



Influence de la familiarité dans la catégorisation rapide de scènes naturelles

Jonathan Hindmarsh

Université de Lille 3 Charles de Gaulle



Université Lille Nord de France

| Pôle de Recherche
et d'Enseignement Supérieur

Sommaire

Résumé.....	4
Abstract	5
Introduction.....	6
Introduction théorique.....	6
Qu'est-ce que la reconnaissance ?	6
Comment mesurer la reconnaissance ?	7
Vitesse de la reconnaissance	8
Niveau super-ordonné et niveau de base	10
Problématique	11
Objectifs.....	11
Hypothèses	12
Méthode	13
Participants.....	13
Critères d'inclusion et d'exclusion	13
Justification du nombre de participants	13
Lieu d'étude	14
Stimuli.....	14
Procédure	16
Pré sélection	16
Choix saccadique	17
Réponse manuelle	18

Contre-balancement	20
Résultats	21
Tâche de dénomination (étude pilote).....	21
Expérience 1 : saccade oculaire	23
Précision.....	23
Temps de réponse.....	24
Expérience 2 : réponse manuelle	26
Précision.....	26
Temps de réaction	27
Comparaison choix saccadique et réponse manuelle.....	28
Discussion	29
Limites et ouvertures.....	33
Conclusion	35
Références	36
Annexes.....	39
Annexe I.....	39
Annexe II	40

Résumé

Il a été montré dans différentes études que le système visuel est capable de reconnaître les visages ou les animaux très rapidement. Mais pour les animaux, le facteur de familiarité n'a pas été testé. Nous avons cherché ici à savoir s'il y avait une différence dans les processus de reconnaissance pour les animaux familiers et les animaux rares. Nous avons préalablement constitué une base constituée de 36 images d'animaux familiers et d'animaux non-familiers grâce à une étude pilote de dénomination. Nous avons ensuite présenté les images sélectionnées à 20 participants grâce à un paradigme de choix saccadique pour obtenir les temps de réponse les plus bas possible, et une tâche de réponse manuelle. Les participants avaient pour but de trouver l'image contenant un animal en effectuant une saccade oculaire, ou en glissant leur doigt sur la cible sur un écran tactile. Les résultats montrent qu'il n'y avait pas de différence significative dans les temps de réaction entre les animaux familiers et les animaux rares lorsque les temps de réaction étaient inférieurs à 200ms, mais que dans la tranche des 240-300 ms il y avait un effet de la familiarité. Ceci tend à corroborer que l'accès à la familiarité impliquerait des processus cognitifs plus tardif. Avant 200 ms, seules les caractéristiques physiques semblent être prises en compte pour déterminer la présence d'un animal dans une scène visuelle.

Mots clés : système visuel, familiarité, saccades oculaires, catégorisation

Abstract

It has been shown in several studies that the visual system is able to identify faces and animals very quickly. However, for animals as targets, the familiarity factor has never been taken into account. We tried to find out if there is a difference in categorizing familiar well known animals as compared to less common animals. We first selected 36 pictures depicting a very common animal and 36 pictures depicting a rare animal with a naming task. We then presented these images to 20 participants with a saccadic choice task in order to get the shortest possible reaction times, and a manual response task. The participants were asked to saccade, or to move their finger, on a touch screen, toward the scene containing an animal. The results show that there is no significant difference between the reaction times for the familiar animals compared to rare animals when response times are lower than 200ms. However, accuracy was higher for familiar animals between the 240ms and 300ms, which tends to support the idea that familiarity would be a late cognitive process. Before 200ms, only the physical features seem to be taken into account in order to detect the presence of an animal in a visual scene.

Key words: visual system, familiarity, saccadic choice task, categorization

Introduction

Introduction théorique

Un des aspects les plus étudiés de la perception est son déroulement temporel. Toutes les études qui s'intéressent à l'un de nos sens rapportent toujours des informations concernant des temps de traitement ou de réaction, que ce soit la vision, l'audition, ou même le toucher. Le cerveau est en mesure de traiter et de comprendre rapidement les informations afférentes des différents récepteurs qui composent notre système perceptif, ce qui lui permet de réagir le plus rapidement possible. Cette rapidité de traitement est évidemment importante pour donner un sens à notre environnement, et pour adapter notre comportement à celui-ci. Ainsi, lorsqu'un danger se présente, plus nous sommes rapides à le comprendre et à réagir, meilleures seront les chances de survie.

La vision nous permet, de manière générale, d'obtenir le plus d'informations sur notre environnement actuel. C'est principalement grâce au système visuel que l'on va être capable de générer une représentation mentale en 3D de notre environnement, et de pouvoir en former des souvenirs pour se rappeler de la scène à un moment ultérieur.

Qu'est-ce que la reconnaissance ?

La reconnaissance est le processus cognitif qui nous permet de comprendre ce qui est visible dans notre champ visuel. Il s'agit d'utiliser les informations du système visuel pour savoir si l'objet qui nous est présenté est connu, ou nouveau, et éventuellement de lui associer une étiquette, un nom.

Le système visuel est capable d'extraire très rapidement l'information d'une scène visuelle afin de reconnaître les objets présents, ce traitement visuel pouvant être effectué en moins de 150 ms (Potter, 1976; Thorpe, Fize, & Marlot, 1996). Différentes structures neurophysiologiques sont impliquées pour permettre cette reconnaissance. Les informations visuelles partent de la rétine, et arrivent jusqu'au cortex visuel. Les informations liées aux formes et aux catégories d'objets sont souvent associées avec des traitements dans la voie ventrale du cortex visuel (alors que les

informations liées à la localisation de ces objets feraient plutôt l'objet d'un traitement de la voie dorsale).

Grâce à des études neuropsychologiques de lésions chez des patients apraxiques et ataxiques, Goodale et Milner (1992) ont développé un modèle séparant la perception pour l'identification, donc la reconnaissance, et la perception pour l'action, grâce à des transformations sensorimotrices. La zone cérébrale importante chez l'homme pour l'identification des objets comporterait donc les projections partant du cortex strié jusqu'au cortex infero-temporal, ce qui constitue la voie ventrale. Le singe possédant un fonctionnement cérébral relativement proche de celui de l'homme, Hubel et Wiesel (1968) ont cherché à montrer comment fonctionnait la reconnaissance chez des singes afin de pouvoir faire des inférences sur le fonctionnement des mêmes capacités cognitives chez l'homme et de mieux déterminer les structures importantes : « La reconnaissance dans le cortex est effectuée par la voie visuelle ventrale, partant du cortex visuel primaire (V1) jusqu'aux aires visuelles extra-striées II (V2) et IV (V4), au cortex infero-temporal (IT), puis jusqu'au cortex préfrontal (PFC), qui est impliqué dans le lien entre la perception et la mémoire et l'action » (traduit de l'anglais). Ces études, effectuées chez l'animal par le moyen d'implantations de sondes dans le cortex, ont été confirmées chez l'homme par des études en neuroimagerie (Ishai et al, 2000 ; Grill-Spetor & Malach, 2004 pour une revue).

Comment mesurer la reconnaissance ?

La vitesse du système visuel peut être mesurée, grâce à différents types de paradigmes. L'un d'entre eux est le protocole de catégorisation go/no-go (Thorpe, Fize et Marlot 1996), qui consiste à présenter des images à un participant, et où la tâche à effectuer est de répondre en appuyant le plus rapidement possible sur un bouton si l'image correspond à une certaine catégorie d'objet donnée à l'avance, par exemple si l'image comporte un animal, ou de ne rien faire dans le cas contraire. Avec ce paradigme, les participants obtenaient en moyenne des temps de réponse de 445 ms de temps de réaction lorsque la cible était effectivement présente dans l'image.

Plus récemment on a utilisé des paradigmes de choix saccadique. Dans ce type de tâche, deux images sont simultanément présentées sur un écran, et le participant a pour but de détecter une certaine cible (par exemple un visage ou un animal), et doit répondre en effectuant une saccade

oculaire vers sa cible. Thorpe et Kirchner (2005) ont proposé ce type de paradigme à la suite des études utilisant le paradigme de go/no-go, afin d'éliminer le temps de traitement du programme moteur de la réponse, puisque les mesures en go/no-go représentent à la fois le temps du traitement visuel mais aussi l'exécution motrice de la réponse (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001 ; Thorpe et al., 1996). Selon plusieurs études, des saccades oculaires peuvent être déclenchées en un laps de temps relativement court de l'ordre de 80 à 100 ms (Busetini, Masson & Miles, 1997 ; Fischer & Weber, 1993 ; Masson, Rybarczyk, Castet, & Mestre 2000), et ces saccades sont d'ailleurs considérées comme porteuses d'informations de valeur concernant la coordination de la vision et des mouvements oculaires, ainsi que de l'attention visuelle.

L'intérêt de ces tâches par rapport à des tâches de dénomination, est qu'elles sont beaucoup plus rapides. Dans ces tâches de catégorisation rapide, il n'y a pas besoin de faire le lien vision-mot, ce qui nécessiterait un traitement supplémentaire, et donc un temps de réponse plus long ne reflétant pas uniquement le temps de reconnaissance du stimulus. Retirer la dénomination réduit donc la nécessité d'un accès lexical. Cette idée est d'ailleurs étayée par des expériences chez l'animal (Fabre-Thorpe, Richard et Thorpe, 1998; Fabre-Thorpe 2003). Les singes sont capables de réaliser ces mêmes tâches avec un haut taux de précision (90%), et plus rapidement que l'Homme.

Vitesse de la reconnaissance

Ces différents paradigmes expérimentaux ont démontré que certains stimuli pouvaient être détectés plus rapidement que d'autres, notamment les visages qui peuvent donner lieu à des temps de réaction aussi bas que 100-110 ms, et les animaux à 120-130 ms, avec des moyennes de l'ordre de 140 ms seulement (Kirchner & Thorpe, 2005 ; Crouzet, Kirchner, Thorpe, 2010 ; Boucart et al in press). De nombreuses études montrent que le traitement des visages peut être effectué très rapidement, et pourrait d'ailleurs même faire l'objet d'un certain avantage comportemental, en capturant l'attention (Langton, Law, Burton, & Schweinberger, 2008; Ro, Russell, & Lavie, 2001; Theeuwes & Van der Stigchel, 2006; Vuilleumier, 2000) : la présence d'un visage parmi les distracteurs engendre des temps de réaction plus longs, car l'attention est capturée par ce visage même s'il n'a pas d'utilité pour la tâche en cours, reflétant donc un processus automatique et

involontaire. En utilisant le paradigme de choix saccadique, Crouzet, Kirchner et Thorpe ont remarqué qu'en présentant simultanément une image de véhicule et de visage, les participants avaient tendances à effectuer des saccades vers les visages, alors que les consignes étaient de regarder vers les véhicules. Ce biais pour les visages semble difficile à éliminer.

Deux explications ont été apportées pour ce phénomène. La première est, selon New, Comsides et Tooby (2007) mais aussi Crouzet, Kirchner et Thorpe (2010), que le cerveau serait doté de mécanismes spécifiques pour traiter les informations biologiquement importantes, concernant donc les visages mais aussi les animaux. Comme il est généralement admis qu'il faut environ 20 ms pour que les structures du tronc cérébral (colliculus supérieur) impliquées dans le contrôle oculomoteur et les muscles de l'œil eux-mêmes s'activent, les mécanismes sous tendant la reconnaissance des visages pourraient s'activer seulement 80 ms après l'apparition du stimulus. Cependant, un temps aussi court ne permettrait pas au système de la voie ventrale d'intervenir, qui sert normalement à la reconnaissance des objets présents dans la scène visuelle. On sait que toutes les fibres nerveuses optiques ne se projettent pas dans le cortex visuel, et certaines structures sous-corticales telles que l'amygdale peuvent traiter certaines informations visuelles (Johnson, 2005), notamment concernant les stimuli à valence émotionnelle ou pouvant engendrer des réactions de peur (Ohman, Carlsson, Lundqvist, & Ingvarn 2007). Il n'est donc pas impossible d'expliquer ces temps de réaction par rapport aux visages avec l'intervention de structures sous-corticales, qui sont moins plastiques que le cortex et supporteraient une hypothèse d'importance au niveau de l'évolution.

La seconde explication pour ces temps de réaction très courts repose sur la manière dont nous utilisons les images qui nous sont présentées. Des travaux sur les premiers traitements de l'image mettent en évidence l'utilisation de statistiques de l'image plus « brutes » (coarse), basées sur des représentations de plus bas niveau (Oliva & Torralba, 2001). Ces représentations ont reçu le nom « d'enveloppe spatiale » (spatial layout). Elles permettent de « résumer la scène », et ont ainsi pour principal intérêt de permettre de regrouper facilement et rapidement des scènes visuelles similaires qui partagent un certain nombre de caractéristiques (comme les rues et les autoroutes). Nous pouvons donc identifier la scène qui nous est présentée dans son ensemble avant de voir clairement les détails, la notion de globalité est donc importante. Oliva et Torralba (2003) montrent également comment ces statistiques de l'image peuvent nous aider à prédire la présence ou

l'absence d'un objet dans une scène avant même d'explorer l'image, ce qui implique donc des temps de reconnaissance relativement bas. Il est donc possible d'imaginer l'intervention de ce genre de traitement pour la reconnaissance de la présence d'un visage ou d'un animal.

La catégorisation rapide de scènes ou d'objets pourrait donc reposer sur un traitement des caractéristiques de bas niveau, c'est-à-dire des informations traitées précocement par le système visuel telles que les statistiques globales de l'image (Greene et Oliva 2009a ; 2009b). Ces caractéristiques globales ont été définies comme étant les couleurs, les orientations, ou les fréquences spatiales (Oliva & Schyns 2000 ; Oliva 2005 ; Gaspar & Rousselet 2009 ; Fabre-Thorpe 2011 ; Kauffmann et al 2014). Certaines catégories de stimulus partagent certaines de ces caractéristiques globales. Par exemple, les arbres partagent relativement tous les mêmes palettes de couleur (tronc, feuillages), et possèdent plus ou moins la même orientation (tronc vertical).

Niveau super-ordonné et niveau de base

Ces études sur les caractéristiques globales des images nous amènent à une notion importante de la reconnaissance : les niveaux de catégorisation. Selon Oliva (2005), la description d'une scène de notre environnement peut se faire à trois niveaux différents :

- Le niveau super-ordonné : il correspond au niveau d'abstraction le plus élevé, et fait usage d'informations relativement simple, par exemple pour faire la distinction entre une scène naturelle ou un environnement urbain.
- Le niveau de base : il correspond au niveau de représentation catégorique le plus commun, par exemple « forêt », « plage », « montagne », « rue » ... Les éléments d'une même catégorie dans ce niveau partagent des caractéristiques communes, comme une forme ou disposition semblable, mais aussi partagent la même fonction.
- Le niveau sub-ordonné : ce niveau exige de faire une analyse plus poussée des différentes structures composant la scène, ainsi que la distinction des différentes régions d'une image et des objets, « ma rue ».

On peut voir que ces trois niveaux de description concernent différents niveaux d'abstraction, et ainsi différentes informations sémantiques. Tout n'est pas pris en compte dans les

niveaux les plus élevés, ce qui implique donc un traitement plus léger. Il semble donc important de prendre en compte ces différents niveaux de catégorisation lors d'études sur la reconnaissance.

Ce principe de niveaux de catégorisation s'applique également aux animaux. Ainsi, au niveau super-ordonné, on va avoir « animal », puis on pourrait avoir « herbivore », « reptile », « oiseau » au niveau de base, et enfin des noms d'animaux plus précis au niveau subordonné tels que « pigeon », « canard », « mouette », ou « aigle ».

Ceci est en accord avec les résultats obtenus par Thorpe et collaborateurs, qui montre que la catégorisation à un niveau super-ordonné (animal ou non animal) engendre des temps de réponse plus rapides que la catégorisation à un niveau de base (Macé et al, 2009 ; Poncet et al, 2014 ; Wu et al, 2005). Selon ces études, le système visuel est capable de détecter la présence d'un animal avant de savoir de quel type d'animal il s'agit exactement : « on reconnaît l'animal avant de reconnaître l'oiseau ». Ces résultats s'expliquent par le fait que les traitements réalisés ne nécessitent que des informations relativement grossières qui sont véhiculées par les fréquences spatiales basses de l'image.

Problématique

Cependant, dans ces études sur la reconnaissance des animaux, les cibles « animal » étaient tirées au hasard parmi une banque de milliers d'images (banque de données Corel). Ainsi, de nombreux facteurs de plus haut niveau n'ont pas été contrôlés, comme le degré de familiarité, c'est-à-dire si l'animal était très connu ou non. Pour aborder cette question, nous avons utilisé une tâche de choix saccadique (Crouzet, Thorpe, Fabre-Thorpe 2012), qui nous permet d'étudier le traitement des scènes naturelles dans une fenêtre temporelle très courte.

Objectifs

Notre objectif est de tester l'influence de certaines caractéristiques sémantiques, telles que le degré de familiarité, grâce à un paradigme de choix saccadique. Nous voulions savoir si la familiarité peut avoir une influence sur les processus les plus précoces de la détection, ce qui

pourrait mettre en évidence l'existence de certaines connexions ou mécanismes de facilitation pour la reconnaissance des stimuli familiers et ainsi optimiser ce traitement.

Pour opérationnaliser le degré de familiarité dans notre étude, nous avons au préalable effectué une étude pilote de sélection d'images, dans laquelle nous avons pris en compte la rapidité du temps de dénomination et l'exactitude de la réponse.

Hypothèses

Nous supposons que comme la familiarité fait l'objet d'un traitement sémantique de plus haut niveau que la simple détection des informations physiques caractérisant la présence d'un animal, elle ne devrait pas avoir d'effet sur catégorisation rapide. Donc, il n'y aurait pas de différence de reconnaissance pour un animal familier par rapport à un animal peu familier ou inconnu. Au niveau des résultats, on s'attend donc à obtenir des temps de réaction équivalents pour les animaux familiers que pour les animaux rares. Afin de tester si le traitement de l'information sémantique nécessite plus de temps nous avons comparé la performance dans une tâche de choix saccadique très rapide à celle d'une tâche de choix avec réponse manuelle plus lente. Les mêmes images et le même paradigme ont été utilisés. Seul le mode de réponse différait. Cette comparaison (réponse saccade versus réponse manuelle) avait donné lieu à des résultats différents sur le traitement de l'information contextuelle chez des patients Alzheimer et des sujets sains appariés en âge (Boucart et al, 2014).

Même si la familiarité devrait partager un lien fort avec la reconnaissance, et donc devrait éventuellement engendrer un effet facilitateur pour les stimuli familiers, des expériences montrent que le système visuel traite le niveau super-ordonné plus rapidement que le niveau de base. Or comme la familiarité résiderait plutôt au niveau sub-ordonné (ex : « c'est mon chien »), on pourrait ainsi s'attendre à ne pas avoir d'effet de facilitation de la familiarité.

Méthode

Participants

Critères d'inclusion et d'exclusion

Afin d'être inclus dans cette étude, les participants devaient remplir différents critères d'inclusion. Ils devaient avoir entre 18 et 40 ans (pour éviter d'éventuels effets du vieillissement), avoir une acuité visuelle normale ou corrigée à la normale grâce à des lentilles de contact (le port de lunette n'était pas compatible avec l'oculomètre), et donner leur consentement libre et éclairé. Les participants étaient des étudiants qui ont été recruté sur la base du volontariat principalement au sein de l'Université Lille 3 Charles-De-Gaulle, par voie d'affichage. Chaque participant a été indemnisé pour sa participation. Le protocole de recherche a reçu l'approbation du comité d'éthique de Lille3 (cf. Annexe I).

Justification du nombre de participants

Pour éviter des effets d'habituation aux images, les participants qui ont fait la tâche de pré-sélection d'images ont ensuite participé à une autre étude (Adriano Guidice) basée sur un choix saccadique portant sur la complexité de la scène visuelle, alors que les participants qui ont fait la pré-sélection de l'étude sur la complexité (Adriano Giudice) ont ensuite passé la tâche de choix saccadique portant sur la familiarité de l'animal.

Comme nous nous sommes intéressés à un facteur de haut niveau qui est le degré de familiarité de l'animal, qui peut tout à fait être subjectif et présenter une certaine variabilité inter-individuelle, nous avons recruté 20 participants pour la tâche de pré-sélection d'images (3 ont été exclus parmi les 23 initialement recrutés), et 15 pour la tâche de choix saccadique et motrice (4 ont été exclus parmi 19). Les participants sont venus deux fois au laboratoire : la première fois pour la pré-sélection d'images, et la deuxième fois pour la tâche de choix saccadique et une tâche de choix avec réponse manuelle.

Lieu d'étude

Les passations se sont déroulées au sein des locaux du SCA-Lab au Pôle recherche de l'Université Lille 2 Droit et Santé.

Stimuli

Les stimuli sélectionnés étaient des photos de scènes naturelles, et ont été choisies depuis internet (Google, Wikipedia Commons). La moitié de ces scènes comportait un animal, appartenant à différentes catégories : mammifère aquatique, oiseau, insecte, herbivore, prédateur, rongeur. Les animaux étaient présentés dans un environnement naturel. L'autre moitié des images comportait des objets variés (bâtiments, véhicules, végétaux, minéraux, instruments de musique, paysages) et servait de distracteurs, car ne contenant pas d'animal, ou d'humain. Chaque stimulus était présenté dans un contexte naturel, c'est à dire que l'objet des photographies n'était pas découpé, le fond était présent.

Chaque image disposait d'une version en couleur et d'une version en nuance de gris. Le facteur principal manipulé était la familiarité des animaux présentés, qui pouvaient être familiers, ou rares voir inconnus. Il y avait 36 images d'animaux familiers et 36 images d'animaux rares, ce qui donne 144 images en comptant les versions en nuances de gris.

La taille de toutes les images a été fixée à 512 pixels de largeur et de hauteur. Des exemples des images utilisées sont présentées dans les figures 1a, 1b, et 1c.



Fig. 1 a : exemple d'animaux familiers. De gauche à droite : dauphin, cheval, coccinelle, canard, tigre, souris.

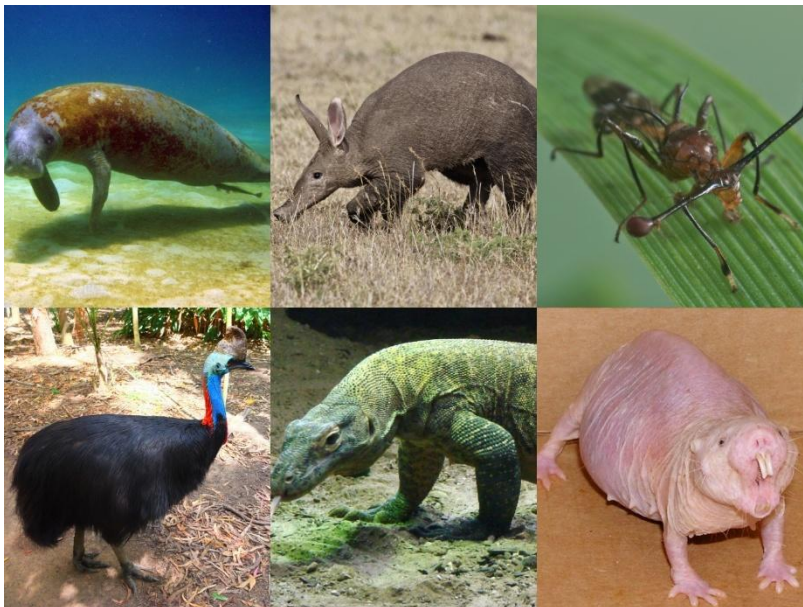


Fig. 1 b : exemple d'animaux rares. De gauche à droite : lamentein, oryctérope, diopsidae, casoar, dragon de komodo, rat-taupe nu.



Fig. 1 c : exemple de distracteurs.

Procédure

Pré sélection

Avant de faire les tâches de choix saccadique et motrice, il fallait vérifier que les images sélectionnées répondaient bien aux critères définis, c'est à dire qu'une image sélectionnée d'un animal familier devait bien correspondre à un animal familier pour les participants. Idem pour les images d'animaux rares. Donc, nous avons préalablement sélectionné un grand nombre d'images, pour pouvoir ne garder que celles qui correspondaient le mieux aux paramètres de l'étude.

Les photos sélectionnées, initialement au nombre de 228, étaient présentées une par une au centre d'un écran d'ordinateur. Pour la sélection des stimuli la tâche du sujet était de dénommer chaque animal, le plus rapidement possible. Nous enregistrions, grâce à un micro, le temps de réponse verbal pour que le participant donne le nom exact de l'animal présenté. La réponse orale du sujet était ensuite entrée au clavier par l'expérimentateur. Ensuite, le sujet devait dire si l'animal présenté lui semblait très connu ou peu connu, sur une échelle de 0 à 9, avec « 9 » pour très connu

(par exemple « vache », « dauphin », « poule »), et « 0 » pour peu connu (par exemple « okapi », « narval », « rat taupe », « dragon de komodo »). Cette réponse était également entrée au clavier par l'expérimentateur.

Les images ont ensuite été classées sur la base de la rapidité du temps de dénomination, sur l'exactitude de la réponse, ainsi que sur le jugement de familiarité. Nous avons comparé, pour les noms donnés, si le nom exact de l'animal était donné (par exemple « guépard ») plutôt que le nom de sa catégorie générale (« félin »).

Suite à ce classement, 72 (36+36) images d'animaux ont été retenues, parmi les 228 initiales. Pour les animaux familiers, nous avons gardé les images ayant donné lieu au temps de réaction le plus rapide, associé à leur nom exact et un haut jugement de familiarité. Pour les animaux rares, nous avons gardé les images qui ont donné lieu à des temps de réponse plus longs, associées à des noms de catégorie générale (ex : « insecte »), et des jugements de familiarités relativement bas. Un nombre équivalent d'animaux de chaque classe (rongeur, insecte, ...) a été sélectionné pour les stimuli familiers et pour les stimuli moins familiers, dans la mesure du possible.

Choix saccadique

Les participants devaient poser leur tête sur une mentonnière afin d'éviter les mouvements de la tête. Les stimuli ont été présentés sur un écran d'ordinateur d'une dimension de 30 pouces (Dell).

Les mouvements des yeux (saccades oculaires) étaient enregistrés grâce à un « Eye-tracker » (Fig. 2), qui était placé sous l'écran, en face du participant à une distance de 60cm. Le RED-M (Senso-Motoric Instruments, Teltow, Germany) est basé sur un système « eye-tracking » utilisant des rayons infra-rouges. Le système enregistrait la position des pupilles grâce au logiciel iViewXTM HED, capable de compenser les mouvements de la tête. Les images des yeux étaient analysées en temps réel en détectant le reflet des pupilles, et en éliminant les artefacts. Une phase de calibration était effectuée pour chaque participant avant le début de la tâche.

Les essais se présentaient de cette façon : une croix de fixation était affichée au centre de l'écran pendant 500 ms, puis après un intervalle (gap) de 200 ms, deux photos apparaissaient simultanément de chaque côté de la croix, à 10° du centre, pendant 400 ms. Le logiciel, écrit en

Matlab, était conçu pour ne pas présenter les images tant que le participant ne fixait pas la croix, afin de minimiser les essais incorrects. Pour chaque paire d'images, il y avait une image cible avec un animal, et un distracteur sans animal. Leur localisation (gauche ou droite) était aléatoire et équiprobable. Les participants avaient pour consigne d'effectuer une saccade oculaire vers l'image comportant l'animal. L'intervalle entre les essais était de 2 secondes, et pour chaque essai le programme enregistrait la latence de la première saccade ainsi que la précision des réponses (si le sujet effectuait bien une saccade vers la cible et non pas vers le distracteur).

Les mêmes images étaient présentées une fois en couleur et une fois en nuances de gris, au hasard, afin de contrôler le facteur de la capture du regard par la couleur (exemple : le rouge de la coccinelle qui attire le regard). La position de l'image comportant un animal était aléatoire (soit gauche, soit droite), et le distracteur occupait donc l'autre position.

Les essais dont le temps de réponse était inférieur à 80ms n'ont pas été retenus.



Fig. 2: Photo de l'eye tracker Red-M; SensoMotoric Instruments, Teltow Germany

Réponse manuelle

Les participants étaient cette fois-ci assis devant un écran d'ordinateur tactile légèrement incliné de 21 pouces, et de résolution 1920*1080.

Dans cette tâche, les mouvements du doigt du participant étaient enregistrés, grâce à l'écran tactile. Les sujets devaient avoir leur index en contact avec l'écran, et répondre avec un mouvement du doigt (soit vers la gauche, soit vers la droite). Pour commencer l'essai, les participants devaient avoir leur doigt sur le « point de départ », qui était symbolisé par un cercle blanc en bas de l'écran. Puis, une croix de fixation apparaissait au centre de l'écran pendant 500 ms, puis après un intervalle de 200 ms, deux images apparaissaient à gauche et à droite de l'écran. A chaque essai, une image comportait un animal, et l'autre image servait de distracteur. Les images restaient affichées 400 ms à l'écran avant de disparaître, et avaient toutes les mêmes dimensions (512px * 512px). Les sujets avaient pour consigne de glisser leur doigt sur l'écran le plus rapidement possible vers l'image qui contenait un animal, soit l'image de gauche ou l'image de droite. Les sujets n'avaient pas besoin de glisser leur doigt jusqu'à l'image elle-même, mais simplement en direction de cette image.

Pour mesurer le temps de réaction, nous prenons en compte le premier instant où le doigt du participant commence à se déplacer sur l'écran tactile, c'est-à-dire au moment où il quitte le point de départ.

Pour mesurer l'exactitude de la réponse, nous avons établi deux zones de réponses : une zone « réponse gauche », et une zone « réponse droite », à partir d'un certain nombre de pixels depuis le centre de l'écran en abscisse. Si, à la fin de l'essai la coordonnée « X » du doigt du participant était inférieure à 900 px, alors la réponse était considérée comme « gauche », et si cette coordonnée était supérieure à 1020 px, alors la réponse était considérée comme « droite ». Une réponse dans l'intervalle 900-1020 px représentait automatiquement une mauvaise réponse dans les résultats.

La présentation des images suivait les mêmes règles que pour la tâche de choix saccadique, c'est-à-dire que les images étaient présentées dans un ordre aléatoire, avec des versions en couleur et en nuances de gris, et une position (gauche ou droite) aléatoire pour l'image comportant un animal.

Les essais dont le temps de réponse était inférieur à 80ms n'ont pas été retenus.

Contre-balancement

Pour éviter les effets d'apprentissage des images, la moitié des participants ont commencé par la tâche de choix saccadique, et l'autre moitié a commencé par la tâche de choix moteur. Dans chaque expérience, les images étaient présentées dans un ordre aléatoire, différent pour chaque sujet. Puisque dans chaque condition, chaque image est présentée deux fois (une fois en couleurs et une fois en nuances de gris), les participants étaient exposés 4 fois à chaque image. De plus, comme certains stimuli représentaient des animaux nouveaux (inconnus) pour les participants, un effet d'apprentissage aurait pu être plus marqué sur ce type de stimuli.

Un effet d'apprentissage pour les stimuli « rares » aurait pu favoriser la confirmation de notre hypothèse et ainsi biaiser l'étude. Un effet d'influence de la familiarité aurait pu être dissimulé par un effet d'apprentissage. Avec ce contre-balancement, nous nous assurons de ne pas avoir d'avantages en termes de temps de réaction ou de précision pour les deux conditions du facteur familiarité à cause d'un effet d'apprentissage.

Résultats

Nous cherchions à savoir si la familiarité d'un animal pouvait avoir une influence sur les temps de catégorisation rapide. Les animaux familiers engendrent-ils des réactions plus rapides que les animaux rares ?

Lors des deux expériences, nous avons enregistré la précision des participants, ainsi que les temps de réponses.

Pour la tâche de choix saccadique, 4 participants ont été exclus à cause de problèmes techniques avec l'eye-tracker, et 2 participants ont été exclus des résultats pour la tâche manuelle car leur taux de bonne réponses étaient bien inférieurs au reste du groupe (plus de 2 écart-types de la moyenne). De plus, 4 autres participants n'ont pas pu se rendre au laboratoire pour réaliser les tâches. Il y a donc 15 individus pour la tâche de choix saccadique, et 18 individus pour la tâche manuelle. Pour la tâche de dénomination de l'étude pilote, 1 participant a été exclus de l'étude à cause d'une performance trop faible, et les résultats d'un autre participant n'ont pas pu être sauvegardés. Il y avait donc 20 participants pour l'étude pilote.

Nous avons analysé la précision des réponses (réponse orientée vers l'animal et non vers le distracteur) et le temps de réponse (saccade oculaire ou réponse manuelle) avec une analyse de variance (ANOVA à mesures répétées). Les facteurs à l'étude étaient la familiarité de l'animal (familier ou rare) et la chromaticité de l'image (couleur ou gris).

Tâche de dénomination (étude pilote)

Dans cette étude pilote qui devait nous apporter des confirmations par rapport aux familiarités des images cibles choisies, nous avons enregistré pour chaque image présentée le temps de réponse du sujet, un jugement de familiarité sur une échelle de 1 à 9, et l'exactitude du nom donné par rapport à l'animal présenté. Les résultats de cette étude sont résumés dans le tableau 1 pour chaque participant.

Nous pouvons ici comparer les temps de réponse (dénomination) pour les stimuli familiers et rares. En moyenne, les 20 participants ont mis 1470ms pour donner le nom des animaux familiers, et 1656ms les animaux rares. Avant de mettre en place un test statistique, nous avons voulu vérifier la normalité des distributions. Nous avons donc appliqué le test de Lilliefors aux deux échantillons. Comme nous obtenons une probabilité critique $p < .10$, nous avons décidé de plutôt utiliser une méthode non-paramétrique pour comparer les deux échantillons. Pour une comparaison de deux échantillons à mesures répétées, nous utiliserons le test du signe, c'est-à-dire en comparant les fréquences des signes des différences (entre les moyennes pour chaque participant pour les animaux familiers et rares). Ce test révèle une différence significative entre les deux échantillons ($t(19) = 3,3541$; $p < 0.001$). Les participants ont donc été plus rapides dans cette tâche pour donner le nom des animaux familiers par rapport aux animaux rares.

P	% correct familier	% correct rare	TR familier	TR rare	Jugement familier	Jugement rare
2	89,17	22,22	1,353	1,520	7,81	5,72
4	64,17	3,70	1,254	1,171	7,09	3,85
5	85,83	13,89	1,470	1,499	7,35	2,89
6	81,67	21,30	1,526	1,658	6,84	2,94
7	95,00	36,11	1,821	1,876	6,87	3,06
8	85,00	20,37	1,629	1,828	5,03	2,87
9	80,00	9,26	1,817	1,868	6,90	1,39
10	90,00	20,37	1,404	1,712	8,09	3,44
11	82,50	13,89	1,313	1,687	7,18	3,14
12	79,17	17,59	1,429	1,688	7,61	3,36
13	85,83	20,37	1,816	1,768	7,97	5,25
14	89,17	25,93	1,305	1,585	7,97	4,34
15	87,50	16,67	1,575	1,813	7,43	3,15
16	92,50	31,48	1,257	1,714	8,27	4,06
17	85,83	18,52	1,576	1,721	8,53	4,81
18	90,83	21,30	1,319	1,428	8,32	4,79
19	90,00	30,56	1,351	1,639	8,35	5,47
20	86,67	18,52	1,405	1,726	8,22	4,77
21	88,33	12,96	1,525	1,739	7,55	5,26
22	96,67	25,93	1,261	1,468	7,70	4,25

Tableau 1 : Pourcentage de bonnes réponses, des temps de réponse et des jugements de familiarité en fonction de la familiarité du stimulus pour chaque participant.

Nous avons également conduit une analyse des corrélations entre le temps de réponse et le jugement de familiarité pour chacune des images utilisées. Les corrélations entre chacune des mesures de cette étude sont résumées dans le tableau 2. Il y a une corrélation négative significative ($p < .05$) entre le jugement de familiarité et le temps de réaction ($r = -0.61$), c'est-à-dire que plus le jugement de familiarité est élevé, plus le temps de réponse est bas. Donc plus l'animal était jugé familier, et plus les participants avaient tendance à répondre rapidement, et plus l'animal présenté était jugé rare, et plus le temps de réponse des participants était long.

Nous pouvons également observer une corrélation positive significative ($p < .05$) entre le pourcentage de bonnes réponses pour les animaux familiers et les animaux rares ($r = 0.73$). Il semble y avoir un lien entre le nombre de noms exacts corrects donnés pour les animaux familiers et pour les animaux rares.

Variable	Correlations (Spreadsheet3) Marked correlations are significant at $p < .05000$ N=20 (Casewise deletion of missing data)					
	%fam	%rare	TR fam	TR rare	jugement fam	jugement rare
%fam	1,00	0,73	-0,55	-0,48	0,66	0,27
%rare	0,73	1,00	-0,40	-0,28	0,45	0,31
TR fam	-0,55	-0,40	1,00	0,78	-0,60	-0,38
TR rare	-0,48	-0,28	0,78	1,00	-0,55	-0,42
jugement fam	0,66	0,45	-0,60	-0,55	1,00	0,56
jugement rare	0,27	0,31	-0,38	-0,42	0,56	1,00

Tableau 2 : corrélations entre les pourcentages de bonne réponse, de temps de réaction et des jugements de familiarité en fonction de la familiarité du stimulus.

Expérience 1 : saccade oculaire

Précision

Pour les 15 participants, la précision globale était de 79%, le maximum étant de 89% et le minimum de 69%. Nous avons également décomposé les taux de bonnes réponses pour obtenir une précision pour chaque facteur.

Si l'on compare les précisions en considérant la familiarité comme facteur, on obtient 82.0% pour les animaux familiers, et 75.2% pour les animaux rares. La différence entre ces deux taux de précision est significative ($F(1, 14) = 11,4146$, $p = 0,0045$) : les participants étaient plus précis pour les images comportant un animal familier que pour les images comportant un animal rare.

Pour la précision en fonction de la chromaticité de l'image, on obtient 77.7% pour les images en couleur, et 79.6% pour les images en nuances de gris. La différence entre ces deux taux de précision n'est pas significative ($F(1, 14) = 1,3831$, $p = 0,2591$) : les participants ne semblent pas avoir mieux ou moins bien répondu pour les images colorées.

Si l'on considère uniquement les images d'animaux familiers, pour les images en couleur il y avait 80.0% de bonnes réponses, et 84.1% pour les images en nuances de gris. Les précisions sont considérées comme similaires, il n'y a pas de différence d'erreurs entre les images en couleur ou en nuances de gris pour les animaux familiers ($t(14) = -1.082$; $p = 0.288$). Pour les animaux rares, il y avait 75.4% de bonnes réponses pour les images en couleur, et 75.2% pour les images en nuances de gris. Là encore, la différence entre les deux précisions n'est pas significative ($t(14) = 0.054$; $p = 0.956$). L'effet de l'ANOVA pour la couleur et la familiarité n'est pas non plus significatif ($F(1, 14) = 1,4199$, $p = 0,2532$).

Temps de réponse

Pour l'ensemble des 15 participants, le temps de réponse moyen (quand la réponse était correcte) était de 263 ms, avec un écart type de 41 ms entre les moyennes de chaque participant. Les temps de réponse pour les bonnes réponses en fonction de la familiarité de l'animal présenté et de la couleur sont résumés dans le tableau 3.

Temps de réaction	Couleur	Gris	Total
Familier	272 +/- 41	267 +/- 41	257 +/- 38

Rare	258 +/- 41	260 +/- 45	253 +/- 41
Total	264 +/- 41	261 +/- 42	263 +/- 41

Tableau 3 : temps de réaction en millisecondes en fonction de la familiarité de l'animal (commun/rare) et la chromaticité (couleur/gris) dans la tâche de choix saccadique.

Le but de l'étude était de déterminer s'il y aurait ou non un effet de la familiarité ou de la couleur sur la latence de la réponse. Pour ce faire, nous avons réalisé une ANOVA à mesures répétées, en prenant comme facteur la familiarité et la couleur.

Aucune différence significative n'a été observée pour la familiarité, ($F(1,14) = 1,0537$, $p=0,3220$), ni pour la chromaticité ($F(1,14) = 0,0572$, $p=0,8144$). Les participants n'ont donc pas été plus rapides pour effectuer des saccades oculaires vers les animaux familiers ou rares, et la couleur n'a pas non plus eu d'effet significatif sur les temps de réponse des participants.

Nous nous sommes ensuite intéressés à la distribution des temps de réponse saccadiques. Le nombre de bonnes réponses a été organisé pour chaque pas de 20 ms et en fonction de chaque condition (la familiarité et la couleur). Ces distributions sont illustrées dans la figure 3.

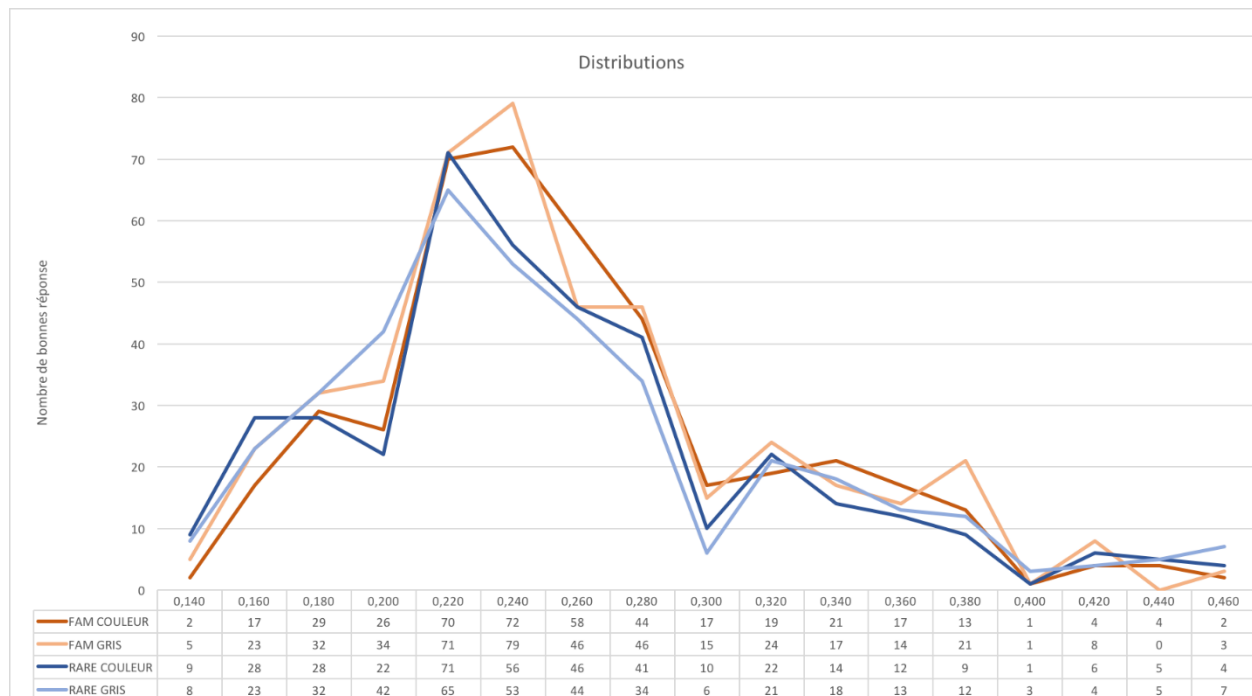


Fig. 3 : distributions des bonnes réponses en fonction par pas de 20 ms en fonction de chaque condition (animal familier en couleur, en nuances de gris, animal rare en couleur et en nuances de gris).

Nous pouvons remarquer que pour les temps de réponse les plus courts (moins de 220 ms), les distributions sont très proches l'une de l'autre, ce qui indique que la familiarité semble ne pas avoir d'influence pendant ces périodes.

Pour la période de 240 ms à 300 ms, nous pouvons voir que les distributions des stimuli familiers sont supérieures à celles des stimuli rares en termes de nombre de bonnes réponses. Puis, à partir de 320 ms, les distributions semblent atteindre un plafond.

Expérience 2 : réponse manuelle

Précision

Pour les 18 participants, la précision globale était de 95.1, le maximum étant de 100% et le minimum de 84%.

En comparant les précisions en prenant la familiarité comme facteur, les participants ont obtenu 94.9% pour les animaux familiers, et 93.7% pour les animaux rares. La différence entre ces deux taux de précision n'est pas significative ($F(1, 16) = 1,711, p = 0,2093$) : les participants étaient aussi précis pour les images représentant un animal familier que pour les images comportant un animal rare.

Pour les précisions en rapport avec la couleur de l'image, on obtient 93.9% pour les images en couleur, et 94.6% pour les images en nuances de gris. La différence entre ces deux taux de précision n'est pas significative non plus ($F(1, 16) = 0,7606, p = 0,3960$), donc les participants ne semblent pas avoir mieux ou moins bien répondu en fonction de la couleur de l'image de manière générale.

Si l'on s'intéresse maintenant seulement aux images d'animaux familiers, pour les images en couleur il y avait 94.7% de bonnes réponses, et 95.0% pour les images en nuances de gris. Les moyennes sont très proches l'une de l'autre, les précisions sont considérées comme similaires ($t(16) = -0.164 ; p = 0.8705$) : il n'y a pas de différence d'erreur entre les images en couleur ou en nuances de gris pour les animaux familiers. Pour les animaux rares, il y avait 93.1% de bonnes réponses pour les images en couleur, et 94.3% pour les images en nuances de gris. Là encore, la différence entre les deux précisions n'est pas significative ($t(16) = -0.612 ; p = 0.5448$). La combinaison des facteurs de familiarité et chromacité n'engendre de différences significatives de précision pour les participants ($F(1, 16) = 0,3703, p = 0,5513$).

Temps de réaction

Pour l'ensemble des 18 participants, le temps de réponse moyen était de 325 ms, avec un écart type de 49 ms entre les moyennes de chaque participant. Le participant avec le temps de réaction moyen le plus rapide est à 238 ms, et le participant avec le temps de réaction moyen le plus élevé est à 413 ms. Les temps de réactions pour les réponses en fonction de la familiarité de l'animal présenté et de la couleur sont résumés dans le tableau 4.

Temps de réaction	Couleur	Gris	Total
Familier	325 +/- 50	323 +/- 55	326 +/- 52
Rare	325 +/- 51	328 +/- 50	327 +/- 50
Total	324 +/- 49	325 +/- 49	325 +/- 49

Tableau 4 : temps de réponse en millisecondes en fonction de la familiarité de l'animal (commun/rare) et la couleur (couleur/gris) dans la tâche tactile.

Aucune différence significative n'a été observée pour la familiarité, ($F(1, 17) = 0,0755$, $p = 0,7868$), ni pour la couleur ($F(1, 17) = 0,2048$, $p = 0,6575$). Les participants n'ont donc pas été plus rapides pour effectuer un mouvement vers les animaux familiers ou rares, et la couleur n'a pas non plus eu d'effet significatif sur les temps de réponse des participants. Il n'y a pas non plus d'interaction significative entre ces facteurs ($F(1, 17) = 0,7183$, $p = 0,4084$).

Comparaison choix saccadique et réponse manuelle

Pour comparer les deux tâches, nous avons fait la moyenne des deux tâches pour chaque participant et pour tous les facteurs. Comme les valeurs s'éloignent trop de la loi normale (choix saccadique, Lilliefors $p < .01$, tâche manuelle, Lilliefors $p < .05$), une méthode non-paramétrique a été utilisée pour comparer ces deux échantillons. Au niveau des taux de précision, on peut observer que la précision était supérieure dans la tâche avec la réponse manuelle (95,1% en moyenne) que pour la tâche de choix saccadique (78,6% en moyenne). Cette différence a été analysée avec le test de Kolmogorov-Smirnov, qui montre que cette différence est significative ($p < 0.001$). Les participants ont donc été plus précis pour la tâche manuelle que pour la tâche de choix saccadique. Pour les temps de réponse, c'est dans la tâche de choix saccadique que les temps de réponse sont les plus bas (262 ms en moyenne, contre 322 ms pour la tâche manuelle). La différence des temps de réponse est significative ($p < 0.005$) selon le test de Kolmogorov-Smirnov.

Discussion

Le but de notre étude était de déterminer l'influence de la familiarité dans la catégorisation rapide de scènes naturelles. Nous avons comparé la performance pour un paradigme de choix saccadique rapide et pour une tâche de catégorisation avec réponse manuelle plus lente. Nous allons discuter des implications théoriques des résultats obtenus pour ces expériences, puis nous adopterons un regard critique sur cette étude avant de proposer de nouvelles perspectives de recherche.

Dans la première expérience, avec les saccades oculaires, nous avons fait l'hypothèse que la familiarité, qui implique un traitement cognitif de haut niveau n'aurait pas le temps d'être traitée avant la réponse (le temps de la première saccade).

Les résultats ont montré qu'il n'y a pas de différences significatives pour les temps de réponse pour les animaux familiers et pour les animaux rares, ce qui confirme donc notre hypothèse. La familiarité n'a pas le temps d'avoir une influence sur notre latence de saccade. Les distributions nous montrent que la familiarité peut avoir un effet, mais pour les temps de réaction plus longs, dans l'intervalle de 240 ms à 300 ms. Pour les réponses les plus rapides, il n'y a aucune différence impliquant la familiarité du stimulus. Ceci nous permet donc de suggérer que la familiarité serait bien un processus cognitif tardif. Le temps de traitement nécessaire pour l'accès à la familiarité se trouverait donc aux alentours de cet intervalle, de 240 ms à 300 ms. Pour les réponses les plus rapides, seules les caractéristiques physiques des images auraient une importance. Le système visuel tirerait partie de ces caractéristiques pour déterminer la présence ou l'absence d'un animal, sans accéder à la nature de cet animal en termes de « connu / non-connu ». Ceci correspondrait au niveau super-ordonné. Le niveau de base correspondrait à la différenciation entre différents types d'animaux (« chien », « félin »). Et donc, la familiarité se trouverait elle au niveau sub-ordonné (ex : « caniche », « mon chien »). Ces résultats montrent donc que nous sommes plus rapides pour accéder au niveau super-ordonné que pour le niveau de base, ce qui est en accord avec les travaux de Thorpe et collaborateurs (Macé et al, 2009 ; Poncet et al, 2014 ; Wu et al, 2005).

Dans la deuxième expérience, avec la réponse manuelle, l'hypothèse de base était qu'avec des temps de réponse plus lents un effet de la familiarité pourrait se manifester.

Nous avons vu dans les résultats qu'il n'y avait pas non plus de différences significatives pour les temps de réaction entre les animaux familiers et les animaux rares. Cependant, dans cette tâche, on se trouve dans des temps de réponse relativement longs, là où on est au plafond dans la distribution des réponses en choix saccadique.

Nous n'avons également pas trouvé d'effets en lien avec la couleur des images, ce qui est en accord avec les travaux de Delorme et al (2000), qui n'ont pas trouvé de différences dans une tâche de catégorisation rapide avec des images en couleur et des images en nuances de gris. Mais Oliva et Schyns (2000) rapportent eux que la couleur peut influencer la reconnaissance lorsque celle-ci est diagnostique, ou prédictive d'une catégorie, comme le jaune pour citron. Cependant, les couleurs n'avaient aucun effet pour la reconnaissance de scènes non-diagnostiques. Cet effet pour les couleurs diagnostiques a aussi été montré dans les études de Humphrey (1989), Tanaka et Presnell (1999) et Wurm et al (1993), qui ont trouvé que les informations sur les couleurs permettent une reconnaissance plus rapide des objets, mais seulement ceux pour lesquels le lien avec la couleur est fort : une carotte était dénommée plus rapidement quand présentée en couleur, mais il n'y avait pas de différences entre les images en couleur et les images en nuances de gris pour des voitures. La couleur semble donc importante pour la reconnaissance, mais seulement pour les objets avec une couleur diagnostique. Pourtois et al (2004) supposent que la reconnaissance des objets au quotidien qui se ferait au niveau de base serait facilitée par la présence de la couleur. Il se peut donc que cette règle des couleurs diagnostiques s'applique également à certains animaux. Par exemple, nous avons utilisé des coccinelles dans les images présentées, dont la couleur rouge pourrait être considérée comme diagnostique, de même pour le tigre, ou le panda. Les autres animaux auraient des couleurs plus variées, et donc n'auraient pas de couleurs diagnostiques, expliquant pourquoi de manière générale nous n'avons pas obtenus d'effets de la couleur (ni pour les taux de précisions, ni pour les temps de réponse, dans les deux paradigmes).

Nous avons aussi observé des taux de précision plus élevés pour la tâche manuelle que pour la tâche de choix saccadique, mais également des temps de réponse plus élevés. Dans la tâche

manuelle, les participants avaient plus de temps pour explorer les images, et si leur première saccade oculaire ne se posait pas initialement sur l'image comportant un animal, il leur était possible de corriger poser leur regard sur l'autre image et corriger leur programme moteur. Ces résultats sont conformes à ceux trouvés par Boucart et al (2014), où la tâche manuelle engendrait des temps de réaction plus longs à ceux de la tâche de choix saccadique.

Avec les données de notre étude pilote, nous avons pu observer une corrélation entre le degré de familiarité perçu d'un animal, et le temps de réponse. Plus l'animal était jugé familier, et plus les participants étaient rapides pour en donner le nom. Comme cette tâche demandait aux participants un traitement supplémentaire, car ils devaient faire un lien visuel-mot nécessitant un accès lexical, les temps de réaction pour cette tâche étaient largement plus grand (entre 1300ms et 1800ms) que pour ceux des deux autres expériences. Avec une fenêtre temporelle de réponse plus longue à cause du temps nécessaire pour l'accès au lexique, nous pouvons observer un effet de la familiarité du stimulus. Cependant, cette expérience nous servait à vérifier que les images utilisées dans les deux autres expériences représentaient bien des animaux comme jugés très familiers ou très rares. Comme il était naturel que pour les animaux rares les participants ne connaissent pas les noms des animaux, il faut prendre en compte un certain temps d'hésitation avant que les participants se décident plutôt à donner le nom de catégorie générale (par exemple « oiseau » pour ibis). Donc, il aurait fallu plus de temps pour faire la dénomination au niveau de base plutôt qu'au niveau subordonné. Or, selon Macé et al (2009) ou Wu et al (2005), la catégorisation au niveau super-ordonné serait plus rapide que la catégorisation au niveau de base. Cependant, nous nous trouvons ici dans une tâche de dénomination, qui n'implique pas les mêmes traitements qu'une tâche de catégorisation.

Pour nos tâches de catégorisation, les résultats obtenus sont en accord avec ces idées d'un avantage de traitement au niveau super-ordonné par rapport au niveau de base. Les participants de cette étude étaient capables de reconnaître l'animal en moyenne aussi vite peu importe le degré de familiarité, et peu importe si le nom de l'animal était connu ou simplement le nom de sa catégorie, comme nous avons pu le voir avec la tâche de dénomination. Comme la consigne des participants était d'effectuer une saccade oculaire ou un geste de la main vers l'image comportant un « animal », seule la reconnaissance de l'animal au niveau super-ordonné était nécessaire. Le nom de l'animal ou même de sa catégorie n'étant pas utile, ceux-ci n'avaient pas le temps d'être activé avant la

réponse, c'est à dire que les participants étaient sûrement conscient de la présence d'un animal avant de savoir précisément de quel type d'animal il s'agissait.

Cette absence d'effet de familiarité pour des temps de réponse aussi bas s'expliquerait par le fait que l'œil serait attiré par les caractéristiques globales qui composent les animaux et qui nous permettraient de les reconnaître rapidement. Même pour les animaux rares qui étaient généralement inconnus des participants, les taux de bonne réponse étaient suffisamment élevés pour ne pas être attribués au hasard. Les participants ont obtenu un score de 75% pour la tâche de choix saccadique, et 93% pour la tâche manuelle, chacun n'étant pas statistiquement différent de la précision pour les animaux familiers. Donc alors même que certains de ces animaux n'avaient sûrement jamais été vu par les participants, ils étaient en mesure de non seulement les catégoriser en tant qu'animaux, mais également aussi rapidement que pour les animaux familiers. Les propriétés communes aux animaux semblent donc être critiques surtout pour les 200 premières millisecondes, et des effets de la familiarité ne pouvant être observés que pour des réponses plus tardives de l'intervalle 240-300ms.

Un effet similaire a été observé par l'équipe de Barbeau et al (2015), lors d'une étude sur la reconnaissance des visages et l'effet de la familiarité. Avec un paradigme de go/no-go et des enregistrements en électroencéphalographie (EEG), ils ont pu observer un effet de la familiarité sur l'onde N250 à partir de 200 millisecondes après la présentation du stimulus (l'amplitude de l'onde est modifiée en fonction de la familiarité du visage présenté, mais pas sa position). L'intérêt de cette étude était de pouvoir observer l'activité du cerveau avant la réponse comportementale des participants, et de montrer l'importance d'une certaine onde cérébrale dans la familiarité. Les effets de la familiarité n'influencent cette onde qu'à partir de 200 ms, et pour les visages. Or, les visages sont souvent reconnus plus rapidement que les animaux (Kirchner & Thorpe, 2005 ; Crouzet, Kirchner, Thorpe, 2010 ; Boucart et al in press). Si cette onde peut également refléter les effets de familiarité pour des animaux, il se pourrait que les effets de familiarité apparaissent plus tard.

Limites et ouvertures

Finalement, il convient de proposer quelques ouvertures et problématiques pour de futures recherches au sujet de la reconnaissance dans la vision et les paradigmes de catégorisation de scènes naturelles. Il existe dans la littérature deux approches pour l'étude d'une fonction cognitive : l'étude de participants sains, et l'étude de patients. On parle également de modèle d'activation (cerveau sain) et de modèle lésionnel (cerveau pathologique). Ces deux approches apportent des informations complémentaires, ce qui permet une compréhension plus riche de la fonction étudiée.

Nous avons également présenté certaines données issues d'imageries cérébrales en EEG de Barbeau et al. (2015) concernant l'activité cérébrale de la reconnaissance des visages familiers et inconnus : des effets de la familiarité pour les visages ont pu être observés au niveau de l'amplitude de l'onde cérébrale N250, il semble donc probable de pouvoir observer un phénomène similaire pour la familiarité des animaux.

Une étude a permis de montrer que l'accès à la familiarité pouvait être modifié dans certaines pathologies comme la schizophrénie (Horn et al, 2015). Les patients qui souffraient de schizophrénie avaient tendance à évaluer des visages comme étant plus familiers que les contrôles, suggérant l'existence d'un symptôme d'hyper-familiarité dans cette pathologie. Il pourrait donc être intéressant de voir quels seraient les effets de différentes pathologies sur la familiarité au niveau des animaux. En utilisant le même paradigme de notre étude, on pourrait s'attendre à avoir des effets d'hypo-familiarité chez les patients Alzheimer, ou d'hyper-familiarité chez les patients schizophrènes.

Et finalement, il pourrait être intéressant de proposer une étude pouvant servir de contrôle à la notre, en utilisant le paradigme du choix saccadique, mais cette fois en demandant un traitement explicite de la familiarité. En présentant deux animaux (un familier et un rare) à l'écran, et en demandant de faire une saccade vers l'animal familier, il serait possible de tester si le temps de réponse correspond bien à la fenêtre temporelle de 240-300 ms que nous avons pu observer dans notre étude. Si l'animal familier est reconnu plus rapidement que l'animal rare, grâce à la familiarité, la majorité des premières saccades dans cette fenêtre temporelle devraient être effectuées dans la direction de l'animal familier.

Conclusion

En conclusion, dans cette étude, nous avons tenté de déterminer la contribution de la familiarité du stimulus dans la catégorisation des scènes naturelles. Inspirés par les travaux de Thorpe et al (2012), nous avons cette fois pris en compte la familiarité et la rareté des animaux, à l'aide d'une étude pilote, puis avec une expérience de choix saccadique et une expérience de choix motrice, chez des participants jeunes disposant d'une bonne vision.

Nous avons donc pu mettre en évidence que la familiarité du stimulus n'a pas d'influence pour les premières étapes du traitement de l'image par le système visuel, et interviendra seulement après la reconnaissance de la catégorie de l'objet. Avant 240 ms, seules les caractéristiques physiques globales sont prises en compte par le système visuel, qui vont attirer l'œil, c'est-à-dire la forme d'un animal ou non-animal. Un effet de familiarité peut être observé dans une tâche de choix saccadique seulement entre 240 et 300 ms.

Ces résultats nous montrent que lors des premières étapes du traitement de l'information visuelle, le cerveau n'utilise que les caractéristiques physiques de l'objet dans le but de prendre une décision rapidement. Ceci se rapporte aux différents niveaux de catégorisation comme définis par Oliva. Le système visuel est capable de comprendre très rapidement la présence d'un animal à un niveau super-ordonné, peu importe si cet animal est connu ou inconnu.

Références

- Barragan-Jason, G., Cauchoix, M., Barbeau, E.J. (2015). The neural speed of familiar face recognition, *Neuropsychologia*, 75, 390-401.
- Boucart, M., Calais, G., Lenoble, Q., Moroni, C., & Pasquier, F., (2014). Differential processing of natural scenes in typical and atypical Alzheimer disease measured with a saccade choice task. *Front Integr Neurosci.* 25:8:60.
- Boucart, M., Lenoble, Q., Quettelard, J., Szaffarczyk, S., Desprez, P., & Thorpe, S.J. (2016 in press). Finding faces, animals and cars in far peripheral vision. *Journal of Vision*.
- Busettini, C., Masson, G. S., & Miles, F. A. (1997). Radial optic flow induces vergence eye movements with ultra-short latencies. *Nature*, 390 (6659), 512–515.
- Crouzet S., Kirchner, H., Thorpe, S. (2010) Fast saccades towards face: Face detection in just 100 ms. *Journal of Vision* 10: 1–17.
- Delorme, A., Richard, G., Fabre-Thorpe, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on color cues: a study in monkeys and humans. *Vision Research*, 40, 2187–2200.
- Fabre-Thorpe, M., Richard, G., Thorpe, S. (1998) Rapid categorization of natural images by rhesus monkeys. *Neuroreport*, 9, 303–308.
- Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. J. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, 13 (2), 1–10.
- Fabre-Thorpe, M. (2003). Visual categorization: accessing abstraction in nonhuman primates. *Philos Trans R Soc Lond B Biol Sci*, 358, 1215–1223.
- Goodale, M., Milner, D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, 15 (1), 20-25.
- Grill-Spector, K., Malach, R., (2004). The human visual cortex. *Annu Rev Neurosci.* 27, 649-77.

- Horn, M., D'Hondt, F., Vaiva, G., Thomas, P., Pins, D., (2015). Categorical perception of familiarity: Evidence for a hyper-familiarity in schizophrenia. *J Psychiatr Res*, 71, 63-9.
- Hubel, D., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. (Lond.)* 195, 215–243.
- Ishai, A., Ungerleider, L.G., Martin, A., Haxby, J.V., (2000). The representation of objects in the human occipital and temporal cortex. *J Cogn Neurosci*. 12 (2), 35-51.
- Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews. Neuroscience*, 6, 787–798.
- Langton, S., Law, A., Burton, A., Schweinberger, S. (2008). Attention capture by faces. *Cognition*, 107, 330–342.
- Macé, M., Joubert, O., Nespoulous, J., Fabre-Thorpe, M. (2009). The time-course of visual categorizations: you spot the animal faster than the bird. *PLoS On*, 4 (6).
- New J, Cosmides L, Tooby J (2007) Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National ...* 104:16598.
- Ohman, A., Carlsson, K., Lundqvist, D., & Ingvar, M. (2007). On the unconscious subcortical origin of human fear. *Physiology Behavior*, 92, 180–185.
- Oliva, A., Schyns, P., (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology* 41, 176–210.
- Oliva, A., Torralba, A., (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, 42, 145–175.
- Peelen, M., Kastner, S. (2014). Attention in the real world : toward understanding its neural basis. *Trends in Cognitive Sciences*, 18 (5)
- Poncet, M., Fabre-Thorpe, M. (2014). Stimulus duration and diversity do not reverse the advantage for superordinate-level representations: the animal is seen before the bird. *Eur J Neurosc*, 39 (9), 1508-16.
- Potter MG (1976). Short-term conceptual memory for pictures. *J Exp Psychol Hum Learn* 2: 509-522

- Torralba, A., Oliva, A. (2003) Statistics of natural image categories. *Network (Bristol, England)*, 14, 391–412.
- Thorpe, S., Fize, D., Marlot, G. (1996). Speed of processing in the human visual system. *Nature*, 381, 520-522.
- Theeuwes, J., & Van der Stigchel, S. (2006). Faces capture attention: Evidence from inhibition of return. *Visual Cognition*, 13 (6), 657-665.
- Vuilleumier, P. (2000). Faces call for attention: Evidence from patients with visual extinction. *Neuropsychologia*, 38, 693–700.
- Wu, C., Crouzet, S., Thorpe, S., Fabre-Thorpe, M. (2015). At 120 msec you can spot the animal but you don't yet know it's a dog. *J Cogn Neurosci*, 27 (1), 141-9.

Annexes

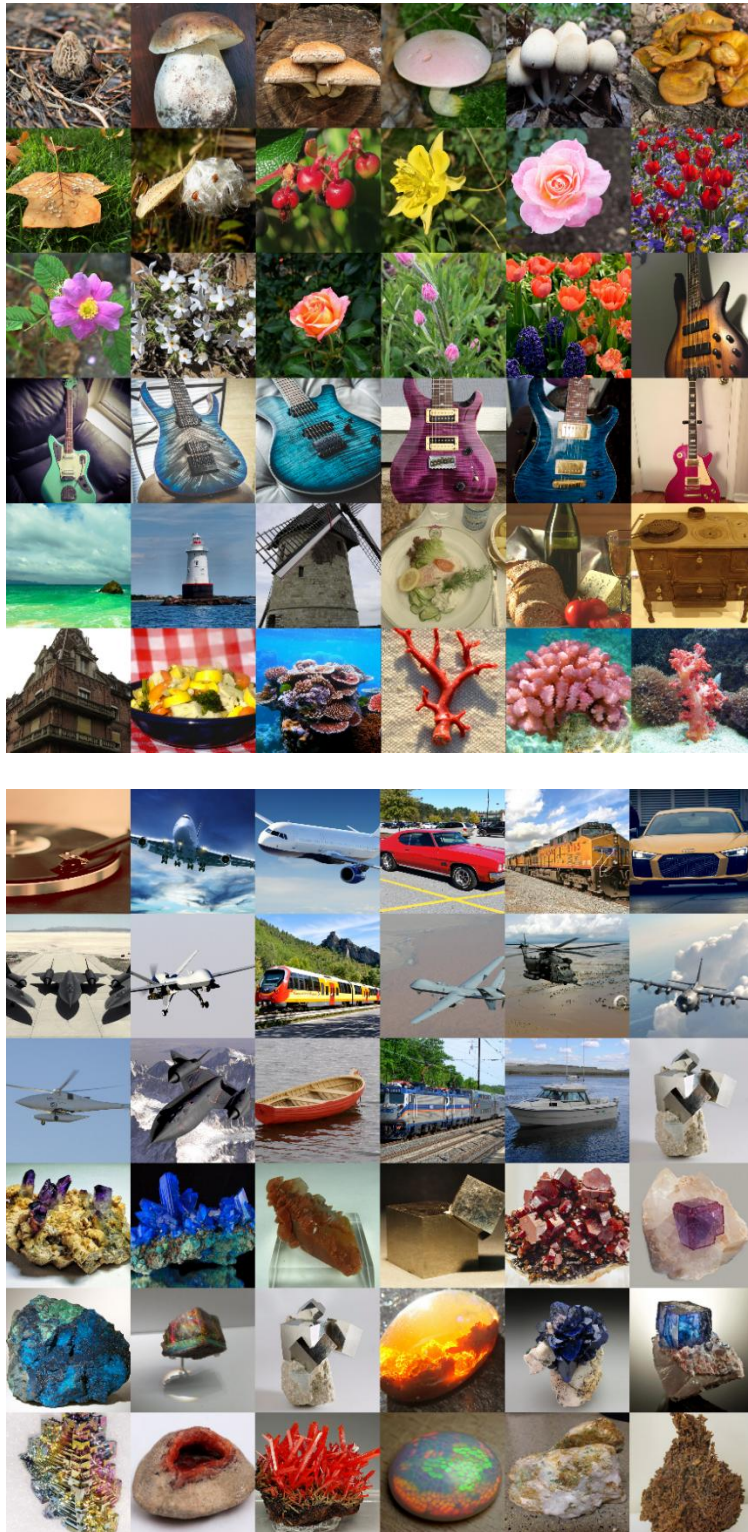
Annexe I

Images d'animaux familiers et rares utilisées suite à l'étude pilote



Annexe II

Images utilisées pour les distracteurs



**Comité d'éthique en sciences comportementales**Président :

Yvonne DELEVOYE-TURRELL

Président adjoint :

Céline DOUILLIEZ

Personne ressource (dossier administratif) :

Aurélie DUCROQUET

Tél : 03.20.41.67.92 -

E-mail : aurelie.ducroquet@univ-lille3.fr

Villeneuve d'Ascq le 22 avril 2016

Références comité d'éthique :	2015-2-S36
Sigle :	IMAGE
Numéro de version et date :	Version 1 du 25/10/2015
Promoteur :	Lille 3
Porteur projet :	Muriel Boucart et Yann Coello

Date de la soumission : 28/10/2015**Date de la réunion du comité d'éthique :** 24/11/2015**Avis du comité d'éthique :** Avis Favorable

Le protocole est accepté en état. Si pour une quelconque raison, vous souhaitez modifier le protocole (en terme de calendrier, inclusion d'un nouveau groupe...), vous êtes tenu d'informer le comité d'éthique par l'envoi d'un avenant expliquant les motivations mais également les modifications apportées au protocole initial.

Cet avenant sera réévalué par le comité d'éthique.

Pr Yvonne DELEVOYE-TURRELL
Présidente du comité d'éthique



La plupart des remarques ont été prises en compte.

Cependant, deux points semblent devoir être corrigés:

- (1) à la page 8, il est fait état de "la moitié des sujets ($N=10$) alors que, compte tenu de la taille de l'échantillon total ($n=40$), on devrait avoir 10 participants (ce point est mineur).
- (2) Le deuxième point concerne la lettre d'information, il semble qu'une description de ce qu'est un "oculomètre" devrait être fournie au participant pour éclairer son avis de participer.
avis favorable quant à l'ajout d'une deuxième condition.