

Année 2004 / 2005

Université Bordeaux 2

Laboratoire Cerveau & Cognition

**Discrimination visuelle rapide**  
**Étude comparative des indices bas-niveau.**

Sébastien Crouzet

Le 17 juin 2005

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Contraintes temporelles en catégorisation visuelle de scènes complexes . . .	2
1.2	Quels attributs de l'images vont être utilisés pour permettre des temps de réponse si rapides ? . . . . .	3
1.3	Les saccades oculaires . . . . .	5
<b>2</b>	<b>Méthodes</b>	<b>6</b>
2.1	Stimuli . . . . .	6
2.2	Expérience préliminaire . . . . .	9
2.2.1	Protocole expérimental . . . . .	9
2.2.2	Résultats . . . . .	10
2.3	Expérience principale : Tâche de discrimination saccadique . . . . .	10
2.3.1	Protocole expérimental . . . . .	10
2.3.2	Matériel . . . . .	11
2.3.3	Enregistrement des réponses . . . . .	11
2.3.4	Calcul des temps de réaction minimum . . . . .	12
2.3.5	Tâche contrôle moteur . . . . .	12
<b>3</b>	<b>Résultats</b>	<b>12</b>
3.1	Tâche contrôle moteur . . . . .	12
3.1.1	Comparaison des cibles . . . . .	12
3.1.2	Comparaison des essais du début et de la fin . . . . .	13
3.2	Tâche de discrimination saccadique . . . . .	13
3.2.1	Analyse des temps de réaction et de la précision . . . . .	13
3.2.2	Distributions de temps de réaction . . . . .	15
3.2.3	Temps supplémentaire de discrimination . . . . .	17
3.3	Tâche manuelle . . . . .	18
<b>4</b>	<b>Discussion</b>	<b>20</b>
4.1	Vers une hiérarchie des indices bas-niveau ? . . . . .	20
4.2	Une scène complexe traitée très rapidement . . . . .	21
4.3	Une scène simple traitée très lentement . . . . .	22
	<b>Bibliographie</b>	<b>23</b>

# 1 Introduction

Notre système visuel reçoit perpétuellement une grande quantité d'information de l'environnement. Celui-ci change constamment et tout ceci doit être traité à très grande vitesse. Si on prend l'exemple d'un prédateur apparaissant dans le champs visuel, on peut supposer qu'il serait avantageux pour la survie de l'espèce de réagir très rapidement. La reconnaissance d'un objet se ferait donc de façon très rapide. La question qui vient ensuite est de savoir comment va se faire cette reconnaissance à grande vitesse. Va-t'elle être guidée par certains indices simples comme le font certains animaux pour reconnaître leurs prédateur ? Ou alors le cerveau est-il assez rapide pour effectuer des traitements complexes de manière quasi-immédiate ?

## 1.1 Contraintes temporelles en catégorisation visuelle de scènes complexes

Le traitement visuel d'une scène complexe se fait de façon très rapide. Ainsi, lors d'une tâche go/no-go dans laquelle des images sont flashées très brièvement (20 ms), les temps de réaction (TR) moyen pour savoir si ces scènes naturelles contiennent un animal, un véhicule ou un aliment sont de l'ordre de 400 ms, avec 94% de bonnes réponses. Et ces réponses correctes sont significativement supérieures aux réponses fausses à partir de 250 ms [Thorpe et al., 1996]. Les TR observés lors de ces tâches comprennent le temps de traitement de l'information visuelle, le temps de catégorisation de l'image, et le temps de l'initiation de la réponse motrice.

Pour s'affranchir de ce dernier, des enregistrements de potentiels évoqués ont été réalisés pendant la tâche et ont montré une activité différentielle très précoce de certaines populations de neurones à partir de 150 ms. Cette activité prendrait naissance dans les aires visuelles extrastriées [Fize, 2000] et reflèterait une réelle décision perceptive indépendante des différences physiques entre les images appartenant à chaque catégorie [Van Rullen and Thorpe, 2001]. Un délai de 150 ms semble donc être suffisant pour effectuer le traitement visuel nécessaire à la détection d'une catégorie d'objet dans l'image. Cette latence est extrêmement courte et l'on peut donc se demander comment transporter et traiter l'information de la rétine au cortex extrastrié en un temps aussi court.

Si l'on se réfère à l'organisation du système visuel, et plus précisément à la voie ventrale en majeure partie responsable de la reconnaissance des objets, plusieurs étapes sont nécessaires à ce traitement. Ainsi, d'après les connaissances accumulées sur la hiérarchie du système visuel, l'information visuelle aurait à traverser environ dix aires de traitement.

Les informations rétinienne seraient envoyées vers les corps géniculés latéraux (CGL) à travers le nerf optique. Ensuite, ces informations traverseraient entre autres le thalamus, V1, V2 puis V4 avant de se projeter vers le cortex inférotemporal. En prenant en compte le temps nécessaire pour la réalisation de chaque étape (environ 10/15 ms, Bullier [2003]), on s'aperçoit que la durée de traitement de 150 ms ne laisse le temps qu'à une première vague d'information. Une grande partie du processus mis en évidence dans ces expériences de catégorisation rapide serait donc de nature feed-forward et massivement parallèle.

Dans le même ordre d'idée, il a été montré que la présence ou non d'un objet dans une scène était corrélée avec des statistiques de second-order, comme la signature spectrale [Torrallba and Oliva, 2003]. Ainsi, la première vague feedforward d'information pourrait semble-t'il permettre, sans étapes de segmentation ou de regroupement mais juste grâce aux statistiques générales de l'image, la détection de la présence de certains objets dans une scène complexe. Ceci pourrait ensuite servir à donner des informations contextuelles descendantes sur cette scène.

## **1.2 Quels attributs de l'images vont être utilisés pour permettre des temps de réponse si rapides ?**

Au niveau comportemental, on pouvait se demander si des indices dans les images présentées pouvaient être responsables de cette rapidité de traitement. Ainsi, dans une tâche go/no-go de catégorisation d'images naturelles, le fait que les images soient en couleur ou en noir et blanc n'aurait pas d'effet sur les performances [Delorme et al., 2000]. On peut donc penser que la couleur n'est pas utilisée pour la catégorisation rapide. Ce qui suggère que ce processus se ferait par un traitement feed-forward de l'information achromatique de la voie magnocellulaire. D'après ces résultats on peut émettre l'hypothèse qu'une tâche où l'on demanderait à un sujet de discriminer deux images uniquement sur les indices bas-niveaux que sont leurs couleurs pourrait entraîner des temps de réponses plus longs. Une autre méthode d'expérimentation a permis, en revanche, de mettre en évidence l'importance de la couleur pour le traitement des images complexes [Gegenfurtner and Rieger, 2000]. Dans une tâche de reconnaissance d'images naturelles, le sujet devait dire si l'image a déjà été vue ou pas, les résultats montrent que la couleur serait impliquée à plusieurs niveaux d'intégration. Au niveau précoce, elle constituerait un indice supplémentaire sur lequel pourrait se baser la segmentation de l'image. Et à un niveau plus tardif, elle serait un indice de plus pour la récupération de l'information en mémoire.

D'autres études ont testé la vitesse de traitement pour catégoriser des images de formes simples comme des ronds et des carrés. Dans une tâche go/no-go, la catégorisation de

forme simple serait 40 ms plus rapide que la catégorisation 'Animal vs non-Animal' [Fize, 2000]. Cependant, les données enregistrées pour la tâche 'Animal' sont beaucoup plus tardives que dans les expériences précédentes. Et la différence est beaucoup moins nette sur les essais corrects les plus précoces. Ces résultats restent donc à confirmer. Au niveau électrophysiologique, il semblerait que les processus mis en jeu pour la décision 'Cible vs Distracteur' soient les mêmes quelle que soit la complexité des stimuli. Lors d'une étude sur les vitesses de traitement chez le migraineux utilisant des stimuli comparables [Aubertin et al., 1999], il a été montré une différence similaire (environ 40 ms) entre les stimuli 'forme simple' et 'Animal vs non-Animal'.

Les différents attributs des stimuli visuels sont habituellement considérés comme étant traités de façon relativement indépendante par différentes zones du cerveau [Livingstone and Hubel, 1988]. Et ces différences se retrouveraient à tous les niveaux du système visuel. Un module s'occuperait de la couleur, un autre de la forme, un autre du mouvement, et ainsi de suite.

Mais une autre question qui peut se poser est celle de la simultanéité du traitement des caractéristiques de bas-niveau des stimuli. Ainsi, il paraît plus économique, pour l'extraction d'une information donnée, de ne traiter qu'un attribut contenant cette information. Une information pourrait souvent être extraite d'un unique attribut et en traiter un autre s'avérerait être redondant [Bullier, 2001]. On pourrait ainsi penser plus en terme de hiérarchie des indices visuels plus qu'en une hiérarchie des aires visuelles. Le système visuel produit constamment des hypothèses de travail à partir de l'information rétinienne. Celle-ci serait donc utilisée pour mettre à jour le modèle interne du monde. Le premier attribut à être traité serait le mouvement, du fait de son passage par la voie magnocellulaire et de son importance pour la segmentation de l'image : "ce qui va ensemble, bouge ensemble". De la même façon, lors d'une tâche consistant en un temps de réponse comportemental, il a été montré que la détection d'un changement de mouvement était plus rapide que celle d'un changement de couleur [Adams and Mamassian, 2004].

L'arrivée plus tardive de la voie parvocellulaire ferait que les attributs concernant les détails plus fins (et aussi la couleur) seraient traités plus tardivement. Et ce traitement pourrait même être influencé par les données extraites du traitement plus précoce du mouvement. Ainsi, plusieurs indices différents doivent être combinés pour la segmentation de la scène visuelle. On pourrait donc abandonner l'idée d'un traitement 'local-to-global' comme le font les machines, au profit d'un traitement basé sur les priorités [Bullier, 2001].

De nombreuses études en psychologie ont permis de mettre en évidence comment les indices bas-niveau peuvent guider les tâches de recherche visuelle. Ainsi, des stimuli qui se différencient par une seule caractéristique ne nécessiteraient que très peu de temps

pour être discriminés. Treisman appelle cela 'pop-out' dans sa théorie de l'intégration de traits [Treisman and Kanwisher, 1998]. Pour elle, chaque trait de bas niveau serait traité par un module différent. Ce traitement serait parallèle et ne nécessiterait pas d'attention focalisée. Ainsi, tous les indices seraient traités en même temps et la détection de la cible se ferait donc très rapidement pourvu qu'un seul trait différencie les différents stimuli. Ce type de processus serait aussi mis en oeuvre lors de recherche visuelle dans des scènes complexes (par exemple un cheval blanc dans un pré).

En revanche, pour Wolfe et sa théorie de l'attention guidée, il n'existerait pas un module dédié à chaque trait au sein de la voie visuelle, mais plutôt un système de guidage attentionnel qui tiendrait compte de ces informations. Il a démontré que les indices pouvant guider cette attention seraient par exemple la couleur, le mouvement ou l'orientation [Wolfe and Horowitz, 2004]. La forme, quant à elle, ferait partie des indices pouvant avoir un effet similaire mais de façon moins nette. Un système de guidage formé par les interactions entre processus ascendants et descendants sélectionnerait alors la région du champs visuel la plus activée.

### 1.3 Les saccades oculaires

La plupart des études sur la catégorisation rapide utilise un paradigme go/no-go manuel, entraînant la prise en compte d'une réponse motrice assez longue. Afin de se rapprocher au plus près des latences du système visuel en réduisant ce temps de réponse motrice, des études du même type peuvent être réalisées en enregistrement de saccades oculaires. Ici, il n'est plus question de catégorisation d'une image, mais de discrimination entre deux images affichées sur l'écran. À l'aide de cette méthode, on observe des TR très courts d'environ 120 ms (contre les 400 ms en manuel) dans une tâche 'Animal vs non-Animal' [Kirchner and Thorpe, soumis]. Cette méthode s'avère très performante pour l'étude des temps de traitement dans le système visuel et pourrait permettre de mesurer les temps requis pour extraire des informations à différents niveaux de complexité.

L'utilisation de cette méthode pour l'étude des latences de discrimination de stimuli ne comportant qu'un indice de bas-niveau pourrait permettre d'évaluer la durée des traitements nécessaires pour extraire chaque type d'indice. On pourrait ensuite comparer les résultats obtenus avec ceux requis par une discrimination 'Animal vs non-Animal'. Pour les indices simples, les informations nécessaires à la prise de décision étant présentes dès les premiers niveaux de traitement, on peut supposer que la complexité de ces tâches est moindre et sera donc plus rapide.

Dans ce rapport, nous allons donc présenter les résultats d'une expérience consistant en

une tâche de discrimination cible/distracteur réalisée sur 10 sujets. Les temps de réaction enregistrés sont soit saccadique, soit manuel. Ceci permettra de comparer plus précisément ces deux types de réponse. Nous présenterons donc les données comportementales pour ces deux modalités.

## 2 Méthodes

### 2.1 Stimuli

L'expérience principale a porté sur 5 types de stimuli différents : couleur, forme, orientation, mouvement et scènes naturelles. Un des principaux objectifs, dans la définition des stimuli, était d'égaliser leurs saillances. Un certain nombre de paramètres a donc été contrôlé (Figure 1). Toutes les images utilisées sont de la même taille (221x221 pixels). Les niveaux de gris moyen sont comparables entre les cibles et les distracteurs, mais aussi entre les conditions. Pour induire une certaine variabilité de contraste local dans les images simples nous avons ajouté des stimuli plus petits dedans.

Les différentes valeurs de luminance ont été mesurée à l'aide d'un Photomètre. L'objectif étant d'équilibrer la luminance des stimuli à l'intérieur et entre les différentes conditions, celui-ci pouvant influencer les temps de réaction dans ce type de tâche [Adams and Mamassian, 2004]. Il a ainsi été mesuré, sur l'écran servant à l'expérimentation, une luminance de 19 candela/ $m^2$  aux endroits précis d'apparition des images pour un gris moyen [128 128 128]. Les différentes images en noir et blanc ont ensuite été créées (ou choisies) afin qu'elles aient une valeur moyenne en niveau de gris proche de 128.

**couleur** Les deux images de la condition couleur sont un rond rouge et un rond vert. L'équiluminance vert/rouge étant variable entre les sujets en fonction de leur ratio de cône L ou M [Gunther and Dobkins, 2002], la mise en place d'un protocole supplémentaire a donc été nécessaire.

Ainsi, à l'aide du Photomètre, il a été déterminé que le rouge correspondant à une luminance de 19 candela/ $m^2$  était le rouge de valeur RGB [201 106 106]. Cette valeur fixée, il restait à faire varier le vert pour trouver une valeur propre à chaque sujet. Cette équiluminance a été déterminé à l'aide d'un Heterochromatic Flicker Photometry (HFP) [Gunther and Dobkins, 2002]. A chaque essai, un carré clignote entre le rouge et le vert (distance par rapport à la croix de fixation et taille sont les mêmes que dans l'expérience). Le sujet doit ensuite ajuster la luminance relative entre les phases de rouge et de vert jusqu'à ce que sa perception du clignotement soit la moins saillante. Une série de 6 essais est réalisée du côté gauche, une autre du côté droit. Le point d'équiluminance est ensuite déterminé par la moyenne des six essais. Pour les dix sujets, les valeurs données à

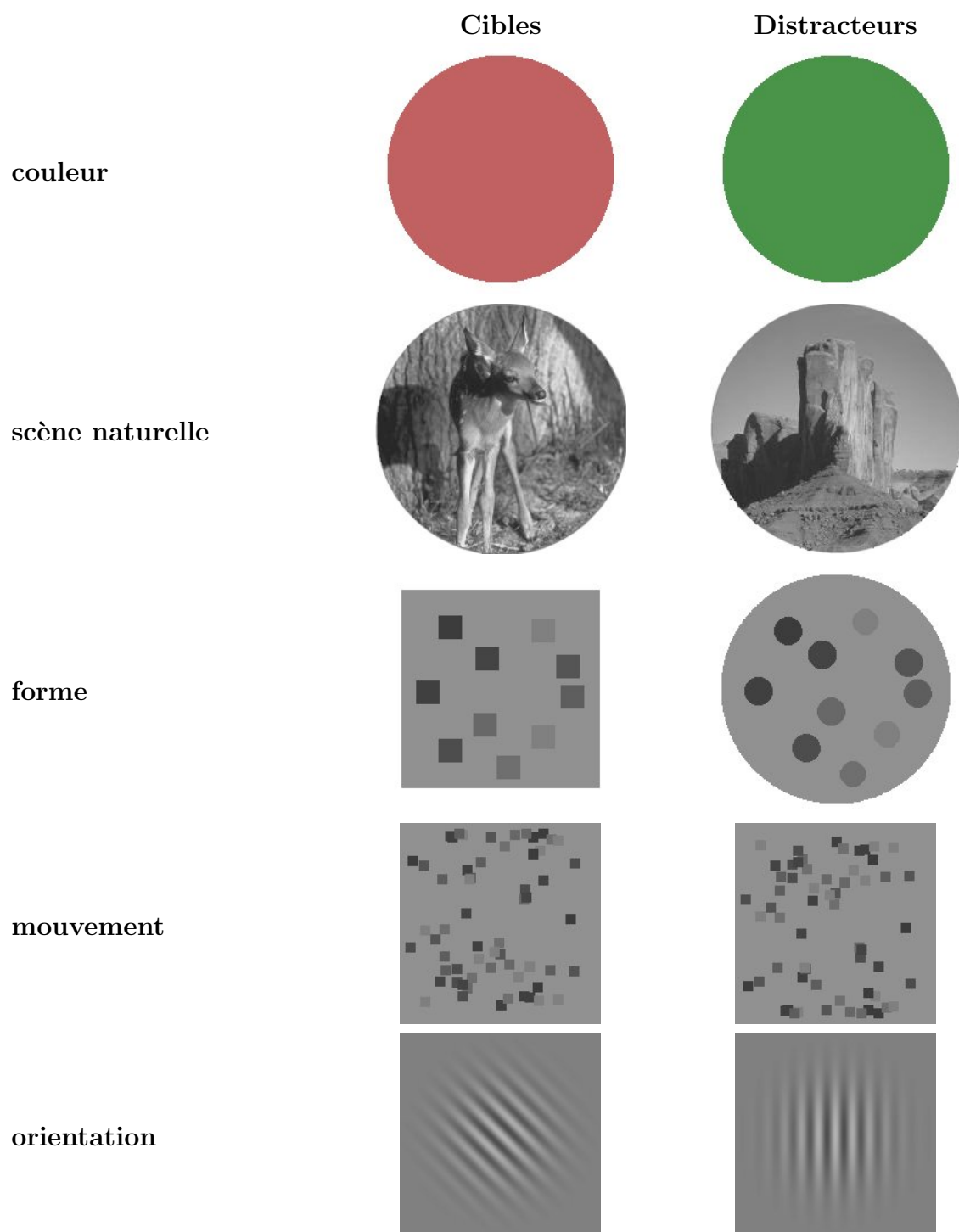


FIG. 1: Ensemble des stimuli présentés classés par condition.



gauche et à droite ont été identiques, une image verte a donc été créée pour chaque sujet.

Caractéristiques : Rouge = (201 106 106)

Vert (valeur moyenne sur l'ensemble des sujets) = (74 147.8 74)

**scènes naturelles** Les deux scènes naturelles ont été sélectionnées, parmi un large panel d'images, telles que leurs histogrammes soient assez similaires, tout en restant proches du gris moyen, et que l'on puisse le moins possible s'aider de traits de bas niveau pour les différencier.

Biche :  $V_{\min} = 22$ ,  $V_{\max} = 252$ ,  $V_{\text{moy}} = 121$

Pierre :  $V_{\min} = 35$ ,  $V_{\max} = 216$ ,  $V_{\text{moy}} = 128$

**forme** On a ici un carré et un rond. Plusieurs carrés ou ronds de plus petite taille ont été insérés à l'intérieur. Ils ont été créés de manière à ce que leurs surfaces soient identiques.

Caractéristiques : Fond = 128

Forme globale = 145

Petites formes = 60, 64, 68, 77, 85, 94, 104, 111, 128, 128

**mouvement** Ces stimuli consistent en une succession de 20 images (chaque image étant présentée durant 20 ms). Les points en mouvement ont les mêmes valeurs de gris que les petits carrés ou ronds de la condition forme.

Caractéristiques : Fond = 145

Vitesse = 9 pixels/images

**orientation** Les images sont ici des patches Gabor.

Caractéristiques : Valeur minimum = 64

Valeur maximum = 191

Valeur moyenne = 128

Inclinaison =  $0^\circ$  ou  $45^\circ$

Fréquence spatiale = 14.7 cycles/images

Une des particularités de cette expérience, en comparaison avec celles de même type menées précédemment [Kirchner and Thorpe, soumis], est que les images sont affichées durant 400 ms au lieu d'être flashées pendant 20 ms. En effet, il a été observé que cette durée d'affichage plus longue permettait de raccourcir les TR moyens, tout en ayant une distribution de TR plus étroite [Kirchner et al., 2005a].

Une autre particularité est que, pour chaque condition, ce sont toujours les deux mêmes images qui sont présentées, une étant toujours la cible et l'autre toujours le distracteur.

Il a donc fallu s'assurer au préalable que les deux images de chaque condition donnent lieu aux mêmes performances de détection. Dans ce but, une expérience préliminaire a été mise en place.

## 2.2 Expérience préliminaire

Pour vérifier que les deux images d'une même condition conduisent à des performances identiques, cette expérience a été mise en place et testée sur trois sujets.

### 2.2.1 Protocole expérimental

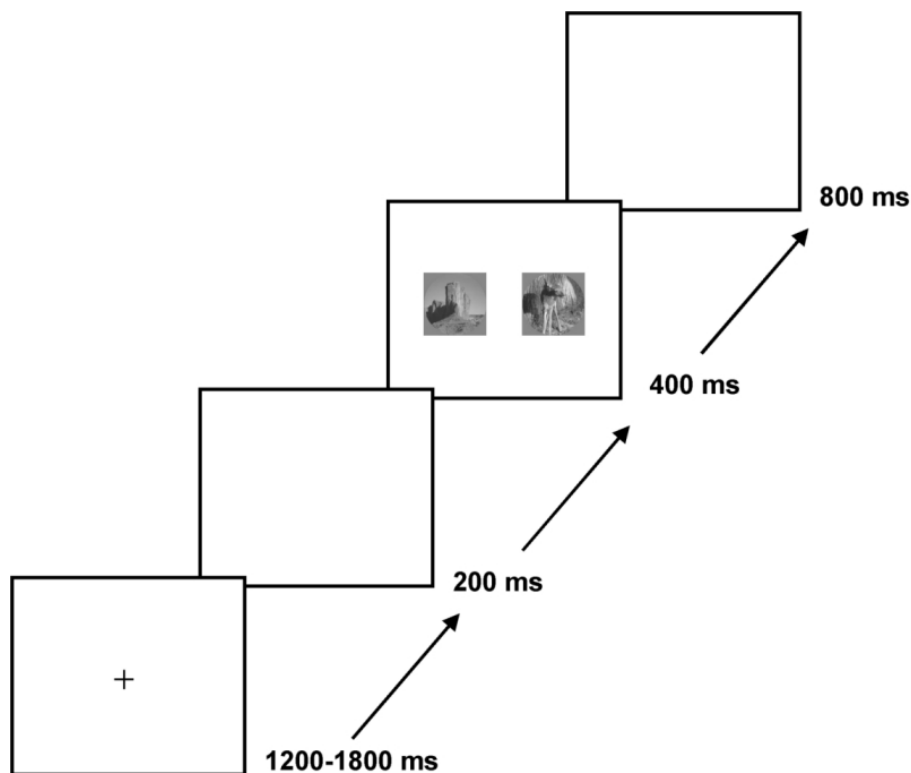


FIG. 2: Déroulement de l'expérience

Une images est présentée à gauche ou à droite (à  $6^\circ$  d'excentricité) de la croix de fixation durant 400 ms (Figure 2). L'hémichamps où l'image apparaît est contrebalancé au sein de chaque condition. La croix de fixation disparaît après un intervalle de temps pseudo-aléatoire (1200-1800 ms), un intervalle de 200 ms est placé entre la disparition de la croix de fixation et l'apparition des images. Cet intervalle est généralement utilisé afin d'accélérer l'initiation des saccades oculaires [Fischer and Weber, 1993]. L'image est ensuite présentée durant 400 ms et le sujet a pour consigne de réaliser une saccade vers le centre de l'image-cible le plus rapidement possible. Les images présentées ici sont les 10

images (2 images x 5 conditions) qui seront utilisées plus tard pour l'expérience principale. Les résultats sont référencés sur la Figure 3.

### 2.2.2 Résultats

CONDITION	CIBLE	SUJET1	SUJET2	SUJET3	MOYENNE
couleur	rouge	133 $\pm$ 3	104 $\pm$ 4	154 $\pm$ 3	132 $\pm$ 13
	vert	148 $\pm$ 7	101 $\pm$ 4	152 $\pm$ 3	133 $\pm$ 14
scène naturelle	animal	119 $\pm$ 2	100 $\pm$ 3	133 $\pm$ 5	116 $\pm$ 10
	non-animal	116 $\pm$ 2	94 $\pm$ 3	132 $\pm$ 4	114 $\pm$ 10
forme	carré	118 $\pm$ 3	99 $\pm$ 3	140 $\pm$ 5	118 $\pm$ 11
	rond	116 $\pm$ 2	98 $\pm$ 3	139 $\pm$ 4	116 $\pm$ 11
mouvement	bas	116 $\pm$ 3	93 $\pm$ 3	145 $\pm$ 4	117 $\pm$ 13
	haut	119 $\pm$ 3	93 $\pm$ 4	138 $\pm$ 4	116 $\pm$ 12
orientation	oblique	134 $\pm$ 3	103 $\pm$ 3	143 $\pm$ 4	127 $\pm$ 12
	droit	140 $\pm$ 3	96 $\pm$ 4	149 $\pm$ 4	128 $\pm$ 15
<b>Moyenne</b>		125 $\pm$ 1	98 $\pm$ 1	143 $\pm$ 1	121 $\pm$ 01

FIG. 3: Résultats de l'expérience préliminaire

On observe globalement une différence entre les conditions. Cependant, une ANOVA pour mesures répétées n'a pas révélé de différences significatives. On n'observe pas non plus de différence si on compare les conditions une à une (par exemple : pas de différence significative entre couleur et mouvement). On n'observe pas non plus de différence à l'intérieur d'une même condition (par exemple : pas de différence significative entre rouge et vert). On peut donc dire que les images sont bien équilibrées à l'intérieur et entre les conditions. Ceci va nous permettre, afin d'avoir plus de données, de ne pas contrebalancer les cibles. Ainsi, pour la condition couleur par exemple, on aura toujours l'image rouge comme cible et l'image verte comme distracteur.

## 2.3 Expérience principale : Tâche de discrimination saccadique

### 2.3.1 Protocole expérimental

Dix volontaires, tous droitiers (âgés de 22 à 34 ans, âge moyen = 25.8, 7 hommes et 3 femmes) ayant une vue normale ou corrigée ont participé à cette tâche de discrimination visuelle.

Le déroulement de la tâche est presque le même que celui de l'expérience préliminaire, en dehors du fait que l'on a cette fois une tâche de discrimination. Deux images seront donc affichées en même temps, et le sujet devra discriminer l'image cible et effectuer une réponse comportementale du côté de la cible (localisation de la cible) (Figure 2). Les deux

images affichées sont toujours les mêmes à l'intérieur d'une même condition. L'une sera toujours la cible, l'autre toujours le distracteur. Cependant, le côté de présentation de la cible sera pseudo-aléatoire à l'intérieur de chaque condition de manière à ce que le sujet ne puisse pas développer de stratégie de réponse.

Au début de chaque session, il est demandé au sujet d'atteindre 75% de bonnes réponses dans chacune des conditions tout en essayant de répondre le plus rapidement possible.

Chaque condition est composée de 3 blocs de 50 essais. Le premier bloc de chaque condition permet au sujet de s'habituer aux stimuli et à la tâche. Il est donc considéré comme un entraînement et ne sera pas retenu dans les résultats. L'ordre de passation des conditions a été contrebalancé entre les sujets et entre les sessions.

### 2.3.2 Matériel

Les participants sont assis sur une chaise. Leur tête est posée sur une mentonnière ajustée de façon à ce que leurs yeux soient au centre de l'écran. Les différentes images servant de stimuli pour l'expérience sont présentées sur un moniteur vidéo (1024 x 768, 100 Hz) sur un fond gris moyen [128 128 128] à une distance de 62 cm de l'écran. Le centre des images apparaît à une distance de 6° de la croix de fixation. La taille des images étant de 221 x 221 pixels, leur taille rétinienne est donc de 7°H x 7.3°V.

### 2.3.3 Enregistrement des réponses

Deux types d'enregistrement des réponses ont été utilisés. Chaque sujet passant les deux types de tâches dans un ordre contrebalancé.

**Réponses saccadiques** La position des yeux est enregistrée par des électrodes EOG horizontales (1 kHz, passe-bas à 90 Hz, passe-bande à 50 Hz, correction de la ligne de base sur l'intervalle  $[-400 : 0]$  ms; NuAmps, Neuroscan Inc.). Les temps de réaction saccadiques (TRS) sont déterminés, après l'enregistrement, comme la différence de temps entre le début de l'image ( $t=0$ ) et le début de la saccade. Pour déterminer cela, le signal différentiel entre les électrodes droite et gauche doit dépasser un seuil d'amplitude de  $\pm 30\mu V$ . Ainsi, le moment du début de la saccade est automatiquement déterminé comme le point d'inflexion du signal étant le plus proche du moment de dépassement du seuil. Chaque essai est ensuite vérifié par l'expérimentateur afin de s'assurer que c'est bien l'inflexion la plus grande (s'il y en a une) qui est considérée comme la saccade.

**Réponses manuelles** Les sujets doivent garder les deux touches 'Ctrl' enfoncées et relâcher celle correspondant au côté où se trouve la cible. Le temps de réponse manuelle

(TRM) correspond donc à la différence entre le début de l’affichage de l’image ( $t=0$ ) et le moment où le sujet relâche la touche.

### 2.3.4 Calcul des temps de réaction minimum

Pour déterminer la valeur des temps de réaction minimum, la distribution des temps de réaction est divisée, pour chaque sujet, en intervalles d’une durée de 10 ms (l’intervalle 120 ms contient les latences allant de 115 ms à 124 ms). On cherche ensuite le premier intervalle qui contient significativement plus de réponses correctes que de réponses fausses (réponse du côté opposé à la cible). Un test du  $\chi^2$  est calculé pour chaque sujet et pour chaque intervalle. Si au moins 10 intervalles consécutifs amènent un test significatif à  $p < .05$ , le premier de ces intervalles est considéré comme le temps minimum. Si le début de la distribution des temps de réaction ne contient pas d’erreurs, le temps minimum est compté à partir du moment où il contient au moins 5 réponses correctes.

### 2.3.5 Tâche contrôle moteur

Cette tâche se déroule comme l’expérience préliminaire : on ne présente qu’une image à la fois au sujet, à droite ou à gauche. Cependant, ici, seules les images cibles (des 5 conditions de l’expérience principale) seront présentées pendant quatre blocs de 60 essais (deux blocs au début de l’expérience et deux à la fin). Ceci sera répété sur l’ensemble des sujets et permettra, tout d’abord, de contrôler le temps nécessaire pour réaliser uniquement la détection de la cible et la réponse motrice, sans tâche de discrimination. Mais elle pourra aussi permettre, puisque deux blocs similaires se déroulent au début et à la fin, d’observer les différences de temps de réaction et de performance indiquant un éventuel effet d’apprentissage pendant la durée de l’expérience.

## 3 Résultats

### 3.1 Tâche contrôle moteur

#### 3.1.1 Comparaison des cibles

Dans cette tâche, où une seule image apparaît à droite ou à gauche, le TRS moyen pour les réponses correctes à travers toutes les conditions et tous les sujets est de 145 ms. Avec des temps allant de 122 ms en moyenne pour le sujet le plus rapide à 183 ms pour le sujet le plus lent. Comme on l’attendait, le taux de réussite est très élevé (plus de 99%) et ne sera donc pas analysé.

On observe globalement une différence significative entre les conditions ( $F_{(4)}=33.9$ ,  $p < 0.001$ ). Si on teste statistiquement les moyennes de SRT pour chaque condition deux à

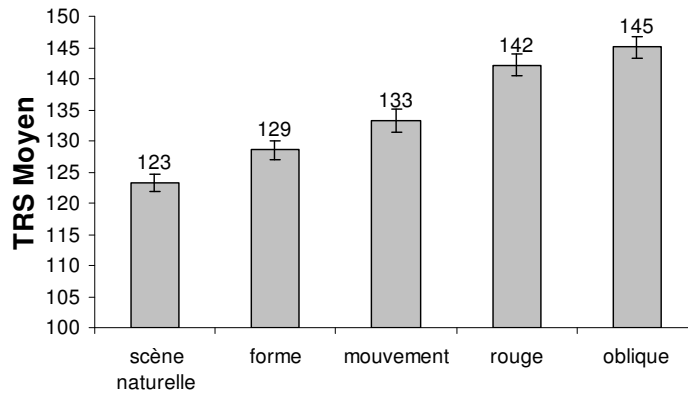


FIG. 4: Tâche contrôle : TRS moyen pour chaque cible

deux, on peut voir, tout d'abord, que la cible de la condition 'scène naturelle' entraîne un TRS moyen significativement plus court que tous les autres, ce qui est tout de même très remarquable. Les cibles 'forme' et 'mouvement' formant un groupe en ayant des temps très similaires entre elles. Les cibles des conditions 'orientation' et 'couleur' formant un autre groupe avec des temps un peu plus long pour ces deux conditions.

### 3.1.2 Comparaison des essais du début et de la fin

La tâche contrôle était composée de 4 blocs, deux avant l'expérience principale et deux après. Si l'on compare les résultats entre ces deux groupes, on trouve une différence significative de 20 ms en moyenne ( $F_{(1)}=156.02$ ,  $p<0.001$ ). En effet, la moyenne de détection de l'image avant la tâche de discrimination est de 154 ms, après la tâche de 134 ms. Il est donc, tout d'abord, intéressant de noter que, malgré la longueur de l'expérience, aucun ralentissement qui pourrait être causé par la fatigue n'est observé. De plus, il n'y a pas d'interaction entre cette amélioration des TRS moyens et le type de cible. On peut donc dire que les temps plus courts du deuxième groupe correspondent à un apprentissage moteur global des saccades au cours de l'expérience. De ce fait, le contrebalancement de l'ordre de passation des conditions à l'intérieur d'une session et à travers les sujets prend toute son importance.

## 3.2 Tâche de discrimination saccadique

### 3.2.1 Analyse des temps de réaction et de la précision

Malgré l'apparente simplicité de la tâche, le pourcentage de réussite global est de 76%. Le TRS moyen pour les réponses correctes est de 191 ms. Ce résultat confirme donc que l'on peut faire une tâche de discrimination saccadique en des temps très courts.

Toutefois, ces valeurs varient beaucoup pour chaque sujet, allant de 65% à 92% pour les taux de réponse corrects et de 147 ms à 213 ms pour les TRS moyen (Figure 5). Cependant, la corrélation entre le TRS moyen et la précision par sujet et à travers toutes les conditions était de faible significativité ( $r=0.615$ ,  $p<0.58$ ). Ceci indique donc un faible 'speed-accuracy trade-off'.

Le TRS minimum, le premier intervalle de 10 ms à contenir significativement plus de réponses correctes que de réponses fausses, à travers l'ensemble des sujets et des conditions est l'intervalle de 130 ms (contenant les réponses faites entre 125 et 134 ms).

Sujet	Précision	TRS moyen	TRS minimum
1	79%	211	160
2	76%	199	170
3	65%	147	/
4	67%	213	200
5	83%	199	130
6	85%	185	160
7	92%	209	160
8	66%	165	140
9	69%	161	140
10	82%	210	150

FIG. 5: Performances, TRS moyen et TRS minimum pour chaque sujet. Le TRS minimum n'a pas pu être calculé pour le sujet 3, en effet 3 intervalles consécutifs au maximum ont été trouvés pour ce sujet et ceci ne rentrait donc pas dans les critères.

### Des différences de temps de traitement entre les conditions ?

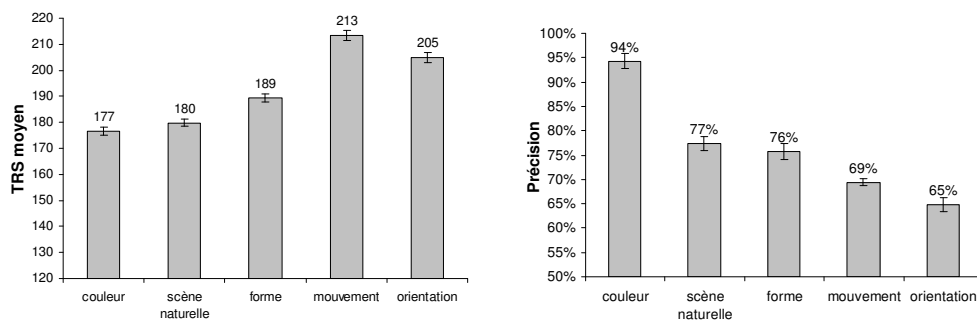


FIG. 6: A gauche : TRS moyen pour chaque condition. A droite : Précision moyenne pour chaque condition

Si l'on compare les conditions deux à deux, on s'aperçoit qu'au niveau des performances, la condition 'couleur' s'avère significativement différente de toutes les autres de part les très bonnes performances qu'elle entraîne (94%). De même pour la condition

'orientation', mais cette fois du fait des mauvaises performances qu'elle entraîne (65%). Les trois autres conditions se situent entre les deux, ne différant pas significativement entre elles. La condition 'mouvement' n'est pas non plus différente de la condition 'orientation'.

Au niveau des temps de réaction moyens, le 'mouvement' se détache des autres de part les temps très longs enregistrés, seule la condition 'orientation' n'étant pas différente. On n'observe pas de différence significative entre les autres conditions.

Comme nous l'avons vu précédemment, il y avait une faible corrélation entre la précision des sujets et leurs TRS sur l'ensemble des données. Cependant, si l'on regarde chaque condition (Figure 7), on trouve une corrélation pour les conditions 'forme' ( $r=0.889$ ,  $p<0.01$ ), 'orientation' ( $r=0.669$ ,  $p<0.05$ ), et une faible corrélation pour 'scène naturelle' ( $r=0.560$ ,  $p<0.09$ ). Il semblerait que l'on ait donc un phénomène de 'speed-accuracy trade-off' sur ces conditions. Ce qui veut dire que les sujets ayant des stratégies de réponses rapides ont tendances à faire plus de faute et qu'à l'inverse, les sujets répondant plus lentement font moins d'erreurs.

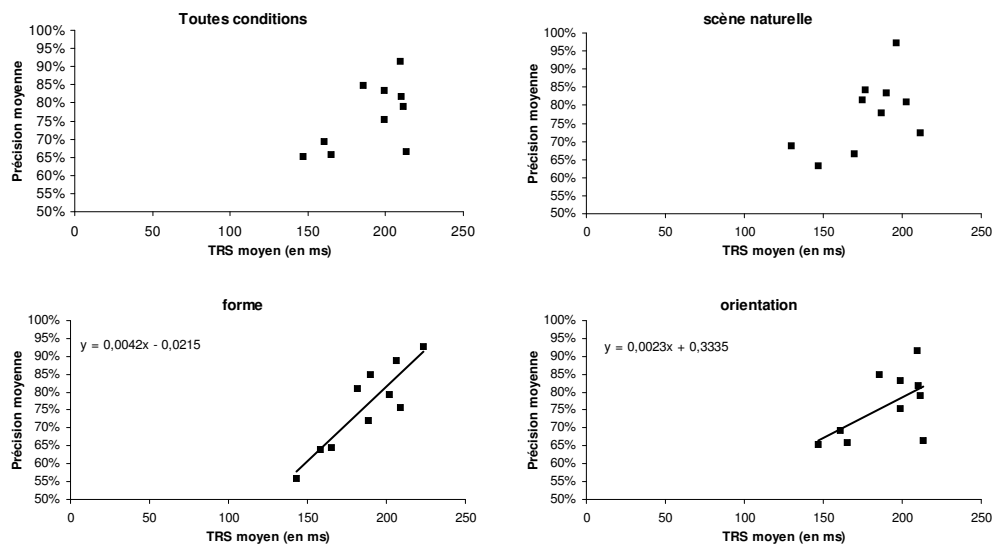


FIG. 7: Précision moyenne en fonction des TRS moyens pour chaque sujet. Le premier graphique (en haut à gauche) se base sur l'ensemble des données. Les autres sont des cas particuliers de 3 conditions.

### 3.2.2 Distributions de temps de réaction

Comme on le voit sur la Figure 8, les distributions de TRS selon les conditions sont très différentes. Si on prend les données à travers toutes les conditions, le moment précis où les réponses correctes prennent le pas sur les réponses incorrectes est bien net (TRS minimum = 130 ms). Ceci s'observe aussi pour les conditions 'scène naturelle' et 'mouvement' (TRS



minimum = 150 ms et 180 ms). La 'couleur', quant à elle, a engendré tellement peu de réponses incorrectes que la distribution de celle-ci décolle sur pratiquement toute la durée des réponses (TRS minimum = 130 ms). Pour finir, on peut penser que la significativité de l'intervalle 220 ms pour l'orientation est causée par le petit pic décrit par la courbe aux alentours de cet intervalle. Cependant, pour cette condition, le moment où les sujets commencent à discriminer réellement la cible du distracteur est beaucoup moins clair que dans toutes les autres conditions.

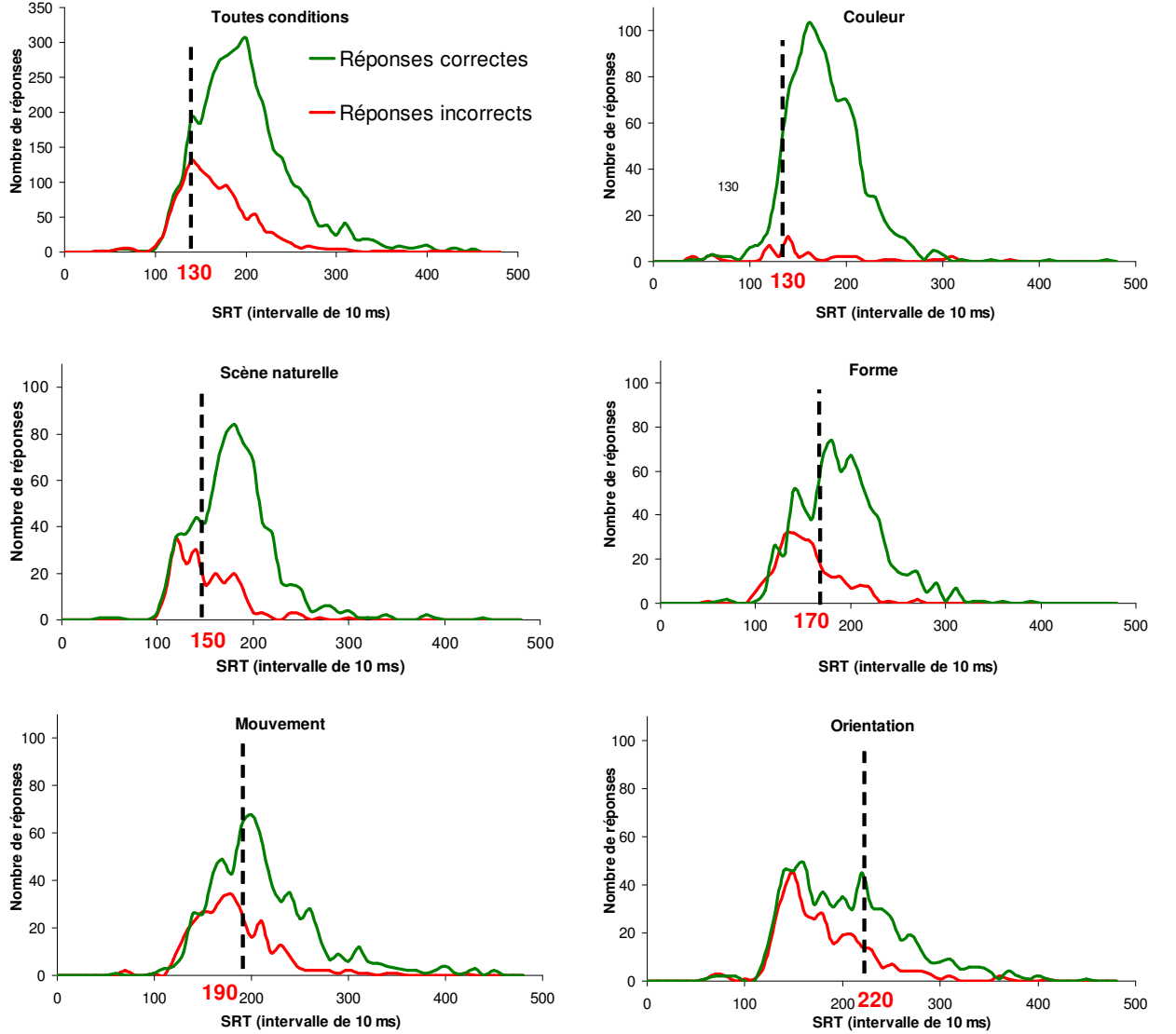


FIG. 8: Distribution des réponses correctes et incorrectes pour chaque intervalle de 10 ms, sur l'ensemble des données puis pour chaque condition

Si maintenant, on observe l'ensemble des distributions des temps de réactions cumulés après avoir soustrait les réponses incorrectes des réponses correctes (Figure 10), on voit bien se dessiner un ordre de traitement de chaque indice dans le temps. La couleur étant

Condition	TRS minimum	Nombre d'occurrences consécutives
couleur	130	15
scène naturelle	150	11
forme	170	13
mouvement	190	9*
orientation	220	7*

FIG. 9: TRS minimum pour chaque condition.\*Conditions pour lesquelles le seuil de 10 intervalles significatifs consécutifs n'a pas été atteint. Toutefois, ce nombre d'occurrences consécutives est assez proche de 10 pour qu'il soit intéressant de retenir ces TR minimum.

disponible bien avant les autres et de façon claire. La forme arrivant en second, puis le mouvement et l'orientation bien plus tard. On peut remarquer que l'information sur le contenu de la scène complexe serait disponible à peu près en même temps que celui de la forme, mais surtout avant ceux du mouvement et de l'orientation.

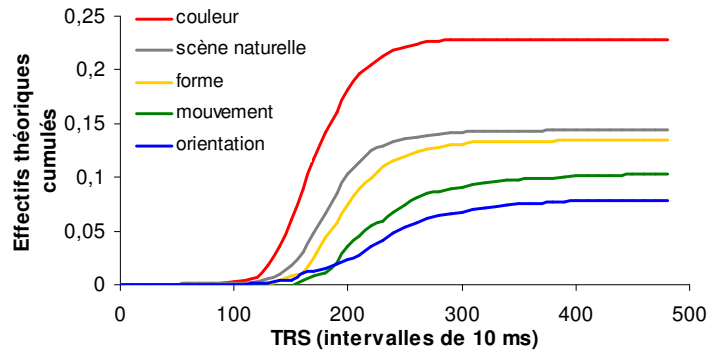


FIG. 10: Distribution cumulée de temps de réaction (chaque pas de temps contient l'ensemble des réponses faites précédemment) de la différence entre réponses correctes et incorrectes par intervalle de 10 ms.

### 3.2.3 Temps supplémentaire de discrimination

Maintenant que l'on a des données sur le temps nécessaire pour discriminer deux images sur un indice simple mais aussi des données sur le temps nécessaire pour simplement détecter cette même image, il pourrait être intéressant de regarder les différences entre ces deux mesures (Figure 11). Cette différence, pour chaque condition, serait le temps supplémentaire pour faire le choix entre les deux images sur la base de l'indice étudié. Globalement, ces différences de temps sont significativement différentes entre elles ( $F_{(4)}=4.761$ ,  $p<.01$ ). Comme on le voit sur la Figure 11, la couleur est l'indice le plus rapidement discriminé. Ce traitement se ferait en environ 35 ms. La forme et la scène naturelle auraient besoin d'environ 20 ms de plus. Il est à noter que l'orientation se situe

dans les mêmes temps moyens que la forme et la scène naturelle mais que les différences entre les sujets sont tellement importantes qu'on ne trouve pas de différence significative entre l'orientation et les autres conditions. Le mouvement quant à lui apparaît comme beaucoup plus lent. Il est à noter encore une fois que la scène naturelle se situe dans des temps comparables à ceux des indices simples.

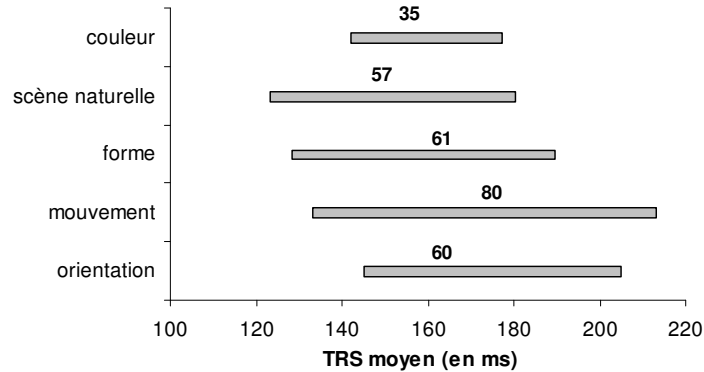


FIG. 11: Différences entre les TRS de la tâche de discrimination et ceux de la tâche de détection simple (tâche contrôle). Le début de la ligne correspond au temps de réaction moyen nécessaire pour simplement détecter la cible. La fin de la ligne correspond au temps de réaction moyen nécessaire pour discriminer la même cible d'un distracteur.

### 3.3 Tâche manuelle

Durant une autre session, les sujets ont tous passé la même tâche que la précédente mais les réponses étaient cette fois manuelles. Le traitement moteur pour relever le doigt d'un bouton étant plus long que le temps de déclenchement d'une saccade, on s'attend donc à des TR plus longs. C'est en effet le cas, la moyenne des TRM (Temps de Réaction Manuelle) sur les réponses correctes est de 349 ms, soit 158 ms en moyenne de plus que pour les réponses saccadiques. La précision est quant à elle de 91%. Celle-ci est bien supérieure à celle observée en choix saccadique. Encore une fois, ces résultats sont très variables selon les sujets (Figure 12). Les TRM vont de 293 ms pour le plus rapide à 411 ms pour le plus lent, la précision allant de 82% à 98% de bonnes réponses.

A travers toutes les données, le TR minimum est de 210 ms, soit 90 ms de plus que dans la tâche de choix saccadique.

Globalement, on observe une corrélation entre la précision des sujets et leurs TRM moyens ( $r=0.701$ ,  $p<0.05$ ). Comme nous l'avons vu précédemment, ceci s'explique par un phénomène de 'speed-accuracy trade-off'.

Au niveau des TRM moyens, on a globalement une différence entre les conditions ( $F_{(4)}=319.675$ ,  $p<0.001$ ). Si l'on compare les conditions deux à deux (Figure 14), elles

Sujet	TRM moyen	Précision
1	347	84%
2	373	94%
3	293	82%
4	396	95%
5	334	89%
6	352	98%
7	320	92%
8	411	94%
9	301	82%
10	351	95%

FIG. 12: Performances et TRM moyen pour chaque sujet dans la tâche manuelle.

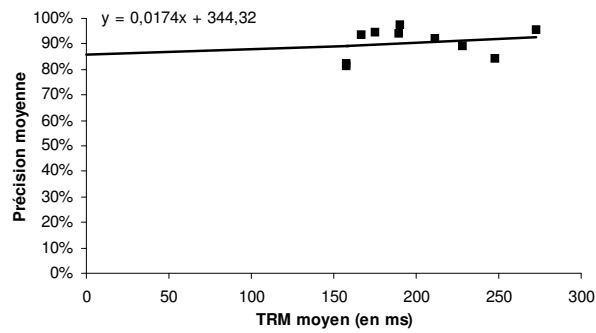


FIG. 13: Précision moyenne en fonction des TRM moyens pour chaque sujet dans la tâche manuelle

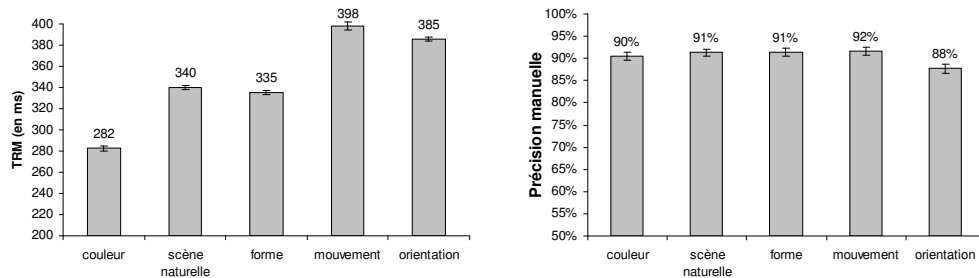


FIG. 14: Tâche manuelle. A gauche : TRM moyen pour chaque condition. A droite : Précision moyenne pour chaque condition

sont toutes significativement différentes entre elles (à  $p < .001$ ) en dehors du couple 'scène naturelle' et 'forme'.

En revanche, en ce qui concerne la précision (Figure 14), les taux de réussite sont très proches et l'on n'observe donc pas de différence statistique entre les différentes conditions. On peut penser que l'on observe ici un effet de plafond du fait de la très bonne précision observée. Ceci est très différent des résultats observés lors de la tâche de choix saccadique où l'on observe une forte chute de la précision sur les conditions 'orientation' et 'forme'.

## 4 Discussion

### 4.1 Vers une hiérarchie des indices bas-niveau ?

Les résultats obtenus lors de cette étude montrent que des différences existent, au niveau des temps de traitement et des performances, entre les différents indices de bas-niveau.

Ainsi, la couleur serait l'indice qui permettrait la discrimination la plus rapide entre deux stimuli. Cette observation est valable pour les deux types de réponses étudiées ici (saccadique ou manuelle). Ceci est plutôt étonnant, compte tenu de nos connaissances sur la physiologie du système visuel. Ainsi, du fait de son passage par la voie parvocellulaire réputée pour être plus lente que la voie magnocellulaire, le traitement de la couleur devrait être plus lent que celui d'autres attributs comme la forme ou le mouvement. De plus, il a été mis en évidence que la couleur n'était pas un indice nécessaire à la catégorisation visuelle rapide dans une tâche 'Animal vs non-Animal'. Cependant, une étude récente a montré la possibilité pour des patients atteints de vision aveugle, ayant des lésions au niveau de l'aire V1, de discriminer des stimuli colorés. On peut donc imaginer que, puisque l'information de la couleur serait présente dès les corps geniculés latéraux, il pourrait exister des connexions entre ce CGL et l'aire V4. Celle-ci étant connue pour être connectée avec le FEF (Frontal Eye Field) impliqué dans le déclenchement des saccades, on peut faire l'hypothèse qu'un chemin de ce type pourrait permettre un traitement aussi rapide.

Les résultats obtenus pour la condition 'forme' par rapport à la condition 'scène naturelle' sont assez différents de ceux obtenus lors d'expériences antérieures (Fize [2000], Aubertin et al. [1999]). On trouve ici un avantage de la scène naturelle sur la forme tandis que l'étude de Fize montrait l'inverse. Cependant, si on prend en compte la différence de temps de traitement entre une simple détection et une discrimination (Figure 11), la forme serait alors traitée dans des temps similaires à ceux de la scène naturelle. Ces différences de résultats peuvent s'expliquer par le fait que le type de réponse n'était pas le même.

Ensuite, cette expérience consistait en une discrimination et non pas en une catégorisation.

Les performances observées pour la condition 'mouvement' sont assez étonnantes. En effet, celui-ci serait notamment traité plus lentement que la forme. Cependant, d'après nos connaissances de la physiologie du système visuel, les informations relatives au mouvement transiteraient par la voie dorsale, et en particulier l'aire MT, de manière très rapide. Ces indices pourraient même être les premiers disponibles [Bullier, 2001]. Une étude récente aurait même montré que les indices de mouvement joueraient un rôle dans la catégorisation rapide des scènes naturelles mais amèneraient des temps de réaction plus longs que pour une présentation statique de la même scène [Kirchner et al., 2005b]. Dans la présente étude, nous confirmons ce résultat pour des stimuli beaucoup plus simples. Cependant, contrairement aux autres conditions étudiées ici, la perception du mouvement ne peut pas être immédiate. Nous avons ici un affichage de 20 images successives en 400 ms pour créer la perception du mouvement. Si l'on considère que 2 images au minimum sont nécessaires pour que les objets aient bougé et que donc le mouvement soit perçus, on pourrait se demander si en augmentant la fréquence d'affichage des images on arriverait à des temps de réaction plus courts. De plus, toujours concernant cette même condition, il a été démontré que la perception du mouvement serait facilitée par la similarité des différentes entités en mouvement [Croner and Albright, 1997]. Les stimuli utilisés dans notre étude étaient tous de la même forme mais on avait fait varier leurs teintes afin de créer une variabilité générale plus importante. Ceci a peut-être aussi participé à diminuer la vitesse de discrimination dans cette condition.

## 4.2 Une scène complexe traitée très rapidement

D'après nos résultats, une discrimination 'Animal vs non-Animal' ne prendrait pas plus de temps qu'une discrimination basée sur des indices de type forme ou mouvement, et même moins de temps qu'une discrimination basée sur des indices d'orientations. De plus, si l'on étudie précisément les "temps de discrimination", on peut voir que ceux-ci sont très similaires. Seule la couleur est discriminée en des temps significativement plus courts.

Si l'on compare les résultats de cette condition à ceux obtenus lors d'une autre étude dans une tâche similaire, on peut observer que les temps de réaction moyens pour la discrimination de l'animal sont du même ordre. Cependant, les TR minimum sont assez différents (20 ms plus long dans notre étude). Ce ralentissement est certainement causé par les nombreuses erreurs tardives observées dans notre étude. Ce phénomène pourrait être dû au fait que les deux images étaient tout le temps les mêmes, cet effet ayant été observé dans une tâche de catégorisation. Il pourrait donc être intéressant de réaliser une

expérience où les stimuli seraient beaucoup plus variables. Cependant, une donnée intéressante dans cette étude est que la scène naturelle a été l'image amenant les TR les plus courts durant la tâche de détection simple. Cependant, ceci pourrait être dû au fait que la variabilité de contrastes était beaucoup plus importante dans cette image que dans les autres conditions.

Les résultats de notre étude peuvent amener à penser que le traitement des indices de bas-niveau dans une scène naturelle se ferait de manière parallèle, les indices seraient tous traités en même temps. Le fait qu'ils soient combinés dans une image complexe, de ce point de vue, ne conduirait pas à une augmentation des temps de discrimination. De plus, on peut penser que la rapidité de discrimination de deux images complexes serait liée au fait que ces images contiennent plusieurs indices. Ces différents indices amèneraient tous une information. Il suffirait ensuite d'avoir extrait assez d'information de ces indices pour que le choix puisse se faire. On pourrait faire ici l'analogie avec la "sommation spatiale" du neurone. Plusieurs indices simultanés lui feraient dépasser le seuil de "discrimination" de la même façon qu'un "gros" indice suffisant à lui seul à lui faire dépasser son seuil. Il n'y aurait pas alors de différence de latence entre ces deux processus pour dépasser le seuil qui permettrait le choix. Cependant, on peut penser que ce mécanisme décisionnel pourrait fonctionner de manière plus pertinente. Ainsi, les indices les plus appropriés à la tâche demandée pourraient être sélectionnés en priorité. Le choix se ferait alors dès que l'on aurait extrait assez d'informations grâce aux indices les plus pertinents [Bullier, 2001]. Tout ceci amène à l'idée que la reconnaissance de l'objet dans la scène pourrait se faire sans phase de segmentation, ou bien avant celle-ci.

### **4.3 Une scène simple traitée très lentement**

Les résultats obtenus pour la condition 'orientation' sont sans conteste les plus surprenant de cette étude. Une discrimination de deux images basée sur ces indices entraîne des temps de réaction très longs en comparaison avec d'autres indices comme la couleur et la forme. En plus de ces TR allongés, on observe une chute des performances par rapport aux autres conditions en réponse saccadique. Il est intéressant d'observer que cette chute n'a pas lieu en réponse manuelle. Au niveau des TR minimum, celui obtenu pour cette condition (220 ms) est supérieur au TR moyen (205 ms), ce qui paraît plutôt surprenant. Ce résultat met en évidence le fait que le TR minimum pourrait correspondre véritablement au moment exacte où la discrimination a été faite, contrairement au TR moyen qui prendrait peut-être trop en compte les réponses 'anticipatoire' qui aurait autant de chance d'aboutir du côté correct qu'incorrect. Cependant, on sait que les informations nécessaires à une tâche de discrimination d'orientations sont disponibles dès le départ du traitement

corticale du flux visuel dans l'aire V1. Nos résultats indiquent donc que le traitement de l'orientation en soit n'est pas difficile (réponses manuelles), mais qu'il le devient aussitôt qu'une réponse saccadique est demandée au sujet. Cet aspect est donc un champ à approfondir dans l'avenir.

En effet, la réalisation motrice d'une saccade est beaucoup plus rapidement mise en place qu'une réponse manuelle. Ceci permet d'affiner les observations et de s'approcher au plus près des temps de traitement en éliminant une grande partie de l'aspect moteur de la tâche manuelle. On peut penser que ceci est à l'origine de la meilleure précision dans cette modalité de réponse. De plus, on peut remarquer que la modalité manuelle n'avait pas amené de différence de précision entre les conditions, contrairement à la tâche saccadique où les conditions 'mouvement' et 'orientation' occasionnaient beaucoup plus d'erreurs que les autres conditions. Ceci pourrait venir du fait que les réponses saccadiques n'amèneraient pas de possibilités de contrôle aussi importantes qu'en réponse manuelle. On peut penser que la nécessité pour le cerveau de contrôler les saccades pour ne pas faire d'erreurs n'est pas fondamentale. Contrairement aux mouvements manuelles qui doivent être beaucoup plus contrôlés. Les mouvements oculaires seraient plus "instinctifs" que les mouvements manuels. Ainsi, l'enregistrement des saccades serait en quelque sorte un moyen plus "pur" d'étudier les latences nécessaires au traitement visuel car moins de processus non pertinents viendraient biaiser les données de l'expérimentation.

Une étude récente a étudié l'effet de l'orientation de l'image sur la catégorisation 'Animal vs non-Animal'. Les résultats montrent que cette orientation n'aurait que très peu d'effet sur les performances [Guyonneau et al., 2005]. Ceci montre que l'information sur l'orientation ne gêne pas la performance dans ce type de tâche. Il serait très intéressant d'étudier plus en profondeur le phénomène de chute de performance observé dans notre étude. On pourrait, par exemple, faire varier la nature des cibles, en testant, une image naturelle à 0° face à cette même image à 45°, un signe "plus" face à un signe "multiplié" pour essayer de déterminer si ce déficit de performance en cas de réponse saccadique s'applique à tout type de cible orientée.

La méthode d'enregistrement des réponses saccadiques paraît donc très adaptée à l'étude des aspects temporels de la vision et il pourrait être intéressant de l'utiliser pour étudier plus précisément le rôle du distracteur dans une tâche de détection d'un indice bas-niveau face à l'absence de cet indice. Par exemple, un rond coloré par rapport à un rond gris, une forme géométrique face à une forme aléatoire, un stimulus statique face à un dynamique, etc. De plus, ce type d'expérience permettrait de faire varier les stimuli à l'intérieur d'une même condition.



## Références

- W. J. Adams and P. Mamassian. The effects of task and saliency on latencies for colour and motion processing. *The Royal Society*, 271 :139–146, 2004.
- A. Aubertin, M. Fabre-Thorpe, N. Fabre, and G. Géraud. Catégorisation visuelle rapide et vitesse de traitement chez le migraineux. *C. R. Acad. Sci. Paris, Sciences de la vie / Life Sciences*, 322 :695–704, 1999.
- J. Bullier. Integrated model of visual processing. *Brain Research Reviews*, 36 :96–107, 2001.
- J. Bullier. Hierarchies of cortical areas. 36 :181, jun 2003.
- L. J. Croner and T. D. Albright. Image segmentation enhances discrimination of motion in visual noise. *Vision Research*, 37(11) :1415–1427, 1997.
- A. Delorme, G. Richard, and M. Fabre-Thorpe. Ultra-rapide categorisation of natural scenes does not rely on color cues : a study in monkeys and humans. *Vision Research*, 40 :2187–2200, 2000.
- B. Fischer and H. Weber. Express saccades and visual attention. *Behavioral and Brain Sciences*, 16 :553–610, 1993.
- D. Fize. *Bases cérébrales de la catégorisation visuelle rapide, étude chronométriques et fonctionnelles*. PhD thesis, 2000.
- K. L. Gegenfurtner and J. Rieger. Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, 10 :805–808, 2000.
- K. L. Gunther and K. R. Dobkins. Individual differences in chromatic (red/green) contrast sensitivity are constrained by the relative number of l- versus m-cones in the eye. *Vision Research*, pages 1367–1378, 2002.
- R. Guyonneau, H. Kirchner, and S. J. Thorpe. Animals roll around the clock : the rotation invariance of ultra-rapid visual processing. *Perception*, 2005.
- H. Kirchner, D. Kerzel, and S. J. Thorpe. First saccades land close to the animal in the scene. In *European Conference on Eye Movements*, Bern, Suisse, 2005a.
- H. Kirchner and S. J. Thorpe. Ultra-rapid object detection with saccadic eye movements : Visual processing speed revisited. soumis.

- H. Kirchner, Q. Vong, S. J. Thorpe, and I. M. Thornton. Moving objects in ultra-rapid visual categorisation result in better accuracy, but slower reaction times than static displays. In *Tübingen Perception Conference*, number 8, page 132, 2005b.
- M. Livingstone and D. Hubel. Segregation of form, color, movement, and depth : Anatomy, physiology, and perception. *Science*, 240, 1988.
- S. J. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381, jun 1996.
- A. Torralba and A. Oliva. Statistics of natural image categories. *Network : Comput. Neural Syst.*, 14, 2003.
- A. M. Treisman and N. G. Kanwisher. Perceiving visually presented objects : recognition, awareness and modularity. *Current Opinion in Neurobiology*, 8, 1998.
- R. VanRullen and S. J. Thorpe. The time course of visual processing : From early perception to decision-making. *Journal of Cognitive Neuroscience*, 13(4), 2001.
- J. M. Wolfe and T. S. Horowitz. What attributes guide the deployment of visual attention and how do they do it ? *Nature Reviews*, 5, jun 2004.