I.1), il comprend 2 séances d'entraînement à la mangeoire, 2 séances instrumentales en FR1 et 4 séances en VI30.
- **Procédure Y** : les sujets ne sont pas soumis au conditionnement instrumental, mais à un entraînement continu à visiter la mangeoire. Les 2 premières séances se déroulent comme des séances classiques (cf. Expérience I.1) avec distribution automatisée des pastilles. En revanche au cours des 6 séances suivantes la distribution alimentaire survient uniquement dans le cas d'un appui récompensé émis au même moment par un rat de la procédure X dans la cage appariée.
- **Procédure Z** : les sujets sont entraînés continuellement sous un programme FR1.

L'étude de l'activité cérébrale est effectuée en trois temps : au début (*early*), au milieu (*medium*) et la fin (*late*) du conditionnement. On constitue ainsi 8 groupes de sujets :

- groupe **X.early/Z.early** : le conditionnement se limite à une séance en FR1, ils sont composés des même sujets ;
- groupe **Y.early** : il est constitué d'animaux appariés à ceux du groupe *X.early* ;
- groupe **X.medium** : le conditionnement est arrêté après la 4^{ème} séance de la procédure X ;
- groupe **Y.medium** : il est constitué d'animaux appariés à ceux du groupe *X.medium* ;
- groupe **Z.medium** : il est constitué d'animaux ayant effectué 4 séances de FR1 ;
- groupe **X.late** : le conditionnement est arrêté après la 6^{ème} séance de la procédure X ;
- groupe **Y.late** : il est constitué d'animaux appariés à ceux du groupe *X.late* ;
- groupe **Z.late** : il est constitué d'animaux ayant effectué 6 séances de FR1.

Les groupes sont entraînés de manière décalée afin que tous les animaux puissent être sacrifiés dans les mêmes conditions et le même jour. Afin d'équilibrer les groupes, quel que soit le type de programme (FR1 ou VI30) la séance instrumentale qui se tient le jour du sacrifice s'arrête au bout de 40 appuis ou après que 20 minutes se soient écoulées. Ainsi, les sujets des groupes *X.medium*, *Z.medium*, *X.late* et *Z.late* font l'expérience d'une séance plus courte que les précédentes.

c. Perfusion, prélèvement et coupe des cerveaux

A la fin de la séance instrumentale, les sujets rejoignent un groupe témoin d'animaux naïfs dans l'animalerie. Ils reçoivent une injection intra-péritonéale de Pentobarbital sodique (Ceva Santé Animale, 1mL/300g) au bout de 90 minutes (correspondant à l'atteinte du plateau d'expression de la protéine Fos). A l'aide d'une pompe intracardiaque, les sujets sont perfusés avec 300 mL de solution saline (0,9%) puis avec 300mL d'une solution de paraformaldéhyde (PFA, 4%) dilué dans du tampon phosphate (TP, 0,1M, pH 7.4). Les cerveaux, une fois prélevés, sont immergés dans du PFA et conservés à 4°C pendant 24h. Ils sont ensuite transférés dans une solution de sucre (30%) pour 48 heures. A l'aide d'un microtome à congélation (Leica SM 2400), les cerveaux sont sectionnés en coupes frontales de 30 µm d'épaisseur.

d. Immunohistochimie dirigée contre la protéine Fos

Les coupes sont déposées dans des plaques de 6 puits de façon à ce qu'une plaque comporte 6 cerveaux provenant de groupes différents. Elles sont rincées sous agitations 4 fois de suite dans du TP (0,1M) puis incubées dans une solution TP + H₂O₂ (0,3%). Cette étape vise à bloquer l'activité des peroxydases endogènes. Au bout de 30 minutes, elles sont soumises à 3 lavages successifs dans du TP puis placées pour 48h dans une solution contenant un anticorps primaire polyclonal de lapin anti-Fos (1/20000, Santa CruzBiotechnologies, USA) dilué dans un tampon de saturation (TP 0,1 M ; albumine de sérum bovin 0,1% ; Triton X-100 0,2% ; sérum de chèvre 0,2%).

Après 4 rinçages, les coupes sont transférées dans une dilution d'un anticorps secondaire biotinylé de chèvre anti-IgG de lapin (1/2000, Jackson Immunoresearch, USA). Au bout de 2h et après 3 rinçages, les coupes sont placées dans une solution TP contenant le complexe Avidine-Biotine-Peroxydase (kit Vectastain ABC, Vector Laboratories, USA) pour une durée de 1h30. Les coupes sont de nouveau rincées 3 fois et transférées dans une solution de 3,3-diaminobenzidine tetrahydrochloride (DAB, 1mg/mL, Sigma) pour une durée de 8 minutes. On ajoute 50µL de H₂O₂ (0,03%) et on laisse la réaction de colorimétrie s'engager pendant une durée de 11 minutes (cf. figure IV.2). Pour finir la réaction de peroxydation est arrêtée par 4 rinçages avec du TP.

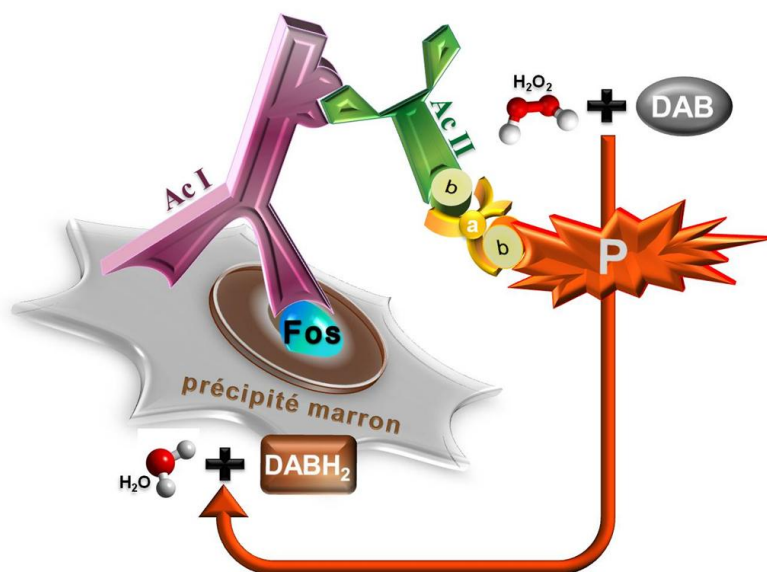


Figure IV.2 Schématisation de la réaction immunohistochimique.

La protéine nucléaire Fos est reconnue par l'anticorps primaire de lapin (Ac I), lui-même reconnu par l'anticorps secondaire (Ac II) anti IgG de lapin biotinylé (b). L'avidine (a) ayant une forte affinité pour la biotine permet la formation d'un complexe anticorps secondaire-biotine-avidine-biotine-péroxydase. L'ajout de l' H_2O_2 permet une réaction de précipitation de la DAB par action de la peroxydase. Il se forme alors un précipité brun au niveau du noyau cellulaire.

Après la réaction d'immunohistochimie, les puits contenant les cerveaux sont affectés d'un numéro aléatoire. Les coupes sont montées sur lames gélatinées et laissées à sécher sous hôte à température ambiante pour 48h. Les lames sont plongées dans deux bains successifs de toluène d'une durée de 10 minutes. Pour finir les coupes incluses dans de la résine (Eukitt®) sont montées sous lamelles.

e. Quantification du marquage

La quantification des noyaux Fos positifs est réalisée en aveugle et de façon semi-automatisée à l'aide du logiciel de traitement d'image Mercator® (Explora Nova) couplé à un microscope optique motorisé (Leica DM6000) et à une caméra (Opsonic Microfire). L'examen des coupes au faible grossissement (x2,5) permet de délimiter les régions d'intérêts en comparaison aux planches de l'atlas de Paxinos et Watson (1998). La procédure de quantification s'effectue en temps réel à l'objectif x10. La détection des noyaux Fos positifs se fait selon un seuillage défini au préalable par l'examineur selon des critères de formes, de couleur et de densité ; ce seuil est vérifié à chaque reprise des comptages.

La quantification est effectuée dans les deux hémisphères sur plusieurs niveaux de coupes variant selon la structure étudiée (cf. figure IV.3). Pour chaque animal, les valeurs sont établies en nombre de noyaux par mm^2 , puis moyennées et transformées par une racine carrée afin d'amenuiser les variations interindividuelles. Enfin, une moyenne est établie pour chaque

groupe correspondant aux 8 conditions comportementales. Afin de diminuer les effets dus à une activité de base, la moyenne de racines carrées des valeurs obtenues chez les animaux témoins est retranchée à la valeur obtenue.

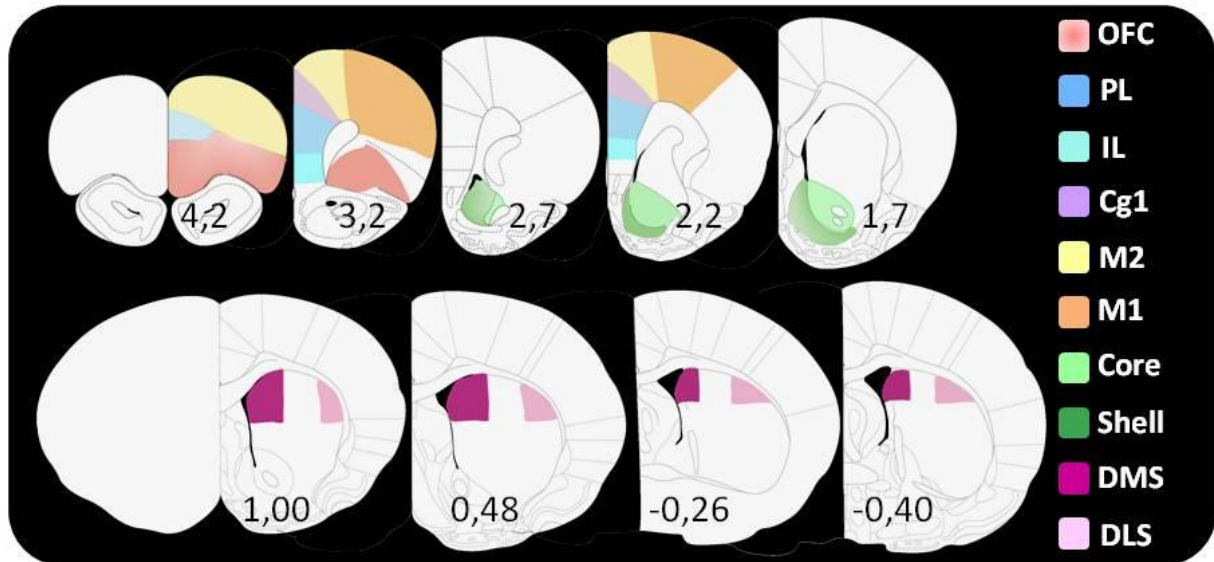


Figure IV. 3 Représentation schématique des sections frontales ayant servi à la quantification des noyaux Fos positifs.

Les coordonnées sont indiquées selon un axe antéropostérieur en fonction de la position du bregma.
Abréviations : cortex cingulaire, Cg1 ; coeur du noyau accumbens, Core ; striatum dorsomédian et dorsolatéral, DMS et DLS ; cortex infralimbique, IL ; cortex moteurs primaire et secondaire, M1 et M2 ; cortex orbitofrontal, OFC ; cortex prélimbique, PL ; coquille du noyau accumbens, Shell.

f. Analyses statistiques

Les résultats obtenus (*taux d'appuis, taux de visites, densité des noyaux positifs*) sont analysés à l'aide des variables indépendantes suivantes :

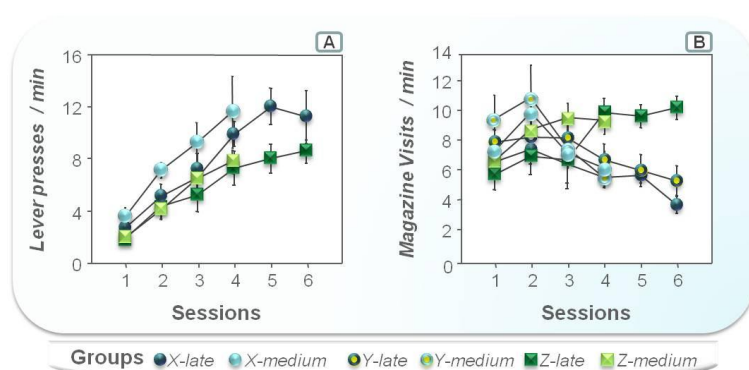
- Le facteur **séance**, correspondant à la répétition des séances au cours du conditionnement. Ce facteur concerne uniquement les groupes *medium* (4 séances) et *late* (6 séances) ;
- le facteur **timing** (*early/medium/late*), correspondant aux trois temps de sacrifice de la procédure d'entraînement ;
- le facteur **procédure** (*X/Y/Z*), correspondant aux trois types de procédures employées pendant le conditionnement ;
- le facteur **groupe**, correspondant à la répartition des sujets au sein des huit groupes ;
- le facteur **structure**, correspondant aux différentes régions cérébrales dans lesquelles sont quantifiés les noyaux Fos positifs.

B. Résultats

A l'exception de deux animaux, tous les sujets soumis aux procédures A et C ont réussi à acquérir la tâche instrumentale. Les effectifs finaux à l'intérieur de chaque groupe sont donc de 8 rats sauf pour les groupes *X.early/Z.early* et *Y.early* qui comprennent 6 sujets.

α. Données comportementales

On peut suivre, sur la figure IV.4, l'évolution des taux d'appuis et de visites au cours du conditionnement des groupes *late* et *medium*.



Le panneau A consacré au taux d'appuis moyen, nous montre que les animaux apprennent sans difficultés à appuyer sur le levier. Les profils évolutifs observés au cours des 4 premières séances sont similaires entre les groupes ayant suivi la même procédure mais sacrifiés à des temps différents : d'une part *X.medium* et *X.late* et d'autre part *Z.medium* et *Z.late*. On constate que les taux d'appuis sont légèrement plus faibles pour la procédure Z pour laquelle chaque appui est renforcé.

Le comportement de visite à la mangeoire est reporté sur le panneau B. On constate également que les profils évolutifs sont similaires chez les groupes soumis à la même procédure. Concernant les rats appariés entre eux, on ne constate pas de différence entre les groupes contrôlant la distribution (groupes X) et ceux qui la subissent (groupes Y). De façon contraire à nos observations pour des taux d'appuis, ce sont les sujets entraînés avec la procédure Z qui visitent le plus fréquemment la mangeoire, en effet chaque appui est récompensé et leurs séances sont donc plus courtes.

* Une analyse de l'ensemble des résultats centrée uniquement sur les quatre premières séances indique que le facteur **timing** n'a aucun effet ($F_s < 2,3$; ns), les groupes soumis à la même procédure mais sacrifiés à des moments différents sont bien comparables en ce qui concerne la phase d'acquisition. En revanche, il y a bien un effet significatif du facteur **séance** sur les taux d'appuis ($F_{3,84} = 40,822$; $p < 0,001$) et les taux de visites ($F_{3,126} = 2,789$; $p < 0,05$). Le type de

procédure influe significativement sur le comportement d'appui ($F_{1,28} = 5,186$; $p < 0,05$) mais pas sur le comportement de visite ($F_{2,42} = 1,062$; ns). Enfin, à l'exception d'une interaction **procédure** x **séance** significative pour les taux de visites ($F_{6,126} = 3,923$; $p < 0,01$), il n'existe aucune autre interaction entre les facteurs **séance**, **procédure** et **timing** ($F_s < 1,8$; ns).

- * Le test SNK confirme que les comportements de visite des groupes Y ne diffèrent pas de ceux des groupes X auxquels ils sont appariés. Il indique également que l'interaction **procédure** x **séance** concerne la quatrième séance et les groupes de la procédure Z.

Les résultats comportementaux observés le jour du sacrifice sont représentés sur la figure IV.5.

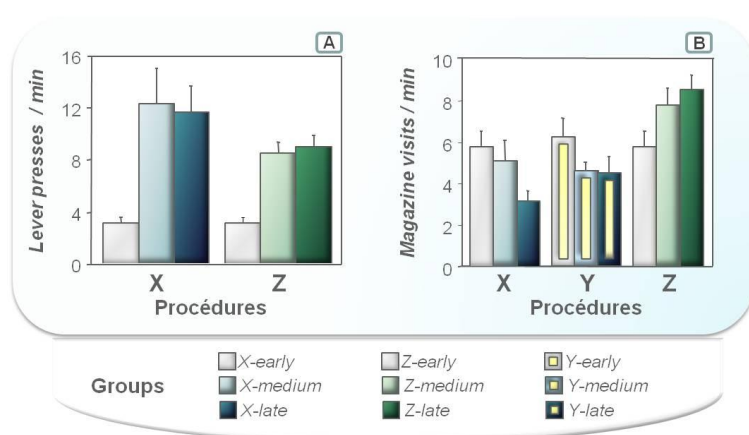


Figure IV.5 Performances comportementales avant le sacrifice : taux d'appuis (A) et de visites (B) ($\pm$ erreur standard). Confrontés au levier pour la première fois, les sujets du groupe Z.early mettent plus de temps pour effectuer 40 appuis. Les taux d'appuis sont plus élevés chez les groupes X, medium et late soumis à des séances en VI30. Au contraire, les animaux de la procédure Y (FR1) visitent plus la mangeoire. Enfin, les animaux de la procédure Y ont les mêmes taux de visites que les sujets auxquels ils sont appariés.

L'ensemble des sujets sélectionnés pour la quantification immunohistochimique de l'expression de la protéine Fos ont effectué 40 appuis. Néanmoins et conformément aux résultats précédents, les sujets émettent des taux appuis différents en fonction de la procédure et de l'état d'avancement de leur conditionnement (panneau A). En effet, on constate une performance faible aux environs de 2 appuis/min pour des animaux confrontés pour la première fois au levier (X.early/Z.early). Ils mettent en moyenne 14 minutes pour effectuer 40 appuis, la séance des groupes X.early/Z.early et Y.early est donc d'une durée trois fois plus longue que celles des six autres groupes avoisinant toutes les 5 minutes. Les taux d'appuis sont plus élevés (9 appuis/min) chez les groupes Z.medium et Z.late qui font l'expérience d'une 4^{ème} ou d'une 6^{ème} séance en FR1. Conformément aux résultats obtenus au cours des précédents chapitres, les taux exprimés pour les groupes X.medium et X.late (soumis à des séances en VI30) se situent aux environs de 12 appuis/min.

- * Une ANOVA simple avec le facteur groupe confirme que les taux d'appuis diffèrent significativement ($F_{5,38} = 5,334$; $p < 0,001$). Une ANOVA double indique un effet significatifs du

facteur **timing** ($F_{2,38} = 8,923$; $p < 0,001$) mais pas d'effet du facteur **procédure** ($F_{1,38} = 2,519$; ns) ni d'interaction entre ces deux facteurs ($F_{2,38} = 0,644$; ns).

- * Le test SNK localise des différences significatives uniquement entre le groupe *X.early/Z.early* et les autres groupes.

Les résultats concernant les taux de visites sont rapportés sur le panneau B. Conformément aux résultats décrits précédemment, les groupes *medium* et *late* soumis à la procédure Z visitent plus fréquemment la mangeoire que les 6 autres groupes, en effet chacun de leurs appui est récompensé. Les groupes soumis à la procédure A ont des taux de visites qui diminuent légèrement en fonction de leur degré d'avancement dans le conditionnement. On remarquera qu'à l'exception du groupe *Y.late*, les groupes Y qui n'effectuent pas de tâche opérante ont des taux avoisinant ceux des groupes X auxquels ils sont appariés. Enfin, les groupes *X.medium/Y.medium* et *X.late/Y.late* reçoivent en moyenne 6 récompenses au cours de cette ultime séance de conditionnement.

- * L'analyse statistique des taux de visites indique un effet significatifs du facteur **procédure** ($F_{2,57} = 11,109$; $p < 0,001$) une absence d'effet du facteur **timing** ($F_{2,57} = 0,443$; ns) mais une interaction significative entre ces deux facteurs ($F_{4,57} = 3,924$; $p < 0,01$).
- * Le test SNK confirme que le comportement de visite des groupes Z diffère significativement des comportements des groupes X et Y.

b. Quantification de la protéine Fos

Les résultats de la quantification des noyaux Fos positifs sont présentés sur la figure IV.6 pour les régions préfrontales et frontales (panneau A) et pour les ganglions de la base (panneau B).

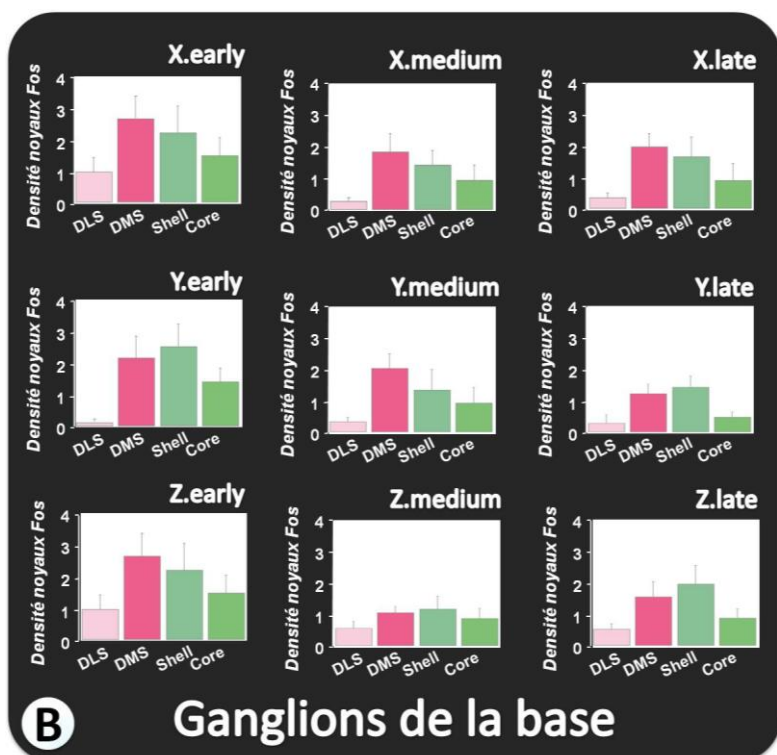
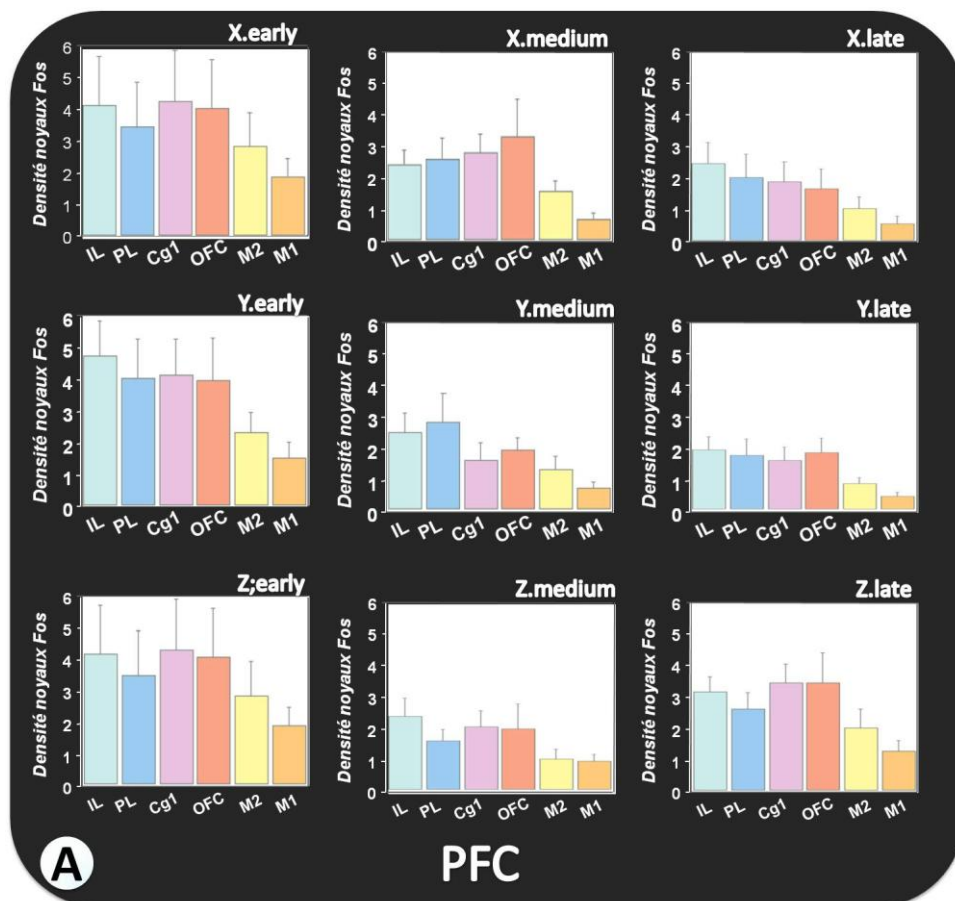


Figure IV.6 Quantification des noyaux Fos positifs au niveau du cortex préfrontal (A) et des ganglions de la base (B).

La densité correspond à la moyenne des valeurs obtenues pour chaque individu en appliquant la fonction racine carrée au nombre de noyaux positifs comptabilisés par mm^2 . Abréviations : cortex cingulaire, Cg1 ; cœur du noyau accumbens, Core ; striatum dorsomédian et dorsolatéral, DMS et DLS ; cortex infralimbique, IL ; cortex moteurs primaire et secondaire, M1 et M2 ; cortex orbitofrontal, OFC ; cortex prélimbique, PL ; coquille du noyau accumbens, Shell.

A la première vision de ces résultats il apparaît une tendance à une baisse générale de l'activation des structures chez les groupes *medium* et *late*. Cette observation n'est néanmoins pas valable pour la procédure Z ou un regain d'activité semble avoir lieu pour le groupe *Z.late*. De façon intéressante, chez les animaux des groupes Y qui n'apprennent pas de tâche d'appui les niveaux d'expression de la protéine Fos sont proches de ceux des animaux « acteurs » auxquels ils sont appariés. On constate chez les groupes *early* des barres d'erreur assez grandes qui pourraient traduire les différences interindividuelles dans l'appréhension d'une nouvelle situation (apparition du levier pour le groupe *X.early/Z.early* ; délai de distribution de la récompense plus long pour le groupe *Y.early*). Au sein des aires préfrontales, l'expression de la protéine Fos semble relativement homogène. Les cortex IL, PL, CG1 et OFC tendent à se distinguer des régions plus dorsales et latérales M1 et M2. Nous n'observons pas de différence entre la quantification des aires infralimbiques et prélimbiques. En revanche, en comparaison à l'activité au sein du DMS, le DLS est peu activé quel que soit la procédure dans laquelle sont engagés les sujets. La diminution du marquage avec l'avancement dans le conditionnement semble moins marquée au sein des ganglions de la base en comparaison aux régions préfrontales.

- * L'analyse statistique des résultats indique qu'il n'y a pas d'effet **procédure** ($F_{2,570} = 0,779$; ns) mais des effets significatifs des facteurs **structure** ($F_{9,570} = 16,308$; $p < 0,001$) et **timing** ($F_{9,570} = 27,858$; $p < 0,001$), les 3 facteurs n'interagissent pas entre eux.
- * Le test SNK localise les différences significatives entre les groupes *early* et les autres groupes *medium* ou *late*. Au sein du PFC les structures ne diffèrent pas à part le M1 et le M2. Au sein des noyaux gris centraux, seuls les DMS et la partie Shell du noyau accumbens ne diffèrent pas significativement dans leur activation.

Chapitre IV : Discussion

L'expérience menée au cours de ce chapitre visait à établir les cinétiques d'activation des régions préfrontales et striatales au cours d'un conditionnement instrumental. Les résultats obtenus confirment premièrement que la procédure appliquée peut induire des différences de comportement d'appui et de visite. Néanmoins, ces différences se reflètent peu dans les niveaux de l'expression de la protéine Fos. En revanche, les niveaux et l'évolution de l'expression de l'IEG varient avec le type de régions observées. Au début du conditionnement, la densité de noyaux Fos positifs est plus élevée au sein des régions préfrontales puis avec l'entraînement elle diminue et atteint des niveaux proches des structures striatales. La densité des noyaux est similaire au sein des cortex prélimbique et infralimbique, au contraire le marquage est plus dense au sein du striatum dorsomédian que dans le striatum dorsolatéral.

Données comportementales

Nos sujets ont été repartis dans des groupes selon trois procédures de conditionnement. La première (X) visait à identifier les structures initialement impliquées dans l'acquisition d'une tâche instrumentale ayant les mêmes caractéristiques que les réponses instrumentales étudiées au cours des chapitres précédents. De fait, nous obtenons une certaine cohérence entre les taux d'appuis et de visites observés chez les trois groupes *early*, *médium*, *late* et le profil évolutif obtenu au cours de l'Expérience I.1. Afin d'étudier l'effet du programme de renforcement, d'autres sujets (procédure Z) ont été entraînés selon un programme à renforcement continu (FR1). On constate que contrairement aux prédictions de Reynolds (1975 ; cf. Introduction générale), leurs taux d'appuis est plus faible que ceux d'animaux entraînés selon un programme VI. Cette différence est sûrement imputable au fait que tous les appuis étant récompensés, l'animal passe plus de temps à la mangeoire pour récupérer la nourriture. Effectivement, chez les groupes Z.medium et Z.late les taux de visites sont plus importants par rapport aux autres groupes. Une dernière procédure visait à contrôler l'effet du comportement opérant sur l'activité cérébrale. Ainsi, les animaux des groupes Y ont seulement été entraînés à visiter la mangeoire. Afin d'équilibrer les niveaux d'exposition à la récompense, nous avons lié ces animaux aux sujets acteurs de la procédure X. Cet appariement nous apporte des informations complémentaires sur le comportement de visite. En effet, la similitude entre les taux de visites dans les deux procédures indique que la visite à la mangeoire est bien de nature pavlovienne. Il semble effectivement que les visites aient lieu

de façon préférentielle au moment de la distribution de la récompense, le bruit provoqué par le fonctionnement du distributeur et la chute de la pastille constituant alors le stimulus conditionnel. De plus, il a été remarqué au sein de l'équipe que lors d'expériences impliquant la distribution d'une solution sucrée par un mécanisme silencieux, les visites à l'abreuvoir pour récupérer la récompense sont faibles. Alors, l'établissement d'un comportement d'appui est retardé, d'où l'utilisation d'un bruit de cliquer pour signaler la distribution (cf. Expérience I.4).

Mécanismes cellulaires et moléculaires des actions et habitudes

La quantification des noyaux Fos positifs nous a permis de déterminer le niveau d'activation de régions préfrontales et striatales au cours de notre entraînement instrumental. Les résultats obtenus pour les groupes acquérant l'action dirigée ne varie pas avec le type de programme de renforcement employé (fixe ou variable). Le groupe confronté au levier pour la première fois présente de hauts niveaux d'activation de ses régions préfrontales. Avec l'entraînement, l'activité baisse de moitié dans l'ensemble des structures préfrontales. On n'observe pas de dissemblance entre le profil d'activation des cortex prélimbique et infralimbique. A travers l'étude des différents groupes, le niveau d'activité au sein des ganglions de la base semble plus stable. Il est intéressant de constater d'une part que le striatum dorsolatéral est très peu activé par la tâche instrumentale et d'autre part qu'à la fin de l'entraînement les niveaux d'activité au sein du striatum dorsomédian sont similaires à ceux du cortex prélimbique. Ainsi, en présence de la récompense le PL semble « autant » impliqué dans la tâche instrumentale que le pDMS.

De façon étonnante, nous n'observons pas de différence de densité entre les animaux qui effectuent une action pour recevoir la récompense et ceux qui vont simplement la chercher dans la mangeoire. Néanmoins, la quantification est faite de manière automatique, sans différenciation entre les différentes couches corticales, sur l'ensemble de la région comprise dans le champ d'observation. Or, il a été démontré récemment que les structures ciblées par les neurones du cortex préfrontal médian dépendaient de leur couches cellulaires (Gabbott et al., 2005). Il n'est donc pas exclu que le nombre égal de noyaux activés concerne des cellules appartenant à des sous régions ou des couches différentes et impliquent ainsi des sous-circuits différents responsables des capacités cognitives engagées lors de nos trois procédures.

Nous nous sommes intéressés ici à la cinétique d'activation des structures préfrontales et striatales au cours de l'entraînement instrumental. Nous avons donc fait le choix de sacrifier nos animaux juste après des séances récompensées. Il est important de noter que les résultats

obtenus ici auraient pu être différents dans le cas d'une séance non récompensée où encore dans le cas d'une séance faisant suite à une dévaluation de la récompense. En effet, elles auraient directement impliqué une représentation mentale de la récompense et non pas le feedback en continu par le biais des propriétés sensorielles de la récompense.

Les mécanismes cellulaires et moléculaires impliqués dans l'acquisition des actions et des habitudes sont à ce jour peu étudiés. Des études portant sur la neuromodulation des activités cérébrales ont démontré que la dopamine et le glutamate avaient un rôle important au sein du striatum en cela qu'ils sont nécessaires à l'établissement d'une potentialisation à long terme et au maintien des niveaux de performance (Andrzejewski et al., 2004 ; Faure et al., 2005 ; Yin et al., 2005a ; Dang et al., 2006). Si la littérature abonde de résultats concernant les cinétiques d'expression des IEG dans différentes tâches d'apprentissage, peu d'études impliquent la nature de la réponse étudiée. Dans leur étude de 2005 Faure et collaborateurs ont mis en évidence une augmentation de l'expression de la protéine Fra-1 au niveau de la substance noire (pars compacta) et de l'aire tegmentale ventrale avec le surentraînement. En revanche, ils observent une évolution différente de niveaux d'expression génique au sein de la région CA1 de l'hippocampe en fonction de la nature de la tâche surentraînée : elle reste stable pour une procédure instrumentale, décroît pour une tâche d'appui associé à un stimulus et croît pour un conditionnement pavlovien.

L'étude menée par Hernandez et collaborateurs (2006) est très intéressante en cela qu'elle soulève la question du choix du marqueur d'activité. En effet, ces auteurs ont montré que la protéine Zif avait des niveaux d'expression qui diminuaient au fur à mesure de l'entraînement (23 séances instrumentales) dans l'ensemble des structures étudiées. Au contraire, on observe un patron évolutif variant avec la structure étudiée pour la protéine Homer 1a. Son expression diminue au sein des régions frontales et cingulaires au cours de l'entraînement. En revanche, elle se maintient au sein du striatum et atteint de hauts niveaux d'expression pour sa partie ventro-latérale.

Dans leur étude de 2010, Shiflett et collaborateurs se sont affranchis du choix de l'IEG en concentrant leurs investigations en amont de l'induction de l'expression génique. La protéine kinase ERK est fortement activée au début d'un conditionnement instrumental et de façon uniforme au sein du striatum dorsal. En revanche, l'inactivation de cette protéine n'entraîne pas les mêmes effets au sein de cette région. Effectuée de façon ciblée au niveau du DMS elle empêche l'acquisition d'une action dirigée. Au contraire lorsque la kinase ERK est

inhibée au niveau du DLS, le caractère dirigé de l'action est conservé mais les performances instrumentales sont fortement diminuées.

Dans leur ensemble, les données accumulées dans ce chapitre attestent de la dynamique d'activation au sein » » des circuits cortico-strataux.

Discussion générale

Discussion générale

Le travail présenté dans ce mémoire avait un certain nombre d'objectifs. Sur le versant purement comportemental, il s'agissait en premier lieu de mettre au point des procédures comportementales permettant d'explorer une réponse dirigée vers un but et sa sensibilité au protocole de conditionnement, et en second lieu de comprendre le rôle du contexte dans le contrôle de la réponse. Sur le plan neurobiologique, il s'agissait d'étudier le décours temporel de l'implication du cortex prélimbique dans les mécanismes de sélection de la réponse et d'étudier l'activation des circuits cortico-striataux. Dans son ensemble, ce travail apporte des données cohérentes contribuant au corpus actuel qui étudie la coordination des actions et habitudes. Ce courant considère que le contrôle comportemental résulte de la sélection en temps réel d'un système plutôt qu'un autre (voir Daw et al , 2005 ; Balleine et O'Doherty, 2010).

Au plan purement comportemental, les données accumulées dans les chapitres 1 et 2 ont montré que les conditions de sélection dépendent de manière étroite des conditions de l'encodage de l'information action-outcome. Dans le chapitre 1, nous avons identifié des conditions d'entraînement et de dévaluation aboutissant systématiquement à l'expression d'une réponse dirigée vers un but. Les données du chapitre 2 nous ont permis de préciser les mécanismes d'encodage aboutissant à une réponse sensible ou pas à des changements de la valeur de l'outcome. Plus précisément, nous avons mis en évidence un mécanisme tout à fait original qui attribue aux informations contextuelles un rôle central. Ces résultats trouvent un écho dans les chapitres 3 et 4 qui suggèrent que les mécanismes de sélection pourraient correspondre à l'activité coordonnée au sein de boucles corticostriatales. En particulier, nous montrons que le cortex préfrontal médian ne joue qu'un rôle temporaire dans l'adaptation de la réponse à des changements de valeur. Les données d'imagerie suggèrent également que l'encodage des relations action-outcome pourrait correspondre à des changements au niveau des interactions neuronales entre cortex préfrontal médian et striatum dorsomédian. En accord avec les données du chapitre 3, nous observons un engagement plus important du cortex préfrontal médian lors des phases initiales de l'acquisition.

Selon la conception classique, le contrôle comportemental résulte de la mise en œuvre séquentielle de deux systèmes de contrôle. Cette conception, d'abord ancrée chez les auteurs

classiques (voir Thorndike, 1911) a connu au travers des analyses théoriques d'A. Dickinson un renouveau spectaculaire (voir Dickinson, 1985). Le point central de cette conception, qui se retrouve dans un vaste champ de la littérature, et dans des procédures comportementales variées (ex : Salinas et al., 1993 ; Pennartz et al., 2009 ; Marsh et al., 2010) est que l'entraînement modifie le format de la représentation utilisée. Sans remettre en cause totalement cette idée, nos résultats suggèrent plutôt que le contrôle comportemental pourrait résulter d'un processus de sélection de la représentation. Plus précisément, les données de cette thèse suggèrent que l'outcome pourrait être encodé selon deux formes distinctes en parallèle, qui diffèrent par la valeur qui lui est attribué. Dans ce cadre, le contexte serait également pris en compte dans l'association et pourrait jouer un rôle dans l'utilisation de la représentation appropriée à la situation. Cette proposition s'écarte un peu des propositions actuelles, qui considèrent que la sélection de la réponse résulte d'un processus en deux étapes du type S-O, O-R (voir Ostlund et Balleine, 2008 ; Balleine et O'Doherty, 2010). Ainsi, ces auteurs proposent que la représentation de l'outcome évoquée par l'association S-O est purement sensorielle, dépourvue de sa valeur, et initie la réponse instrumentale. Par contre l'association O-R serait propre à faire exprimer toutes les caractéristiques motivationnelles de la récompense et contrôlerait le taux de production de cette réponse. Nous mettons en évidence au chapitre 2 un rôle critique du contexte au moment de l'encodage de l'outcome dévalué et suggérons que deux représentations O-R distinctes, correspondant à un même outcome dévalué ou non, peuvent coexister chez un même animal. Selon le contexte de test, l'une ou l'autre pourrait contrôler le taux de production de l'action. Une telle proposition doit évidemment être approfondie, et des expériences sont d'ores et déjà programmées dans l'équipe pour en généraliser les conditions expérimentales (type de contexte, entraînement, dévaluation, etc...).

Les données concernant le rôle du cortex préfrontal médian confirment que cette région joue un rôle déterminant pour l'adaptation de la réponse à des changements de valeur (Balleine et Dickinson, 1998 ; Killcross et Coutureau, 2003). L'implication temporaire de cette région démontre que les processus ayant lieu lors de l'encodage de la réponse sont tout à fait primordiaux pour la sélection ultérieure du type de réponse. Peut être le point le plus sujet à controverse correspond aux mécanismes par lesquels les animaux peuvent acquérir dès le début une réponse qui ne dépend pas d'une représentation de l'outcome, comme cela semble être le cas pour des rats privés de leur fonction préfrontale. Il s'agit d'une proposition provocatrice dans le sens qu'elle remet en cause directement la conception des habitudes

comme devant être acquises par un lent apprentissage. Il convient cependant de noter qu'une telle proposition a également été développée dans des modèles computationnels du contrôle (voir Daw et al., 2005). Là encore, notre proposition selon laquelle le contrôle comportemental résulte de la sélection d'une représentation différente en fonction de la valeur pourrait s'appliquer. De fait, il est tentant de proposer que l'inactivation du cortex prélimbique empêche, dès le début de l'entraînement, l'encodage de multiples représentations de la récompense. Ainsi, l'action resterait toujours dominée par la valeur de la représentation initialement acquise. Une approche corrélative est susceptible d'apporter des éléments déterminants à cette proposition. En particulier, l'enregistrement dans le cortex prélimbique des changements électrophysiologiques liés au codage de la valeur de l'outcome pourrait nous informer sur cette question. C'est le cas également des données d'imagerie cérébrale du type de celles du chapitre 4. Et cette approche pourrait s'enrichir de la dimension anatomique. Dans un ensemble d'expériences encore trop préliminaires pour être présentées dans ce mémoire, nous avons mis en œuvre une approche d'anatomie fonctionnelle en double marquage (voir pour exemple Lee et al., 2005 ; Petrovich et al., 2005) qui couple la révélation de la protéine Fos à celle d'un marqueur anatomique (Retrobeads, Lumafluor) rétrograde injecté dans le striatum dorsomédian. Par cette approche, encore en cours d'analyse, nous espérons pouvoir mettre en évidence des changements quantitatifs dans la dynamique des interactions entre régions frontales et régions striatales.

Ce travail a mis en œuvre une approche utilisant le corpus théorique et méthodologique des apprentissages associatifs. L'ambition de cette approche est d'étudier les processus de haut niveau par l'étude des formes élémentaires de la cognition. Ainsi, la recherche décrite dans ce mémoire devrait modestement contribuer à la description des mécanismes, à la base de processus de choix et de prise de décision. En effet, la prise de décision est un processus hautement intégratif dans lequel les aspects motivationnels liés au but sont intégrés aux aspects informationnels (valeur, lien causal...). En utilisant des tâches structurellement analogues, les données récentes suggèrent que les processus cognitifs mis en jeu sont extrêmement bien conservés chez les mammifères, des rongeurs jusqu'à l'homme (Tricomi et al., 2009 ; Balleine & O'Doherty, 2010). En plus de l'étude de la cognition dans sa dimension phylogénétique, cette recherche peut également apporter des informations cruciales concernant la dimension ontogénétique de la cognition. Il est envisageable de rechercher un lien entre le développement des fonctions cognitives associées au contrôle de l'action et la maturation des régions cérébrales impliquées (voir par exemple Klossek et al., 2008).

Enfin, la question du contrôle comportemental est un thème central en psychiatrie. En effet, les déficits dans ce processus sont à la base d'un certain nombre de propositions actuelles concernant des pathologies telles que la toxicomanie (Everitt et Robbins, 2005 ; Hogarth et al., 2010), les troubles obsessionnels compulsifs (Nielen et al., 2009 ; Kloft et al., 2010) ou encore la schizophrénie (Corlett et al., 2007 ; Hall et al., 2009).

Références bibliographiques

A

ADAMS, C.D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q J Exp Psychol* **34B**, 77-98.

ADAMS, C.D. AND DICKINSON, A. (1981). Instrumental responding following reinforcer devaluation. *Q J Exp psychol* **33B**, 109-121.

ALEXANDER, G. E., DELONG, M. R. and STRICK, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* **9**, 357-381.

AMALRIC, M., BAUNEZ, C. and NIEOULLON, A. (1995). Does the blockade of excitatory amino acid transmission in the basal ganglia simply reverse reaction time deficits induced by dopamine inactivation? *Behav Pharmacol* **6**, 508-519.

ANDRZEJEWSKI, M. E., SADEGHIAN, K. and KELLEY, A. E. (2004). Central amygdalar and dorsal striatal NMDA receptor involvement in instrumental learning and spontaneous behavior. *Behav Neurosci* **118**, 715-729.

ASHBY, F. G., TURNER, B. O. and HORVITZ, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn Sci* **14**, 208-215.

B

BALLEINE, B. W. (2005). Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. *Physiol Behav* **86**, 717-730.

BALLEINE, B. and DICKINSON, A. (1991). Instrumental performance following reinforcer devaluation depends upon incentive learning. *Q J Exp Psychol* **43B(3)**, 279-296.

BALLEINE, B. and DICKINSON, A. (1992). Signalling and incentive processes in instrumental reinforcer devaluation. *Q J Exp Psychol B* **45**, 285-301.

BALLEINE, B. W. and DICKINSON, A. (1998a). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**, 407-419.

BALLEINE, B. W. and DICKINSON, A. (1998b). The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. *Anim Learn Behav* **26**:46-59.

BALLEINE, B., GERNER, C. and DICKINSON, A. (1995). Instrumental outcome devaluation is attenuated by the anti-emetic ondansetron. *Q J Exp Psychol B* **48**, 235-251.

BALLEINE, B. W., KILLCROSS, A. S. and DICKINSON, A. (2003). The effect of lesions of the basolateral amygdala on instrumental conditioning. *J Neurosci* **23**, 666-675.

BALLEINE, B. W. and O'DOHERTY, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* **35**, 48-69.

BALLEINE, B. W. and OSTLUND, S. B. (2007). Still at the choice-point: action selection and initiation in instrumental conditioning. *Ann N Y Acad Sci* **1104**, 147-171.

BELIN, D., JONKMAN, S., DICKINSON, A., ROBBINS, T. W. and EVERITT, B. J. (2009). Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. *Behav Brain Res* **199**, 89-102.

BERRIDGE, K. C. (1991). Modulation of taste affect by hunger, caloric satiety, and sensory-specific satiety in the rat. *Appetite* **16**, 103-120.

BERRIDGE, K. C. and KRINGELBACH, M. L. (2008). Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology (Berl)* **199**, 457-480.

BERRIDGE, K. C., ROBINSON, T. E. and ALDRIDGE, J. W. (2009). Dissecting components of reward: 'liking', 'wanting', and learning. *Curr Opin Pharmacol* **9**, 65-73.

BLUNDELL, P., HALL, G. and KILLCROSS, S. (2003). Preserved sensitivity to outcome value after lesions of the basolateral amygdala. *J Neurosci* **23**, 7702-7709.

BOAKES, R. A., WESTBROOK, R. F., ELLIOTT, M. and SWINBOURNE, A. L. (1997). Context dependency of conditioned aversions to water and sweet tastes. *J Exp Psychol Anim Behav Process* **23**, 56-67.

BRODMANN, K. (1909). Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues, Johann Ambrosius Barth Verlag, Leipzig.

C

CALABRESI, P., PICCONI, B., TOZZI, A. and DI FILIPPO, M. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends Neurosci* **30**, 211-219.

CALZAVARA, R., ZAPPALA, A., ROZZI, S., MATELLI, M. and LUPPINO, G. (2005). Neurochemical characterization of the cerebellar-recipient motor thalamic territory in the macaque monkey. *Eur J Neurosci* **21**, 1869-1894.

CATANIA, A. C. and REYNOLDS, G. S. (1968). A quantitative analysis of the responding maintained by interval schedules of reinforcement. *J Exp Anal Behav* **11**, Suppl:327-383.

COLWILL, R. M. (2006). Passage of time leaves response-outcome associations intact. *Behav Processes* **72**, 91-99.

COLWILL, R.M. and RESCORLA, R.A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *J Exp Psychol* **11(1)**, 120-132.

COLWILL, R.M. and RESCORLA R.A. (1988). The role of response-reinforcer associations throughout extended instrumental training. *An Learn Behav.* **13(1)**, 105-111

COLWILL, R. M. and RESCORLA, R. A. (1990). Effect of reinforcer devaluation on discriminative control of instrumental behavior. *J Exp Psychol Anim Behav Process* **16**, 40-47.

CORBIT, L. H. and BALLEINE, B. W. (2000). The role of the hippocampus in instrumental conditioning. *J Neurosci* **20**, 4233-4239.

CORBIT, L. H. and BALLEINE, B. W. (2003). The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res* **146**, 145-157.

CORBIT, L. H. and BALLEINE, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J Neurosci* **25**, 962-970.

CORBIT, L. H. and JANAK, P. H. (2010). Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. *Eur J Neurosci* **31**, 1312-1321.

CORBIT, L. H., MUIR, J. L. and BALLEINE, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *J Neurosci* **21**, 3251-3260.

CORBIT, L. H., MUIR, J. L. and BALLEINE, B. W. (2003). Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *Eur J Neurosci* **18**, 1286-1294.

CORBIT, L. H., OSTLUND, S. B. and BALLEINE, B. W. (2002). Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *J Neurosci* **22**, 10976-10984.

CORLETT, P. R., MURRAY, G. K., HONEY, G. D., AITKEN, M. R., SHANKS, D. R., ROBBINS, T. W., BULLMORE, E. T., DICKINSON, A. and FLETCHER, P. C. (2007). Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain* **130**, 2387-2400.

COUTUREAU, E., GALANI, R., GOSSELIN, O., MAJCHRZAK, M. and DI SCALA, G. (1999). Entorhinal but not hippocampal or subicular lesions disrupt latent inhibition in rats. *Neurobiol Learn Mem* **72**, 143-157.

COUTUREAU, E., MARCHAND, A. R. and DI SCALA, G. (2009). Goal-directed responding is sensitive to lesions to the prelimbic cortex or basolateral nucleus of the amygdala but not to their disconnection. *Behav Neurosci* **123**, 443-448.

D

DALLEY, J. W., CARDINAL, R. N. and ROBBINS, T. W. (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neurosci Biobehav Rev* **28**, 771-784.

DANG, M. T., YOKOI, F., YIN, H. H., LOVINGER, D. M., WANG, Y. and LI, Y. (2006). Disrupted motor learning and long-term synaptic plasticity in mice lacking NMDAR1 in the striatum. *Proc Natl Acad Sci U S A* **103**, 15254-15259.

DAW, N. D., NIV, Y. and DAYAN, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* **8**, 1704-1711.

DAW, N. D., O'DOHERTY, J. P., DAYAN, P., SEYMOUR, B. and DOLAN, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876-879.

DAYAN, P. and DAW, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci* **8**, 429-453.

DE LA CASA, G. and LUBOW, R. E. (1995). Latent inhibition in conditioned taste aversion: the roles of stimulus frequency and duration and the amount of fluid ingested during preexposure. *Neurobiol Learn Mem* **64**, 125-132.

DE WIT, S. and DICKINSON, A. (2009). Associative theories of goal-directed behaviour: a case for animal-human translational models. *Psychol Res* **73**, 463-476.

DELAMATER, A. R. and OAKESHOTT, S. (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Ann N Y Acad Sci* **1104**, 1-20.

DERUSSO, A. L., FAN, D., GUPTA, J., SHELEST, O., COSTA, R. M. and YIN, H. H. (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Front Integr Neurosci* **4**.

DIAS-FERREIRA, E., SOUSA, J. C., MELO, I., MORGADO, P., MESQUITA, A. R., CERQUEIRA, J. J., COSTA, R. M. and SOUSA, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science* **325**, 621-625.

DICKINSON, A. (1985). Actions and habits: the development of behavioural autonomy. *Phil Trans R Soc Lond.* **308**, 67-78.

DICKINSON, A., and BALLEINE, B. (1994). Motivational control of goal-directed action, *An Learn Behav.* **22**, 1-18.

DICKINSON, A., and BALLEINE, B. (2002). The role of learning in the operation of motivational systems. In G. C.R. (Ed.), *Learning, motivation and emotion* (Vol. 3, p. 497-533). New York: John Wiley & Sons.

DICKINSON, A., BALLEINE, B., and WATT, A. (1995). Motivational control after extended instrumental training. *An Learn Behav.* **23(2)**, 197-206.

DICKINSON, A. and MULATERO, C.W. (1989). Reinforcer specificity of the suppression of instrumental performance on a non-contingent schedule. *Behav Process.* **19**, 167-180.

DICKINSON, A., NICHOLAS, D.J. and ADAMS C.D. (1983). The effect of the instrumental contingency and susceptibility to reinforcer devaluation. *Q J Exp Psychol.* **35B**, 35-51

DI FILIPPO, M., PICCONI, B., TANTUCCI, M., GHIGLIERI, V., BAGETTA, V., SGOBIO, C., TOZZI, A., PARNETTI, L. and CALABRESI, P. (2009). Short-term and long-term plasticity at corticostriatal synapses: implications for learning and memory. *Behav Brain Res* **199**, 108-118.

DOMJAN, M. and WILSON, N. E. (1972). Contribution of ingestive behaviors to taste-aversion learning in the rat. *J Comp Physiol Psychol* **80**, 403-412.

DOMJAN, M. (1977). Attenuation and enhancement of neophobia for edible substances. In L. Baker, M. Best, and M. Dojam (Eds.), *Learning mechanisms in food selection* (pp. 151-179). Waco: Baylor University Press.

DONOGHUE, J. P. and HERKENHAM, M. (1986). Neostriatal projections from individual cortical fields conform to histochemically distinct striatal compartments in the rat. *Brain Res* **365**, 397-403.

E

EDELINE, J. M., HARS, B., HENNEVIN, E. and COTILLON, N. (2002). Muscimol diffusion after intracerebral micro-injections: a reevaluation based on electrophysiological and autoradiographic quantifications. *Neurobiol Learn Mem* **78**, 100-124.

ELKINS, R. L. (1973). Attenuation of drug-induced bait shyness to a palatable solution as an increasing function of its availability prior to conditioning. *Behav Biol* **9**, 221-226.

EVERITT, B. J., FRAY, P., KOSTARCZYK, E., TAYLOR, S. and STACEY, P. (1987). Studies of instrumental behavior with sexual reinforcement in male rats (*Rattus norvegicus*): I. Control by brief visual stimuli paired with a receptive female. *J Comp Psychol* **101**, 395-406.

EVERITT, B. J. and STACEY, P. (1987). Studies of instrumental behavior with sexual reinforcement in male rats (*Rattus norvegicus*): II. Effects of preoptic area lesions, castration, and testosterone. *J Comp Psychol* **101**, 407-419.

EVERITT, B.J., ROBBINS, T.W. (2005) Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci* **8**:1481-1489.

F

FAURE, A., CONDE, F., CHERUEL, F. and EL MASSIOUI, N. (2006). Learning-dependent activation of Fra-1: involvement of ventral hippocampus and SNc/VTA complex in learning and habit formation. *Brain Res Bull* **68**, 233-248.

FAURE, A., HABERLAND, U., CONDE, F. and EL MASSIOUI, N. (2005). Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. *J Neurosci* **25**, 2771-2780.

FAURE, A., LEBLANC-VEYRAC, P. and EL MASSIOUI, N. (2010). Dopamine agonists increase perseverative instrumental responses but do not restore habit formation in a rat model of Parkinsonism. *Neuroscience* **168**, 477-486.

FLORESCO, S. B., ZHANG, Y. and ENOMOTO, T. (2009). Neural circuits subserving behavioral flexibility and their relevance to schizophrenia. *Behav Brain Res* **204**, 396-409.

FERSTER C.B. and SKINNER B.F. (1957). Schedules of Reinforcement. Englewood Cliffs, NJ: Prentice-Hall, 741 p.

FOUQUET, N., OBERLING, P. and SANDNER, G. (2001). Differential effect of free intake versus oral perfusion of sucrose in conditioned taste aversion in rats. *Physiol Behav* **74**, 465-474.

FRANCOIS, C., YELNIK, J., PERCHERON, G. and FENELON, G. (1994). Topographic distribution of the axonal endings from the sensorimotor and associative striatum in the macaque pallidum and substantia nigra. *Exp Brain Res* **102**, 305-318.

FUSTER, J. M. (2002). Frontal lobe and cognitive development. *J Neurocytol* **31**, 373-385.

G

GABBOTT, P. L., DICKIE, B. G., VAID, R. R., HEADLAM, A. J. and BACON, S. J. (1997). Local-circuit neurones in the medial prefrontal cortex (areas 25, 32 and 24b) in the rat: morphology and quantitative distribution. *J Comp Neurol* **377**, 465-499.

GABBOTT, P. L., WARNER, T. A., JAYS, P. R., SALWAY, P. and BUSBY, S. J. (2005). Prefrontal cortex in the rat: projections to subcortical autonomic, motor, and limbic centers. *J Comp Neurol* **492**, 145-177.

GARCIA, J. (1989). Food for tolmán, cognition and cathexis in concert. In T.Archer & L.-G. Nilsson (Eds.), *Aversion, avoidance, and anxiety* (pp. 45-85). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

GISQUET-VERRIER, P. and DELATOUR, B. (2006). The role of the rat prelimbic/infralimbic cortex in working memory: not involved in the short-term maintenance but in monitoring and processing functions. *Neuroscience* **141**, 585-596.

GLASCHER, J., HAMPTON, A. N. and O'DOHERTY, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex* **19**, 483-495.

GRILL, H. J. and NORNGREN, R. (1978). The taste reactivity test. I. Mimetic responses to gustatory stimuli in neurologically normal rats. *Brain Res* **143**, 263-279.

GRILL, H. J. and NORNGREN, R. (1978). The taste reactivity test. II. Mimetic responses to gustatory stimuli in chronic thalamic and chronic decerebrate rats. *Brain Res* **143**, 281-297.

- HABER, S. N., FUDGE, J. L. and MCFARLAND, N. R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* **20**, 2369-2382.
- HADDON, J. E., GEORGE, D. N. and KILLCROSS, S. (2008). Contextual control of biconditional task performance: evidence for cue and response competition in rats. *Q J Exp Psychol (Colchester)* **61**, 1307-1320.
- HALL, J., ROMANIUK, L., MCINTOSH, A. M., STEELE, J. D., JOHNSTONE, E. C. and LAWRIE, S. M. (2009). Associative learning and the genetics of schizophrenia. *Trends Neurosci* **32**, 359-365.
- HAMMOND, L. J. (1980). The effect of contingency upon the appetitive conditioning of free-operant behavior. *J Exp Anal Behav* **34**, 297-304.
- HAMPTON, A. N., BOSSAERTS, P. and O'DOHERTY, J. P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* **26**, 8360-8367.
- HEIDBREDER, C. A. and GROENEWEGEN, H. J. (2003). The medial prefrontal cortex in the rat: evidence for a dorso-ventral distinction based upon functional and anatomical characteristics. *Neurosci Biobehav Rev* **27**, 555-579.
- HEIMER, L., ALHEID, G.F., DE OLMOS, J.S., GROENEWEGEN, H.J., HABER, S.N., HARLAN, R.E. and ZAHM, D.S. (1997). The accumbens: beyond the core-shell dichotomy. *J Neuropsychiatry Clin Neurosci* **9**:354-381.
- HERNANDEZ, P. J., SCHILTZ, C. A. and KELLEY, A. E. (2006). Dynamic shifts in corticostriatal expression patterns of the immediate early genes Homer 1a and Zif268 during early and late phases of instrumental training. *Learn Mem* **13**, 599-608.
- HILARIO, M. R., CLOUSE, E., YIN, H. H. and COSTA, R. M. (2007). Endocannabinoid signaling is critical for habit formation. *Front Integr Neurosci* **1**, 6.
- HITCHCOTT, P. K., QUINN, J. J. and TAYLOR, J. R. (2007). Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. *Cereb Cortex* **17**, 2820-2827.
- HOGARTH, L., DICKINSON, A. and DUKA, T. (2010). The associative basis of cue-elicited drug taking in humans. *Psychopharmacology (Berl)* **208**, 337-351.
- HOLLAND, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process* **30**, 104-117.
- HOLLAND, P. C. (2008). Cognitive versus stimulus-response theories of learning. *Learn Behav* **36**, 227-241.
- HOLMAN, E.W. (1975). Some conditions for dissociation of consummatory and instrumental behavior in rats. *Learn Motiv.* **6**, 358-366.
- HOLMES, N. M., MARCHAND, A. R. and COUTUREAU, E. (2010). Pavlovian to instrumental transfer: a neurobehavioural perspective. *Neurosci Biobehav Rev* **34**, 1277-1295.
- HOOVER, W. B. and VERTES, R. P. (2007). Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. *Brain Struct Funct* **212**, 149-179.

J

JAY, T. M. and WITTER, M. P. (1991). Distribution of hippocampal CA1 and subicular efferents in the prefrontal cortex of the rat studied by means of anterograde transport of Phaseolus vulgaris-leucoagglutinin. *J Comp Neurol* **313**, 574-586.

JOEL, D. and WEINER, I. (1994). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *Neuroscience* **63**, 363-379.

JONKMAN, S., KOSAKI, Y., EVERITT, B. J. and DICKINSON, A. (2010). The role of contextual conditioning in the effect of reinforcer devaluation on instrumental performance by rats. *Behav Processes* **83**, 276-281.

JONKMAN, S., MAR, A. C., DICKINSON, A., ROBBINS, T. W. and EVERITT, B. J. (2009). The rat prelimbic cortex mediates inhibitory response control but not the consolidation of instrumental learning. *Behav Neurosci* **123**, 875-885.

K

KABLE, J. W. and GLIMCHER, P. W. (2009). The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733-745.

KELLEY, A. E., BALDO, B. A., PRATT, W. E. and WILL, M. J. (2005). Corticostriatal-hypothalamic circuitry and food motivation: integration of energy, action and reward. *Physiol Behav* **86**, 773-795.

KIEFER, S. W. and ORR, M. R. (1992). Taste avoidance, but not aversion, learning in rats lacking gustatory cortex. *Behav Neurosci* **106**, 140-146.

KILLCROSS, S. and COUTUREAU, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* **13**, 400-408.

KIM, H., SHIMOJO, S. and O'DOHERTY, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol* **4**, e233.

KIRSCH, I., LYNN, S. J., VIGORITO, M. and MILLER, R. R. (2004). The role of cognition in classical and operant conditioning. *J Clin Psychol* **60**, 369-392.

KLOFT, L., KISCHKEL, E., KATHMANN, N. and REUTER, B. (2010). Evidence for a deficit in volitional action generation in patients with obsessive-compulsive disorder. *Psychophysiology*.

KLOSSEK, U. M., RUSSELL, J. and DICKINSON, A. (2008). The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *J Exp Psychol Gen* **137**, 39-51.

KOLOMIETS, B. P., DENIAU, J. M., MAILLY, P., MENETREY, A., GLOWINSKI, J. and THIERRY, A. M. (2001). Segregation and convergence of information flow through the cortico-subthalamic pathways. *J Neurosci* **21**, 5764-5772.

L

LAROCHE, S., DAVIS, S. and JAY, T. M. (2000). Plasticity at hippocampal to prefrontal cortex synapses: dual roles in working memory and consolidation. *Hippocampus* **10**, 438-446.

LEE, H.J., GROSHEK, F., PETROVICH, G.D., CANTALINI, J.P., GALLAGHER and M., HOLLAND, P.C. (2005). Role of amygdalo-nigral circuitry in conditioning of a visual stimulus paired with food. *J Neurosci* **25**:3881-3888.

LEX, B. and HAUBER, W. (2010a). The role of dopamine in the prelimbic cortex and the dorsomedial striatum in instrumental conditioning. *Cereb Cortex* **20**, 873-883.

LEX, B. and HAUBER, W. (2010b). The role of nucleus accumbens dopamine in outcome encoding in instrumental and Pavlovian conditioning. *Neurobiol Learn Mem* **93**, 283-290.

LOPEZ, M., BALLEINE, B. and DICKINSON, A. (1992). Incentive learning following reinforcer devaluation is not conditional upon the motivational state during re-exposure. *Q J Exp Psychol B* **45**, 265-284.

LUBOW, R.E. (1989). Latent inhibition and conditioned attention theory. New York: Cambridge University Press.

LUBOW, R.E. (2009). Conditioned taste aversion and latent inhibition: a review. In *Conditioned Taste Aversion*. pp.57-37.

M

MARSH, R., HAO, X., XU, D., WANG, Z., DUAN, Y., LIU, J., KANGARLU, A., MARTINEZ, D., GARCIA, F., TAU, G. Z., YU, S., PACKARD, M. G. and PETERSON, B. S. (2010). A virtual reality-based fMRI study of reward-based spatial learning. *Neuropsychologia* **48**, 2912-2921.

MAJCHERZAK, M. and DI SCALA, G. (2000). GABA and muscimol as reversible inactivation tools in learning and memory. *Neural Plast* **7**, 19-29.

MATSUMOTO, K., SUZUKI, W. and TANAKA, K. (2003). Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science* **301**, 229-232.

MATSUMOTO, K. and TANAKA, K. (2004). The role of the medial prefrontal cortex in achieving goals. *Curr Opin Neurobiol* **14**, 178-185.

MELCER T and TIMBERLAKE W (1985). Poison avoidance and patch (location) selection in rat. *Animal Learning and Behavior* **13**, 60-68.

MEREDITH, G. E., BALDO, B. A., ANDREZJEWSKI, M. E. and KELLEY, A. E. (2008). The structural basis for mapping behavior onto the ventral striatum and its subdivisions. *Brain Struct Funct* **213**, 17-27.

MILLER, E. K. and COHEN, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* **24**, 167-202.

MOGENSEN, G. J., JONES, D. L. and YIM, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol* **14**, 69-97.

MULDER, A. B., NORDQUIST, R. E., ORGUT, O. and PENNARTZ, C. M. (2003). Learning-related changes in response patterns of prefrontal neurons during instrumental conditioning. *Behav Brain Res* **146**, 77-88.

N

NANEIX, F., MARCHAND, A. R., DI SCALA, G., PAPE, J. R. and COUTUREAU, E. (2009). A role for medial prefrontal dopaminergic innervation in instrumental conditioning. *J Neurosci* **29**, 6599-6606.

NELSON, A. and KILLCROSS, S. (2006). Amphetamine exposure enhances habit formation. *J Neurosci* **26**, 3805-3812.

NIELEN, M. M., DEN BOER, J. A. and SMID, H. G. (2009). Patients with obsessive-compulsive disorder are impaired in associative learning based on external feedback. *Psychol Med* **39**, 1519-1526.

NIV, Y., DAW, N. D. and DAYAN, P. (2006). Choice values. *Nat Neurosci* **9**, 987-988.

NORDQUIST, R. E., VOORN, P., DE MOOIJ-VAN MALSEN, J. G., JOOSTEN, R. N., PENNARTZ, C. M. and VANDERSCHUREN, L. J. (2007). Augmented reinforcer value and accelerated habit formation after repeated amphetamine treatment. *Eur Neuropsychopharmacol* **17**, 532-540.

NUTT, D. J., LALIES, M. D., LIONE, L. A. and HUDSON, A. L. (1997). Noradrenergic mechanisms in the prefrontal cortex. *J Psychopharmacol* **11**, 163-168.

O

O'DOHERTY, J., KRINGELBACH, M. L., ROLLS, E. T., HORNAK, J. and ANDREWS, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* **4**, 95-102.

OSTLUND, S. B. and BALLEINE, B. W. (2005). Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J Neurosci* **25**, 7763-7770.

OSTLUND, S. B. and BALLEINE, B. W. (2008). Differential involvement of the basolateral amygdala and mediodorsal thalamus in instrumental action selection. *J Neurosci* **28**, 4398-4405.

P

PARENT, A. and HAZRATI, L. N. (1994). Multiple striatal representation in primate substantia nigra. *J Comp Neurol* **344**, 305-320.

PARKER, L. A. (2006). The role of nausea in taste avoidance learning in rats and shrews. *Auton Neurosci* **125**, 34-41.

PAXINOS, G. and WATSON, C. (1998). The Rat Brain in Stereotaxic Coordinates. *Academic Press, San Diego, CA*.

PENNARTZ, C. M., BERKE, J. D., GRAYBIEL, A. M., ITO, R., LANSINK, C. S., VAN DER MEER, M., REDISH, A. D., SMITH, K. S. and VOORN, P. (2009). Corticostriatal Interactions during Learning, Memory Processing, and Decision Making. *J Neurosci* **29**, 12831-12838.

PETROVICH, G.D., HOLLAND, P.C., GALLAGHER, M. (2005). Amygdalar and prefrontal pathways to the lateral hypothalamus are activated by a learned cue that stimulates eating. *J Neurosci*, **5**(36), 8295-302.

PICKENS, C. L. (2008). A limited role for mediodorsal thalamus in devaluation tasks. *Behav Neurosci* **122**, 659-676.

PICKENS, C. L. and HOLLAND, P. C. (2004). Conditioning and cognition. *Neurosci Biobehav Rev* **28**, 651-661.

PICKERING, M. and JONES, J. F. (2002). The diaphragm: two physiological muscles in one. *J Anat* **201**, 305-312.

PREUSS, T. M. (1995). Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *J Cog Neurosci*, **7**, 1-24.

R

RAKIC, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci* **10**, 724-735.

REED, P. (2007). Response rate and sensitivity to the molar feedback function relating response and reinforcement rate on VI+ schedules of reinforcement. *J Exp Psychol Anim Behav Process* **33**, 428-439.

RESCORLA, R. A. (1988). Pavlovian conditioning. It's not what you think it is. *Am Psychol* **43**, 151-160.

RESCORLA, R.A. (1991). Associative relations in instrumental learning : the eighteenth barlett memorial lecture. *Q J Exp Psychol* **43B**(1), 1-23.

RESCORLA, R. A. (1992). Depression of an instrumental response by a single devaluation of its outcome. *Q J Exp Psychol B* **44**, 123-136.

RESCORLA, R. A. (1994). A note on depression of instrumental responding after one trial of outcome devaluation. *Q J Exp Psychol B* **47**, 27-37.

RESCORLA, R.A. and WAGNER, A.R. (1972). A theory of pavlovian conditioning : variations in the effectiveness of reinforcement and non-reinforcement. *Classical conditioning II : current research and theory*, pp. 64-99.

REYNOLDS, G.S. (1975). A primer of operant conditioning. Rev ed by Scott Foresome in Glenviews III.

ROBBINS, T. W. (2007). Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. *Philos Trans R Soc Lond B Biol Sci* **362**, 917-932.

RODRIGUEZ, M., LOPEZ, M., SYMONDS, M. and HALL, G. (2000). Lithium-induced context aversion in rats as a model of anticipatory nausea in humans. *Physiol Behav* **71**, 571-579.

ROLLS, E. T. (2005). Taste, olfactory, and food texture processing in the brain, and the control of food intake. *Physiol Behav* **85**, 45-56.

ROSAS, J. M. and CALLEJAS-AGUILERA, J. E. (2007). Acquisition of a conditioned taste aversion becomes context dependent when it is learned after extinction. *Q J Exp Psychol (Colchester)* **60**, 9-15.

ROSE, J.E. and WOOLSEY, C.N. (1948). The orbitofrontal cortex and its connections with the mediodorsal nucleus in rabbit, sheep and cat. *Res Publ Ass Nerv Ment Dis* **27**:210-232.

S

SALINAS, J. A., PACKARD, M. G. and MCGAUGH, J. L. (1993). Amygdala modulates memory for changes in reward magnitude: reversible post-training inactivation with lidocaine attenuates the response to a reduction in reward. *Behav Brain Res* **59**, 153-159.

SCHILMAN, E. A., UYLINGS, H. B., GALIS-DE GRAAF, Y., JOEL, D. and GROENEWEGER, H. J. (2008). The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. *Neurosci Lett* **432**, 40-45.

SCHULTZ, W. and DICKINSON, A. (2000). Neuronal coding of prediction errors. *Annu Rev Neurosci* **23**, 473-500.

SHANKS, D. R. (2010). Learning: from association to cognition. *Annu Rev Psychol* **61**, 273-301.

SHIFLETT, M. W., BROWN, R. A. and BALLEINE, B. W. (2010). Acquisition and performance of goal-directed instrumental actions depends on ERK signaling in distinct regions of dorsal striatum in rats. *J Neurosci* **30**, 2951-2959.

SEAMANS, J. K. and YANG, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* **74**, 1-58.

SMITH, Y., BEVAN, M. D., SHINK, E. and BOLAM, J. P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience* **86**, 353-387.

STRIEDTER, G. F. and NORTHCUTT, R. G. (2006). Head size constrains forebrain development and evolution in ray-finned fishes. *Evol Dev* **8**, 215-222.

T

TANAKA, S. C., BALLEINE, B. W. and O'DOHERTY, J. P. (2008). Calculating consequences: brain systems that encode the causal effects of actions. *J Neurosci* **28**, 6750-6755.

TANAKA, S. C., DOYA, K., OKADA, G., UEDA, K., OKAMOTO, Y. and YAMAWAKI, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci* **7**, 887-893.

THORNDIKE, E. L. (1911). *Animal intelligence*. Macmillan.

TRICOMI, E., BALLEINE, B. W. and O'DOHERTY, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci* **29**, 2225-2232.

U

UYLINGS, H. B., GROENEWEGEN, H. J. and KOLB, B. (2003). Do rats have a prefrontal cortex? *Behav Brain Res* **146**, 3-17.

V

VALENTIN, V. V., DICKINSON, A. and O'DOHERTY, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J Neurosci* **27**, 4019-4026.

VERTES, R. P. (2004). Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse* **51**, 32-58.

VERTES, R. P. (2006). Interactions among the medial prefrontal cortex, hippocampus and midline thalamus in emotional and cognitive processing in the rat. *Neuroscience* **142**, 1-20.

VOORN, P., VANDERSCHUREN, L. J., GROENEWEGEN, H. J., ROBBINS, T. W. and PENNARTZ, C. M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci* **27**, 468-474.

W

WALTON, M. E., CROXSON, P. L., BEHRENS, T. E., KENNERLEY, S. W. and RUSHWORTH, M. F. (2007). Adaptive decision making and value in the anterior cingulate cortex. *Neuroimage* **36 Suppl 2**, T142-154.

WALTON, M. E. (2009). Evaluating and revaluating outcomes in the frontal lobe (commentary on Kennerley and Wallis). *Eur J Neurosci* **29**, 2060.

WANG, S. H., OSTLUND, S. B., NADER, K. and BALLEINE, B. W. (2005). Consolidation and reconsolidation of incentive learning in the amygdala. *J Neurosci* **25**, 830-835.

WASSERMAN, E. A. and MILLER, R. R. (1997). What's elementary about associative learning? *Annu Rev Psychol* **48**, 573-607.

WILSON, C.L., SHERMAN, J.E. and HOLMAN, E.W. (1981). Aversion to the reinforcer affects conditioned reinforcement and instrumental responding. *J Exp Psychol* **2**, 165-174.

WISE, S. P. (2008). Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci* **31**, 599-608.

Y

YIN, H. H. and KNOWLTON, B. J. (2004). Contributions of striatal subregions to place and response learning. *Learn Mem* **11**, 459-463.

YIN, H. H. and KNOWLTON, B. J. (2006). The role of the basal ganglia in habit formation. *Nat Rev Neurosci* **7**, 464-476.

YIN, H. H., KNOWLTON, B. J. and BALLEINE, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci* **19**, 181-189.

YIN, H. H., KNOWLTON, B. J. and BALLEINE, B. W. (2005a). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur J Neurosci* **22**, 505-512.

YIN, H. H., KNOWLTON, B. J. and BALLEINE, B. W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav Brain Res* **166**, 189-196.

YIN, H. H., MULCARE, S. P., HILARIO, M. R., CLOUSE, E., HOLLOWAY, T., DAVIS, M. I., HANSSON, A. C., LOVINGER, D. M. and COSTA, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci* **12**, 333-341.

YIN, H. H., OSTLUND, S. B. and BALLEINE, B. W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur J Neurosci* **28**, 1437-1448.

YIN, H. H., OSTLUND, S. B., KNOWLTON, B. J. and BALLEINE, B. W. (2005b). The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* **22**, 513-523.

YIN, H. H., ZHUANG, X. and BALLEINE, B. W. (2006). Instrumental learning in hyperdopaminergic mice. *Neurobiol Learn Mem* **85**, 283-288.

Annexes

BEHAVIORAL NEUROSCIENCE

Transient role of the rat prelimbic cortex in goal-directed behaviour

Delphine A. S. Tran-Tu-Yen,^{1,2} Alain R. Marchand,^{1,2} Jean-Rémi Pape,^{1,2} Georges Di Scala^{1,2} and Etienne Coutureau^{1,2}

¹CNRS, Centre de Neurosciences Intégratives et Cognitives, UMR 5228, Talence, F-33405, France

²Université de Bordeaux, UMR 5228, Talence, F-33405, France

Keywords: action, muscimol, outcome devaluation, prefrontal cortex, rat

Abstract

Lesion studies show that goal-directed actions mediated by action–outcome (A–O) associations and habits mediated by stimulus–response (S–R) associations can be dissociated during instrumental training, with the prelimbic region of the medial prefrontal cortex being involved in the former and the infralimbic region in the latter. The present work further investigates the role of the prelimbic region in acquisition vs. expression of goal-directed instrumental behaviour, using reversible neuronal inactivation and outcome devaluation procedures. In a first experiment, inactivating the prelimbic cortex at the time of testing did not alter the sensitivity to devaluation, indicating that this region was not essential for the expression of A–O associations. In a second experiment, the prelimbic cortex was inactivated throughout the training phase. At the time of testing the performance was insensitive to devaluation, indicating that the acquired response was not goal-directed but mediated by an S–R association. These data challenge the view that the habit system replaces the goal-directed system as training progresses. They show that the prelimbic cortex plays a transient but crucial role in the acquisition of goal-directed responding and that the A–O and S–R systems can operate in a competitive fashion early in training.

Introduction

To survive in a changing environment, organisms have to make adapted decisions by selecting actions while constantly monitoring their consequences. According to current views, instrumental conditioning in both rodents (Dickinson, 1985; Balleine & Dickinson, 1998) and primates, including humans (Valentin *et al.*, 2007; Tanaka *et al.*, 2008), may capture many of the basic cognitive mechanisms of these selection processes.

Instrumental training can no more be considered as reflecting a simple stimulus–response (S–R) association, as it can rely on flexible associative processes (Colwill & Rescorla, 1985; Balleine & Dickinson, 1998) requiring the representation of the outcome as a goal. Indeed, in rats moderately trained to press a lever for food, reducing the appetitive value of the food (outcome devaluation) or the relation between lever press and outcome delivery (contingency degradation) results in a reduction of lever pressing. Thus, early in acquisition, responses are adaptable and goal-directed, i.e. mediated by an action–outcome (A–O) association. In contrast, in overtrained rats or in rats trained with specific schedules of reinforcement, performance is sensitive to neither outcome devaluation nor contingency degradation (Adams, 1982; Dickinson *et al.*, 1983), thus reflecting a habit S–R association. Furthermore, recent data challenge the classical theory that the S–R system gradually replaces the A–O association through

training (Dickinson, 1985) and indicate instead that both systems are present in parallel and/or competition early as well as late in training (Daw *et al.*, 2005).

These two associative processes rely upon neural circuits involving the medial prefrontal cortex (mPFC) (Balleine, 2005; Daw *et al.*, 2005). However, the mPFC appears heterogeneous (for review see Heidbreder & Groenewegen, 2003), comprising (dorsally to ventrally): the cingulate cortex, prelimbic cortex (PL) and infralimbic cortex. This heterogeneity is reflected in instrumental conditioning. Indeed, animals with pretraining lesions or pharmacological inactivation affecting the infralimbic cortex (Coutureau & Killcross, 2003; Killcross & Coutureau, 2003) maintain A–O responding under conditions of overtraining that normally promote S–R responding. By contrast, lesions of the PL (Corbit & Balleine, 2003; Killcross & Coutureau, 2003; Coutureau *et al.*, 2009) result in S–R-based instrumental performance that is insensitive to outcome devaluation early in training, i.e. under conditions promoting A–O responding.

The mPFC is generally thought to participate in the ‘on line’ control of behaviour (Miller & Cohen, 2001; Dalley *et al.*, 2004). However, lesion experiments have suggested its involvement in the acquisition rather than the expression of goal-directed instrumental behaviours (Ostlund & Balleine, 2005). The present experiments were thus aimed at clarifying this issue using a reversible inactivation approach with the GABA_A agonist muscimol. With this approach, implantation surgery is performed at a single time-point prior to the experiments in all animals, thereby allowing better control of time-dependent phenomena across learning. Rats trained to lever press for a food

Correspondence: Dr E. Coutureau, ¹CNRS, as above.
E-mail: e.coutureau@cnic.u-bordeaux1.fr

Received 16 February 2009, revised 4 May 2009, accepted 9 June 2009

reward were tested after devaluation induced by food aversion. In Experiment 1, the prefrontal area was inactivated just before testing (i.e. once the instrumental response had been acquired), whereas in Experiment 2 it was inactivated throughout instrumental training but not at test.

Materials and methods

Experiment 1: prefrontal cortex inactivation (pretest)

Subjects

The subjects were 28 experimentally naive, male Long-Evans rats (325–375 g), obtained from a commercial supplier (Centre d'Élevage et de Recherche Janvier, France). Rats were housed two per cage (polycarbonate, 49 × 26 × 20 cm) in a temperature-controlled room (22 °C). Before the beginning of the experiment, animals were accustomed for 1 week to the vivarium, where they had *ad libitum* access to food and water. Experiments were conducted during the light portion of the 12 h light/dark cycle. All procedures were carried out in accordance with the French Directive (87/148, Ministère de l'Agriculture et de la Pêche) and the European Communities Council Directive (86/609/EEC).

Surgical procedure

For implantation of guide cannulas, rats first received intraperitoneal injections of Xylazine (13 mg/kg; Rompun®) and Diazepam (3 mg/kg; Valium®) and were then anaesthetized with Ketamine (100 mg/kg; Virbac). They were then placed on a stereotaxic frame (Kopf Instruments, Tujunga, CA, USA) in a flat skull position. The bone of the skull above the region to be implanted was removed using a high-speed drill. Guide cannulas (8 mm long, 36 gauge; Le Guellec, France) were lowered into the PL at the following coordinates (in mm from Bregma): anteroposterior, +3.2; lateral, ± 0.6; ventral, –2.5. Guide cannulas were held in place by a block of dental cement (Palavit®G) overlying three small skull-screws. Removable stylets were inserted in the cannulas to protect them from dust. Rats were left in a warm postsurgical room for 24 h with *ad libitum* access to food and water. They were then replaced in the vivarium for 1 week of recovery and were individually handled every day.

Behavioural apparatus

Training took place in eight identical operant boxes, individually housed in sound- and light-attenuating chambers (Imetronic, France). Boxes measured 40 × 30 × 35 cm and were made of grey PVC on the three sides except the opening face, which was made of transparent Perspex. A ventilation fan produced a background noise of 55 dB. The box was illuminated from its ceiling by four LED lamps and its floor was constructed of 27 parallel 5 mm diameter stainless-steel bars spaced 1 cm apart. At the centre of the left-hand wall, food pellets (45 mg, F0165; Bio_Serv) could be delivered into a recessed

magazine by a food dispenser located outside the operant box. A flat-panel retractable lever was located at the left of the magazine. A personal computer equipped with SKAA_PROG software (Imetronic) controlled the equipment and recorded the data.

Behavioural procedure

Food restriction began the day before the first magazine-training session. During all the behavioural procedure rats had *ad libitum* access to water and received 15 g of their habitual food each evening after training to maintain them at ~90% of their original weight. Training consisted of one session per day (see Fig. 1).

Magazine-training. Prior to instrumental training, two 40 min magazine-training sessions were performed. Each session began with illuminating and ended with turning off the house light. The lever was not available and 40 food rewards were delivered according to a random time 60 s schedule.

Instrumental training. Each rat then received six instrumental training sessions. A session began with chamber illumination and lever insertion. When 40 pellets had been delivered or after 45 min had passed, the session ended, the lever was retracted and the lights were turned off. For the first two sessions, a continuous reinforcement schedule was used (training session with a fixed ratio 1) and for the four following sessions, a variable 30 s interval schedule was used.

Outcome devaluation. After instrumental training, rats received six devaluation sessions. These sessions took place in the operant box and proceeded as a magazine-training session so that the rats could get access to and eat 40 food pellets. At the end of the session, rats were taken out of their boxes and received an intraperitoneal injection before being replaced in their homecages in the vivarium. Animals were divided into two different groups: a devalued group for which LiCl (0.15 M, 10 mL/kg) was injected to induce aversion to the food and a non-devalued group into which NaCl vehicle (0.9%, 10 mL/kg) was injected. After completion of the devaluation phase, rats were further divided into four different groups: a muscimol-devalued (M-D) group, a muscimol-non-devalued (M-ND) group, a vehicle-devalued (V-D) group and a vehicle-non-devalued (V-ND) group.

Microinfusion. Starting 4 days before the first microinfusion, rats were brought each day to the microinfusion room where they were familiarized with being wrapped in a cloth. The day before microinfusion, stylets were removed and cannulas were cleaned using an 8-mm-long dental nerve broach (Micro-Mega, France).

On the day of injection, intracerebral microinfusions were made bilaterally into the PL. The rats were wrapped and gently handled, in order to insert stainless steel injection cannulas (39 gauge and 9.5 mm long; Le Guellec) into the guides. These cannulas were connected through catheters to two Hamilton syringes placed on a dual-syringe infusion pump (2.2 model; Harvard Apparatus). Microinfusions were made at the rate of 0.2 µL/min. Depending on the group, the rats received either 0.25 µL of a GABA_A receptor agonist, muscimol hydrobromide (Sigma-Aldrich) diluted at 1 µg/µL in artificial cere-

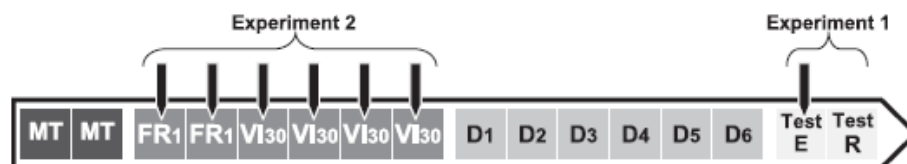


FIG. 1. Schematic representation of the different phases of the behavioural procedure. Black arrows indicate the time of muscimol infusions before the extinction test during Experiment 1 or before each training session during Experiment 2. MT, magazine-training session; FR1, training session with a fixed ratio 1; VI30, training session with a variable interval of 30 s; D1–D6, devaluation sessions; Test E, test in extinction; Test R, reacquisition test.

brospinal fluid (CMA Microdialysis, USA), or 0.25 μ L of artificial cerebrospinal fluid. The injection cannulas were maintained in place for 2 min after microinfusion to allow diffusion of the solution.

Behavioural testing. After the microinfusion, the rats were replaced in their homecages for 5 min before being placed in the operant box and tested. First, they were submitted to an operant lever-press test in extinction. During the 15 min session, the house light was illuminated and the lever was inserted but lever presses were not rewarded. The test session ended with retraction of the lever and turning off the light. Rats were replaced in their homecages in the vivarium. Finally, animals received a rewarded reacquisition test at 1 h after the extinction test. This session was identical to a variable 30 s interval schedule training session.

Experiment 2: prelimbic cortex inactivation during training

Subjects, behavioural apparatus and procedure

For Experiment 2, 48 experimentally naive, male Long-Evans rats (310–400 g) were used. The surgical procedure, behavioural apparatus, schedules of behavioural sessions and food restriction procedure were the same as described for Experiment 1. Experiment 2 only differed in the timing of PL inactivation. Indeed, microinfusions took place during the phase of response acquisition (see Fig. 1). Thus, after each magazine-training session, subjects were brought daily to the microinfusion room and familiarized with being wrapped. On the last day, stylets were removed and guide cannulas were cleaned. Rats were divided into four groups: a M-D group, M-ND group, V-D group and V-ND group. Before each session of the instrumental training rats were microinfused, as in Experiment 1, and waited for 5 min in their home cage before being placed in the operant box. The next phases of outcome devaluation and behavioural testing proceeded without microinfusion.

Histology

After behavioural testing, animals from Experiments 1 and 2 received a lethal dose of sodium pentobarbital (Ceva Santé Animale, France) and were perfused transcardially with 60 mL of a 0.9% NaCl solution, followed by 120 mL of a 10% formaldehyde solution. Brains were

stored in a 30% sucrose/formalin solution for 72 h before being cut in 60 μ m sections with a freezing microtome (Leica SM2400). After being collected onto gelatin-coated slides, brain sections were left to dry for 48 h and were finally stained with thionin. Infusion sites were located under the microscope by D.A.S.T.-Y. and E.C. who were blind to the experimental conditions. Section reconstructions were drawn with reference to the atlas of Paxinos & Watson (1998).

Data analysis

For each session, performance was evaluated as the rate of lever presses per minute. A consumption index was calculated throughout the devaluation sessions, which corresponded to the amount of consumed pellets as a proportion of total delivered pellets. Statistical analyses were performed using ANOVAs, complemented by a Student–Newman–Keuls (SNK) *post hoc* test. In all tests, the alpha risk for rejection of the null hypothesis was maintained at 0.05.

Results

Histology

Figure 2A and B provides a schematic representation of the infusion sites within the PL in Experiments 1 and 2, respectively. Most of the injection sites were located just below the cingulate cortex, allowing a large diffusion into the PL. The final group sizes in Experiment 1 were as follows: V-ND group, $n = 7$; V-D group, $n = 7$; M-ND group, $n = 6$; M-D group, $n = 7$. The final group sizes in Experiment 2 were as follows: V-ND group, $n = 11$; V-D group, $n = 8$; M-ND group, $n = 9$; M-D group, $n = 9$.

Experiment 1

Instrumental training

All of the rats readily consumed the novel food pellets and succeeded in acquiring the instrumental task, the rate of lever presses increasing throughout training for all four groups. A two-way ANOVA, conducted using group and session as factors, showed a session effect ($F_{5,115} = 109.423$, $P < 0.001$) and indicated that the four planned

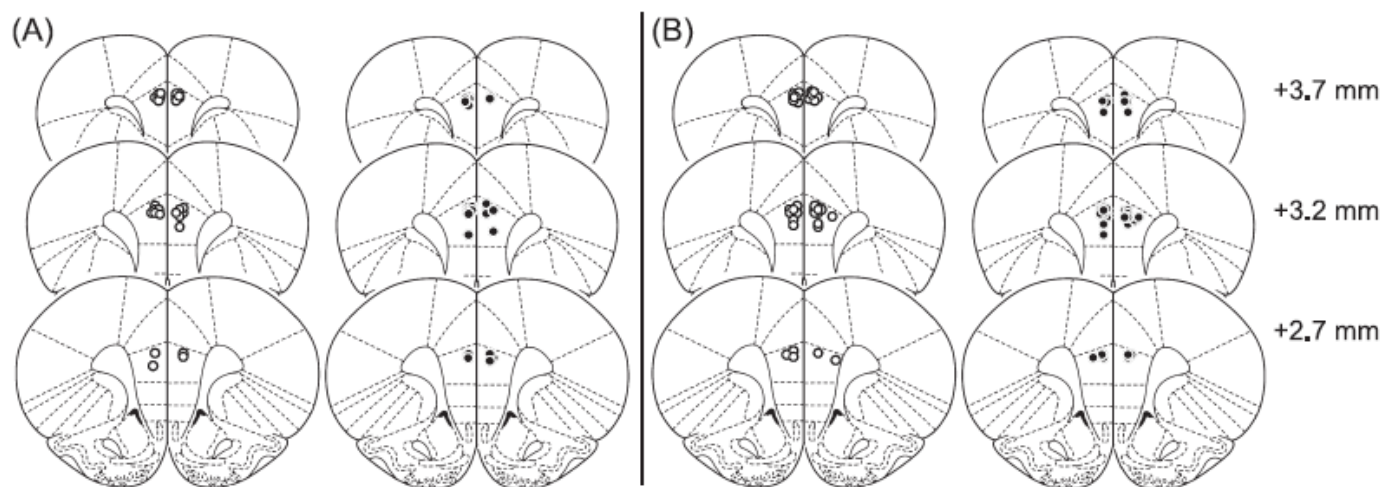


FIG. 2. Schematic representation of infusion sites within the prelimbic region of the mPFC in Experiment 1 (A) and Experiment 2 (B). Open circles, vehicle groups; filled circles, muscimol groups. Outlines are reproduced from Paxinos & Watson (1998) and represent sections ranging from 2.7 to 3.7 mm anterior to Bregma. See text for further details.

groups (V-ND, V-D, M-ND and M-D) did not differ ($F_{3,23} = 0.737$, ns), with no interaction between group and session factors ($F_{15,115} = 0.585$, ns) (Fig. 3A).

Outcome devaluation

The consumption indices of all groups throughout the devaluation sessions are shown in Fig. 3B. The control rats injected with saline (non-devalued) always consumed all of the food pellets so that their indices were 1. By contrast, the consumption indices for subjects treated with LiCl decreased throughout the sessions. This result indicates that animals of both devalued groups acquired a conditioned taste aversion to the food pellet. A two-way ANOVA, with devaluation and session as factors, revealed an effect of devaluation ($F_{1,25} = 49.761$, $P < 0.001$), an effect of session ($F_{5,125} = 46.371$, $P < 0.001$) and a significant interaction between the two factors ($F_{5,125} = 46.346$, $P < 0.001$). Further analysis using the SNK test showed that the planned M-D and V-D groups were homogeneous as they acquired the food aversion at the same rate and reached the same final low level of consumption.

Operant test in extinction

As shown in Fig. 4A, lever-press activity was lower in the V-D group in comparison to the V-ND group during the extinction test. This result indicates that devaluation of the food outcome reduced the instrumental response, thereby attesting that the training procedure resulted

in a goal-directed action. Regarding muscimol-treated groups, M-D animals showed a lower performance than M-ND animals, indicating that the instrumental response was still sensitive to outcome devaluation, and thus goal-directed. The ANOVA attested the main effect of devaluation ($F_{1,23} = 9.979$, $P < 0.01$), no treatment effect ($F_{1,23} = 2.781$, ns) and the absence of interaction between these two factors ($F_{1,23} = 0.441$, ns).

Reacquisition session

During the reacquisition session, the V-D group showed a low instrumental performance in comparison to the V-ND group (Fig. 4B). This pattern of results was the same for muscimol-treated animals. This low rate of lever pressing in both devalued groups confirmed the existence of a conditioned aversion to the reward. A two-way ANOVA, conducted with treatment and devaluation as factors, revealed an absence of effect of the treatment ($F_{1,23} = 0.227$, ns), a significant devaluation effect ($F_{1,23} = 78.702$, $P < 0.001$) and no interaction between the two factors ($F_{1,23} = 0.092$, ns).

Experiment 2

Instrumental training

All of the rats that received microinfusions produced a lower instrumental performance than the subjects of Experiment 1. Nevertheless, the rats obtained the maximum of 40 rewards throughout

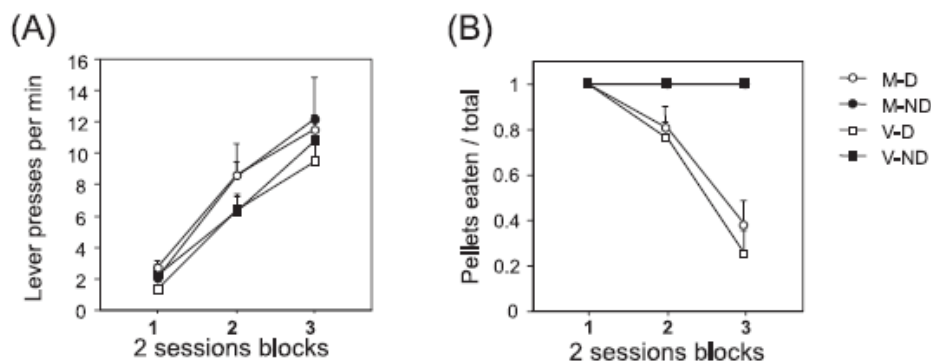


FIG. 3. Experiment 1. (A) Levels of lever pressing during training (mean + SEM; two sessions blocks). Note that the four planned groups acquired the instrumental response at the same rate. (B) Indices of consumption during outcome devaluation (mean + SEM; two sessions blocks). Whereas the indices for non-devalued rats remained equal to 1, the indices for devalued rats decreased throughout the sessions, indicating acquisition of conditioned food aversion to the pellet. M-D, open circle; M-ND, filled circle; V-D, open square; V-ND, filled square.

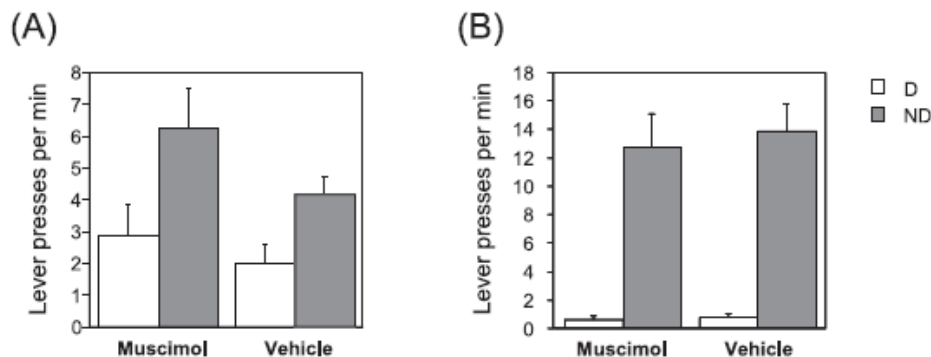


FIG. 4. Experiment 1. (A) Levels of lever pressing during the extinction test (mean + SEM). Sensitivity to outcome devaluation in vehicle control groups was also present in muscimol groups, indicating that the response was goal-directed. (B) Levels of lever pressing during the reacquisition test (mean + SEM). Both devalued groups showed a low performance when the reward was delivered attesting their strong conditioned food aversion to the pellet. D, devalued (white bars); ND, non-devalued (grey bars).

training, indicating that they had acquired the instrumental task. As shown in Fig. 5A, subjects microinfused with muscimol in the PL were able to learn the task at the same rate as rats microinfused with vehicle. Indeed, a two-way ANOVA, with treatment and session as factors, revealed a session effect ($F_{5,175} = 57.279$, $P < 0.001$) but no effect of treatment ($F_{1,35} = 1.628$, ns) and no interaction between the two factors ($F_{5,175} = 0.310$, ns). A further ANOVA with devaluation to be performed as an additional factor ensured that there was no group effect ($F_{3,33} = 0.579$, ns).

Outcome devaluation

As in Experiment 1, subjects in non-devalued groups showed stable consumption indices, whereas those of devalued groups decreased throughout sessions. Thus, independent of treatment, animals in both devalued groups acquired a conditioned taste aversion to the food pellet (Fig. 5B). A three-way ANOVA, with treatment and devaluation as between-subject factors and session as a repeated measure, revealed significant effects of the devaluation and session factors ($F_{1,33} = 99.516$, $P < 0.001$ and $F_{5,165} = 73.345$, $P < 0.001$, respectively) and a significant session $\times$ devaluation interaction ($F_{5,165} = 73.679$, $P < 0.001$). As regards the effects of muscimol, there was no treatment effect ($F_{1,33} = 2.669$, ns) but a significant session $\times$ treatment interaction ($F_{5,165} = 2.997$, $P < 0.05$) and a significant session $\times$ treatment $\times$ devaluation interaction ($F_{5,165} = 3.049$, $P < 0.05$). Thus, the treatment with muscimol could have altered the acquisition of aversion

to the food pellets. The SNK test indeed showed that the consumption index of the M-D group was significantly higher than that of the V-D group. However, this may reflect the fact that the V-D group in this experiment acquired the aversion at a slightly faster rate compared with the devalued group of Experiment 1.

Operant test in extinction

Results of the extinction test are shown in Fig. 6A. As in Experiment 1, animals of the V-D group pressed less on the lever than non-devalued rats, showing that vehicle-treated subjects were sensitive to outcome devaluation and that the response produced was goal-directed. However, rats in the M-D group pressed on the lever as much as rats in the M-ND groups. Thus, animals for which PL had been inactivated during training were not sensitive to outcome devaluation, indicating that the acquired response was not goal-directed. A two-way ANOVA, conducted with treatment and devaluation as factors, revealed an effect of treatment ($F_{1,33} = 4.382$, $P < 0.05$), of devaluation ($F_{1,33} = 6.09$, $P < 0.05$) and a trend towards an interaction between the two factors ($F_{1,33} = 3.943$, $P = 0.0554$). Further analysis using the SNK test attested that instrumental performances of V-D and V-ND groups significantly differed. By contrast, M-ND and M-D groups did not differ. Finally, the SNK test showed that performances in the V-ND, M-ND and also in the M-D groups were significantly higher than in the V-D group.

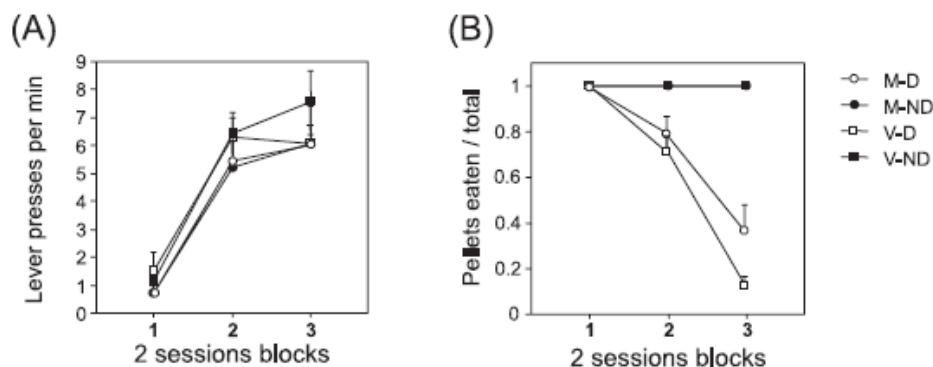


FIG. 5. Experiment 2. (A) Levels of lever pressing during training (mean + SEM; two sessions blocks). Animals injected with vehicle or muscimol acquired the instrumental response at the same rate. (B) Indices of consumption during outcome devaluation (mean + SEM; two sessions blocks). Microinfusion had no effect on the pattern of results obtained for both the vehicle and muscimol groups; the indices for non-devalued rats remained equal to 1 and the indices for devalued rats decreased throughout the sessions. M-D, open circle; M-ND, filled circle; V-D, open square; V-ND, filled square.

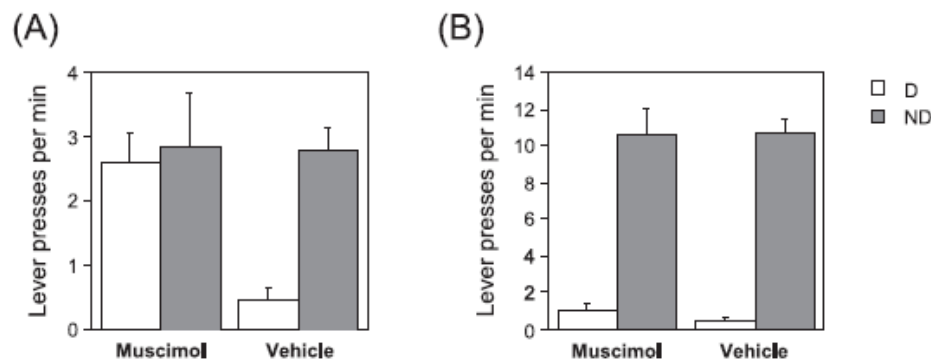


FIG. 6. Experiment 2. (A) Levels of lever pressing during the extinction test (mean + SEM). Sensitivity to outcome devaluation in vehicle control groups was absent in muscimol groups, showing that the response of these animals was not goal-directed. (B) Levels of lever pressing during the reacquisition test (mean + SEM). Both devalued groups showed a low performance when the reward was delivered attesting their strong conditioned food aversion to the pellet. D, devalued (white bars); ND, non-devalued (grey bars).

Reacquisition session

The results obtained in the rewarded conditions singularly differed from those obtained in the non-rewarded conditions (Fig. 6B). Indeed, if animals of the V-D group still showed a low instrumental performance, subjects of the M-D group acted in the same way and stopped pressing on the lever when the pellet reward was delivered. These results attest that both vehicle- and muscimol-treated animals had a strong conditioned aversion to the reward. A two-way ANOVA, conducted with treatment and devaluation as factors, confirmed the absence of effect of treatment ($F_{1,33} = 0.75$, ns), an effect of devaluation ($F_{1,33} = 133.566$, $P < 0.001$) and the absence of interaction between the two factors ($F_{1,33} = 0.165$, ns).

Discussion

The present study made use of local reversible inactivation to investigate the time-dependent role of the PL in the control of goal-directed behaviour. The results demonstrate that normal functioning of the PL is not required at the time of testing for the expression of goal-directed behaviour. In contrast, PL inactivation during training prevented goal-directed responding at test, as response was rendered insensitive to outcome devaluation. These results have important implications that are discussed below.

Methodological issues

The finding that PL inactivation throughout instrumental training alters goal-directed responding is in agreement with various data using pretraining PL lesions (Corbit & Balleine, 2003; Killcross & Coutureau, 2003; Ostlund & Balleine, 2005). Neuronal activity within the prelimbic region was silenced using muscimol-reversible inactivation that has a narrower time window (~8 h) and allows less time for lesion-induced plasticity (for review see Majchrzak & Di Scala, 2000). It could be argued that the multiple injection procedure of Experiment 2 might have induced tissue damage (e.g. gliosis). However, it is unlikely that permanent damage to this region could explain our results because control animals received exactly the same amount of injections. We observed a general decrease in the rate of responding of both muscimol- and artificial cerebrospinal fluid-injected animals, which might be due to repetitive contention. With microinfusions, we cannot completely exclude that some of the injected volume actually spread toward the dorsally located cingulate cortex or the ventrally located infralimbic cortex (Edeline *et al.*, 2002) but we selected a dose (1 $\mu\text{g}/\mu\text{L}$) and a volume (0.25 μL) that should have limited such a spread. In addition, we purposely implanted the guide cannulas quite dorsally in order to prevent diffusion to the infralimbic cortex.

In the present study, we used food aversion rather than sensory-specific satiety in order to devalue the outcome. Devaluation could therefore be performed separately from the testing phase so that any effect of muscimol on devaluation itself was excluded. In addition, in order to rule out a putative impact of contextual control of devaluation (Colwill & Rescorla, 1985; Boakes *et al.*, 1997; Holland, 2004; Nelson & Killcross, 2006), the devaluation took place in the same operant box as during instrumental training. The LiCl concentration chosen (0.15 M) allowed for session by session monitoring of the acquisition of conditioned aversion. This aversion was strong and long-lasting, as all animals refrained from lever pressing in the final rewarded test. This test, unlike the extinction test, did not require any representation of the consequences of lever pressing but relied upon the feedback provided by the aversive outcome to stop pressing the lever. It is worth noting that such a

feedback function does not require an intact PL (see Experiment 1 and Corbit & Balleine, 2003).

Our results, unlike those of Ostlund & Balleine (2005), were obtained in a procedure where rats were trained and tested with a single lever. This may be important as previous research indicates that habit responding cannot be obtained using a multiple responses and reinforcers procedure (Colwill & Rescorla, 1985; Holland, 2004). The single-lever situation moreover fits into a framework of single-response systems that has been extensively examined and modelled (e.g. Daw *et al.*, 2005).

Prefrontal control of goal-directed action

In primates, including humans, and also in rats, the prefrontal cortex contributes to executive functions such as the planning of action and decision making (Miller & Cohen, 2001; Chudasama & Robbins, 2006; Rich & Shapiro, 2007). A number of recent studies aimed at clarifying the implication of the prefrontal cortex in operant conditioning. The present data show that the acquisition of a representation of A-O necessitates the activity of the PL, whereas the expression of such a representation does not depend on this activity. This result therefore extends previous research that used massive excitotoxic lesions of the whole mPFC including the PL, infralimbic cortex and some of the anterior cingulate cortex (Ostlund & Balleine, 2005). Using microinfusions allowed us to better control for long-term effects of lesions and to use similar schedules of training and testing across experiments. The PL region, which had been previously identified as crucial for goal-directed responding (Killcross & Coutureau, 2003), was the primary target of our microinjections. It may therefore be safely concluded that the PL is essential during encoding of the A-O associations.

The A-O representation probably involves the association of a number of elements that include the instrumental response and the outcome as a goal (see Rescorla, 1991). Our results emphasize the importance of PL activity to bind these elements into a comprehensive representation. Pretraining inactivation (Experiment 2) or lesion of the PL (Corbit & Balleine, 2003; Killcross & Coutureau, 2003) leads to an instrumental responding that is insensitive to outcome devaluation, suggesting that the association built into this condition is a different (possibly simpler) S-R habit association. Such a pattern of results contrasts with available data concerning the dorsomedial striatum. Either pretraining lesions of the dorsomedial striatum (Yin *et al.*, 2005b) or blockade of the *N*-methyl-D-aspartate receptors (Yin *et al.*, 2005a) within the posterior dorso medial striatum (pDMS) alter the sensitivity of the rats to outcome devaluation. Moreover, posttraining lesions or inactivations of this region also alter instrumental sensitivity to devaluation, suggesting that the pDMS could be a locus of expression of the A-O association (Yin *et al.*, 2005b). Given the dissociation between the pDMS and PL concerning their time-dependent contribution to A-O learning, it is likely that early in training the PL and pDMS functionally interact (Hernandez *et al.*, 2006) through glutamatergic input from the PL to the pDMS (Alexander *et al.*, 1986; Vertes, 2004; Gabbott *et al.*, 2005) and plasticity through an *N*-methyl-D-aspartate receptor-dependent mechanism (Costa, 2007; Wanjerkhede & Bapi, 2008).

The integration of A-O and S-R responding probably involves parallel corticostriatal loops (Balleine, 2005). A PL/dorsomedial striatum loop might account for A-O responding, whereas an infralimbic/lateral striatal circuit appears to be involved in the control of S-R habits (Coutureau & Killcross, 2003; Yin *et al.*, 2004, 2006; Faure *et al.*, 2005). The control of instrumental responding by A-O vs. S-R systems is thought to depend on the perceived

- Nelson, A. & Killcross, S. (2006) Amphetamine exposure enhances habit formation. *J. Neurosci.*, **26**, 3805–3812.
- Nordquist, R.E., Voorn, P., de Mooij-van Malsen, J.G., Joosten, R.N., Pennartz, C.M. & Vanderschuren, L.J. (2007) Augmented reinforcer value and accelerated habit formation after repeated amphetamine treatment. *Eur. Neuropsychopharmacol.*, **17**, 532–540.
- Ostlund, S.B. & Balleine, B.W. (2005) Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J. Neurosci.*, **25**, 7763–7770.
- Paxinos, G. & Watson, C. (1998). *The Rat Brain in Stereotaxic Coordinates*. Academic Press, San Diego, CA.
- Rescorla, R.A. (1991) Associative relations in instrumental learning: the eighteenth Bartlett memorial lecture. *Q. J. Exp. Psychol.*, **43B**, 1–23.
- Rich, E.L. & Shapiro, M.L. (2007) Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *J. Neurosci.*, **27**, 4747–4755.
- Simmons, J.M., Ravel, S., Shidara, M. & Richmond, B.J. (2007) A comparison of reward-contingent neuronal activity in monkey orbitofrontal cortex and ventral striatum: guiding actions toward rewards. *Ann. NY Acad. Sci.*, **1121**, 376–394.
- Tanaka, S.C., Balleine, B.W. & O'Doherty, J.P. (2008) Calculating consequences: brain systems that encode the causal effects of actions. *J. Neurosci.*, **28**, 6750–6755.
- Valentin, V.V., Dickinson, A. & O'Doherty, J.P. (2007) Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.*, **27**, 4019–4026.
- Vertes, R.P. (2004) Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse*, **51**, 32–58.
- Wanjerkhede, S.M. & Bapi, R.S. (2008) Modeling the sub-cellular signaling pathways involved in reinforcement learning at the striatum. *Prog. Brain Res.*, **168**, 193–206.
- Wickens, J.R., Horvitz, J.C., Costa, R.M. & Killcross, S. (2007) Dopaminergic mechanisms in actions and habits. *J. Neurosci.*, **27**, 8181–8183.
- Yin, H.H. & Knowlton, B.J. (2006) The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.*, **7**, 464–476.
- Yin, H.H., Knowlton, B.J. & Balleine, B.W. (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.*, **19**, 181–189.
- Yin, H.H., Knowlton, B.J. & Balleine, B.W. (2005a) Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.*, **22**, 505–512.
- Yin, H.H., Ostlund, S.B., Knowlton, B.J. & Balleine, B.W. (2005b) The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.*, **22**, 513–523.
- Yin, H.H., Knowlton, B.J. & Balleine, B.W. (2006) Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav. Brain Res.*, **166**, 189–196.