



Année universitaire 2018 - 2019

Étude du traitement des visages à travers des mesures de saccades oculaires

Sous la direction de :

Carole Peyrin, DR CNRS, LPNC
Louise Kauffmann, Post-doctorante, LPNC & GIPSA-lab

Présenté par Sarah Khazaz

Mémoire présenté le 24 Juin 2019, devant un jury composé de :

David Méary, Maître de conférences, LPNC
Nathalie Guyader, Maître de conférences GIPSA-lab
Carole Peyrin, DR CNRS, LPNC
Louise Kauffmann, Post-doctorante, LPNC & GIPSA-lab

Mémoire de Master 2 Recherche Sciences cognitives, mention cognition naturelle et artificielle

Abstract :

Previous work has shown that faces elicit extremely fast saccades. This effect would be mediated by rapid processing of their global information. The present study further examined whether these effects also reflected a holistic processing of faces. We performed two experiments using a saccadic choice task: participants were presented simultaneously with two images and had to perform a saccade toward the target stimulus (face or vehicle). In the first experiment, images were displayed either upright or inverted (upside-down) and on a second experiment, images were spatially scrambled. Indeed, inverted and scrambled conditions being used to disrupt holistic processing of stimuli. Results revealed that while stimuli inversion generally impaired performances (overall higher error rate and longer latencies of saccades toward the target when stimuli were upside-down than upright), saccades toward face targets were still faster and had lower error rates than saccades toward vehicle targets, irrespective of the inversion condition. Furthermore, same face features were targeted in both conditions. Results also revealed that scrambled images which generally impaired performance for objects do not for faces. Overall, these results suggest that faces could be processed early in a holistic way.

Introduction

Les visages sont des stimuli extrêmement saillants et attractifs pour les humains. Dès les premières heures après la naissance, les nourrissons présentent une préférence pour les visages. Plus précisément, il semblerait que ce soit l'organisation de la structure d'un visage qui attire leur regard. En effet, les expériences de Morton et Johnson (1991) ont montré que la simple présentation de formes géométriques agencées de la même manière qu'un visage suffit à capter l'attention des nouveau-nés. Il existerait alors un mécanisme spécifique de reconnaissance des visages basé sur l'organisation structurale des visages. C'est grâce à l'orientation innée et l'exposition répétée vers les visages qu'un second mécanisme de traitement des visages se développerait afin de devenir expert dans la reconnaissance des visages (Mondloch, Le Grand, & Maurer, 2002 ; Nelson, 2001).

Chez l'adulte, de nombreuses études rapportent une supériorité du traitement des visages par rapport à d'autres types de stimuli. Par exemple, des travaux en électroencéphalographie (EEG) rapportent que l'exposition à des images de visages (vs. non-visage) provoque, 170 ms après la présentation du stimulus, une réponse spécifique de plus grande amplitude dans les régions occipito-temporales de l'hémisphère droit (Bentin et al., 1996 ; Xu et al., 2005 ; pour une revue voir Rossion & Jacques, 2008). Plusieurs études en imagerie par résonance magnétique fonctionnelle (IRMf) ont mis en évidence des régions cérébrales sélectives au traitement des visages au sein du gyrus occipital inférieur (qui permettrait l'extraction des traits des visages) et du gyrus fusiforme dans le cortex temporal inférieur, cette dernière région étant plus connue sous le nom de "Fusiform Face Area" (Kanwisher, McDermott & Chun, 1997 ; Haxby et al., 2000 ; Silson, Croen, Kravitz & Baker, 2016). Le cerveau aurait donc développé des zones qui lui permettraient de traiter de façon spécifique les visages en comparaison à d'autres types de stimuli. Plus récemment, des études comportementales en oculométrie ont révélé que lors de l'exploration de scènes, la présence de visages attirait immédiatement le regard des observateurs. De plus, ceux-ci consacraient une grande partie de leur temps d'exploration aux visages (Foulsham, Cheng, Tracy, Henrich,

& Kingstone, 2010 ; Hirvenkari et al., 2013 ; Coutrot & Guyader, 2014 ; Judd, Ehinger, Durand et Torralba, 2009 ; Marat, Rahman, Pellerin, Guyader, & Houzet, 2012 ; Crouzet, & Thorpe, 2011). L'oculométrie apparaît donc comme une technique de choix pour étudier la rapidité de la reconnaissance des visages.

Kirchner & Thorpe (2006 ; voir aussi Crouzet, Kirchner & Thorpe, 2010 ; Crouzet & Thorpe, 2011) ont en particulier développé un paradigme de choix saccadique. Ce paradigme consiste à présenter deux images simultanément (e.g., un visage et un véhicule) de chaque côté de l'écran. Les participants ont pour consigne de diriger leur regard le plus rapidement possible vers l'image contenant un stimulus cible (e.g., le visage) et l'autre image devient donc un distracteur (e.g., véhicule). Afin de réaliser cette tâche, le système visuel doit intégrer simultanément les informations relatives aux deux stimuli, les catégoriser et programmer puis effectuer un mouvement oculaire en direction de la cible. En utilisant ce paradigme, les auteurs ont observé que les latences des saccades en direction de la cible étaient plus rapides lorsque celle-ci était un visage que lorsqu'elle appartenait à une autre catégorie (e.g., un véhicule). De façon surprenante, les latences des saccades vers les visages étaient extrêmement courtes. En effet, les auteurs se sont intéressés au temps de réaction saccadique minimum en fonction de la cible, c'est-à-dire à la latence de saccade minimum à partir de laquelle on observe une plus grande proportion de saccades correctes (i.e. dirigées vers la cible) que de saccades erreur (i.e. dirigées vers le distracteur). Ils ont observé que les latences minimums de saccades vers les cibles visages étaient de l'ordre de 100-110 ms, contre 140 ms vers les cibles d'autres catégories. Ces latences de réponses vers les visages sont à peine expliquées par ce qu'on sait du temps de traitement nécessaire aux différentes aires visuelles pour permettre la reconnaissance, et soulignent donc l'impressionnante rapidité du système visuel pour reconnaître un visage. Autre résultat important dans ce paradigme, les participants faisaient plus d'erreurs de saccades vers le distracteur lorsque celui-ci était un visage que lorsqu'il appartenait à une autre catégorie. Cela suggère qu'il est plus difficile d'inhiber volontairement une saccade vers un visage que vers une autre catégorie de stimuli. Dans l'ensemble, les résultats obtenus avec le paradigme de choix saccadique suggèrent que le traitement des visages influence la programmation des saccades en entraînant des saccades ultra-rapides et presque automatiques vers eux.

Kauffmann et al. (2019) ont très récemment utilisé le paradigme de choix saccadique dans l'objectif de préciser comment le traitement des visages est susceptible d'influencer la programmation des mouvements oculaires. Pour cela, les auteurs se sont intéressés à l'amplitude des saccades, c'est-à-dire la distance entre leur point de départ et d'arrivée. Ce paramètre est connu pour être déterminé avant l'initiation de la saccade et ne peut plus être modifié qu'une fois la saccade initiée. Il constitue donc un bon moyen d'étudier les aspects de programmation des saccades. Les résultats ont montré que l'amplitude des saccades vers la cible était plus petite lorsque celle-ci était un véhicule que lorsque c'était un visage, et cela même lorsque les participants avaient pour tâche d'effectuer leur saccade vers un point précis sur la cible. Ces résultats suggèrent donc que le contenu des stimuli visuels influence la programmation de l'amplitude des saccades lors d'une tâche de choix saccadique. Les auteurs ont proposé que dans le cadre d'une présentation de deux images (un visage et un véhicule) présentées de chaque côté d'un écran, deux programmations en parallèle de saccades seraient mises en compétition (i.e. une saccade dirigée vers l'image de gauche et la deuxième dirigée

vers la droite) et pourraient s'inhiber de façon mutuelle. Les différences d'amplitude des saccades pourraient donc s'expliquer par un poid plus important pour les saccades programmées vers le visage lors de ces processus de compétition/inhibition, ayant pour effet de réduire l'amplitude des saccades vers le véhicule. Ainsi, pour une tâche de choix saccadique, la saillance et la pertinence d'un visage interféraient sur la planification et l'initialisation des saccades vers les véhicules. Cette interférence se traduirait non seulement par une difficulté à inhiber les saccades vers le visage en tant que distracteur, mais aussi par des amplitudes de saccades plus courtes vers les véhicules que vers les visages.

Le biais en faveur des visages observés à travers le paradigme de choix saccadique persiste même lorsque les stimuli sont présentés à de grandes excentricités de 80° (Boucart et al., 2016) ou lorsqu'ils sont filtrés pour ne préserver que les basses fréquences spatiales (BFS; informations globales dont la résolution spatiale est faible) ou hautes fréquences spatiales (HFS; informations détaillées ; Guyader et al., 2017). Cette dernière étude a notamment permis de montrer que les saccades sont initiées plus rapidement et précisément vers des visages filtrés en BFS comparé aux visages filtrés en HFS. Ce résultat vient confirmer la préférence des régions corticales sélectives aux visages pour traiter rapidement leur contenu en BFS (Goffaux et al., 2011). Le recours aux BFS pour la détection rapide des visages est également en accord avec un certain nombre de travaux qui suggèrent que le traitement des BFS permettrait en particulier le traitement de la configuration des visages : deux yeux alignés de façon horizontale, au-dessus du nez, lui-même au-dessus d'une bouche. En effet, la perception des visages est différente de la perception des autres objets en ce sens qu'elle se fait de façon globale ou de façon « holistique », comme le suggère la théorie de la Gestalt (Galton, 1883). Cette théorie appliquée au traitement des visages revient dans la littérature à employer des termes tels que configural ou relationnel pour décrire les caractéristiques émergentes d'un visage qui ne deviennent apparentes que lorsque deux ou plusieurs de ses caractéristiques de base (i.e. les relations spatiales entre les yeux, le nez et la bouche qui forment un visage) sont traitées en même temps. Le traitement holistique d'un visage reviendrait à placer cette configuration de caractéristiques de bases (i.e. les traits du visage) dans un module prédéfini où la représentation holistique des visages est stockée. L'avantage du traitement holistique permettrait de reconnaître très rapidement un stimulus comme un visage avant d'émettre une décision rapide de familiarité du visage perçu (pour revues voir Peterson & Rhodes, 2006 ; Maurer, Le Grand et Mondloch, 2002 ; Mc Kone et al., 2013). C'est au cours du développement et avec l'exposition aux visages que l'expertise se met en place. Les visages représentent une catégorie d'objet où tous les exemplaires sont homogènes. Dans la catégorie des visages, il existe donc un prototype plus fort que pour d'autres catégories d'objet, comme par exemple les véhicules. Par exemple, lors d'une présentation de visage, un traitement de premier ordre va rapidement évaluer la configuration des éléments du visage, puis un traitement de second ordre va émettre grâce à l'expertise un jugement de familiarité en comparant les distances des traits au prototype stocké (Diamond & Carey, 1986). À l'inverse, la reconnaissance d'un type de véhicule serait moins automatique et mettrait en jeu d'autres mécanismes moins rapides. Ainsi le traitement ultra-rapide des visages pourrait en partie reposer sur leur traitement holistique, c'est à dire un traitement comme un tout encodant en parallèle la configuration des différents éléments des visages.

Dans la littérature, une des manières les plus communes de perturber le traitement holistique consiste à retourner les visages. L'inversion ne modifie pas les caractéristiques de bas niveau du stimulus (e.g., spectre d'amplitude, contraste, fréquences spatiales), mais change son orientation canonique. Lorsque des objets sont retournés à l'envers, il devient plus difficile de les reconnaître. De façon importante, les visages seraient des stimuli plus sensibles à cet effet d'inversion (Yin, 1969 ; pour revues voir Rossion & Gauthier, 2002 ; et McKone & Yovel, 2009). L'illusion Thatcher (Thomson, 1980) illustre bien ce phénomène. L'effet Thatcher consiste à retourner des éléments isolés d'un visage, notamment les yeux et la bouche, tout en les replaçant dans le contexte d'un visage. Comme on peut le voir sur la Figure 1, lorsque les éléments modifiés sont intégrés dans le visage de Margareth Thatcher à l'envers, on ne remarque aucune anomalie de la configuration de son visage, mais lorsqu'on replace l'image à l'endroit, l'incohérence est flagrante. Lorsqu'un visage est retourné, il est difficile de s'apercevoir des modifications des éléments locaux puisque notre sensibilité à la configuration des visages serait atténuée. Plus précisément, Rakover et Teucher (1997) avancent l'idée que l'inversion des visages empêche le traitement des éléments isolés du visage, ce qui perturberait le jugement de familiarité. Ainsi un visage à l'envers ne serait pas qualitativement perçu et exploré de la même manière qu'un visage à l'endroit.



Figure 1. L'effet Thatcher consiste à retourner des éléments isolés d'un visage, notamment les yeux et la bouche, tout en les replaçant dans le contexte d'un visage. Ici les éléments modifiés sont intégrés dans le visage de Margareth Thatcher à l'envers (à gauche), puis l'image est replacée à l'endroit (à droite). Figure adaptée de Thompson et al. (1980).

Certains auteurs mettent même en avant le fait que les visages inversés ne correspondent plus à un pattern de visage, du fait que la perception de leur configuration serait alors perturbée (Rossion & Boremanse, 2008). Pour d'autres auteurs, un traitement holistique ne pourra pas être effectué et par conséquent les visages seraient traités comme d'autres objets, de façon analytique, c'est à dire élément par élément (Tanaka & Farah, 1998). D'un point de vue neurofonctionnel, la sensibilité aux visages semble également réduite lorsqu'on les retourne de haut en bas. Des études utilisant l'électroencéphalographie révèlent que le pic d'amplitude négatif (N170) provoqué par des visages inversés dans les régions temporo-occipitales est retardé et moins intense que celui observé pour les visages à l'endroit (Rossion & Gauthier, 2002 ; Rossion et al., 2000). Par ailleurs, les visages à l'endroit

provoquent une plus grande activation au sein de la “Fusiform Face Area” par rapport aux visages à l'envers (Brants et al., 2011 ; Gilaie-Dotan et al., 2010 ; Mazard et al., 2006; Yovel & Kanwisher, 2004, 2005 ; pour revue voir Duchaine & Yovel, 2005). La littérature actuelle sur le traitement holistique se fonde essentiellement sur des données issues d'expériences sur la reconnaissance et non sur la perception rapide des visages. Par conséquent, ces études ne nous permettent pas de rendre compte du traitement très rapide des visages conférés par le traitement holistique, qui est propre aux visages.

Notre étude avait pour objectif d'apporter des informations supplémentaires concernant la manière dont les visages sont perçus par rapport à d'autres autres objets (par exemple, des véhicules) en utilisant des mesures de mouvements oculaires. Plus précisément, nous avons cherché à savoir si le biais en faveur des visages observé dans le paradigme de choix saccadique pouvait être expliqué par un traitement holistique/configural d'un visage ou bien par l'attractivité d'éléments particuliers et saillants au sein du visage. Nous avons donc réalisé deux expériences en utilisant le paradigme de choix saccadique. Nous présentions à des participants deux images simultanément, une à droite et une à gauche. Les participants devaient effectuer le plus rapidement possible une saccade vers l'image contenant une catégorie cible, soit un visage, soit un véhicule. Dans la première expérience, nous avons présenté les images à l'endroit ou à l'envers afin de perturber dans ce dernier cas le traitement holistique des visages. Dans une deuxième expérience, nous avons dégradé la configuration spatiale des visages en créant des images dites “scrambled”, c'est à dire en divisant une image en carrés et en repositionnement les carrés aléatoirement dans l'image. Cette manipulation permettait de perturber le traitement holistique d'un visage, tout en préservant les éléments qui le composent (yeux, bouche ou nez). Nous nous attendions à répliquer les résultats des précédentes études de choix saccadique (Crouzet et al., 2010 ; Crouzet et Thorpe, 2011 ; Boucart et al., 2016 ; Guyader, Chauvin, Boucart et Peyrin, 2017 ; Kauffmann et al., 2019) lorsque les images étaient présentées à l'endroit dans leur version originale (Expérience 1). Plus précisément, nous nous attendions à ce que les saccades soient initiées plus rapidement vers une cible “visage” que vers une cible “véhicule”, mais aussi à une plus grande proportion d'erreurs de saccades en direction des distracteurs “visage” lorsque la cible était un “véhicule”. Par ailleurs, en accord avec les résultats de Kauffmann et al. (2019), nous faisions également l'hypothèse que l'amplitude des saccades vers une cible “véhicule” soit plus courte que l'amplitude des saccades vers une cible “visage”. Sous l'hypothèse que le biais en faveur des visages résulte d'un traitement holistique, nous nous attendions à ce que l'inversion d'un visage (Expérience 1) ou la perturbation de sa configuration spatiale via des images “scrambled” (Expérience 2) ait pour conséquence de réduire ce biais. En d'autres termes, nous nous attendions à une interaction entre le type de cible et la condition de présentation des images sur les différentes mesures (latences, erreur de saccade et amplitude).

Expérience 1

Matériel et méthode

Participants. Vingt-quatre participants (17 femmes et 7 hommes ; moyenne d'âge \pm écart type = 22 ± 3.09 ans, dont 1 gaucher) avec une vision normale ou corrigée de l'Université Grenoble Alpes ont participé à l'expérience pour des points d'examen. Tous les

participants ont signé un consentement éclairé pour participer à l'expérience, qui a été conduite en accord avec le code éthique de l'association médicale mondiale pour les études impliquant l'humain (déclaration d'Helsinki).

Stimuli. Les stimuli utilisés ont été élaborés à partir de 60 images en couleur téléchargées de la base d'image Pixabay sous licence libre Creative Commons (<https://pixabay.com/fr/>). La moitié des images contenaient des visages et l'autre moitié des véhicules (e.g. voiture, train, scooter). Toutes les images avaient une taille de 300 x 300 pixels soit $11.5^\circ \times 11.5^\circ$ d'angle visuel pour une présentation à une distance de 60 cm. Les stimuli ont été soigneusement choisis pour que les visages et les véhicules aient en moyenne la même position spatiale et la même taille dans les images. Pour cela, l'objet principal (visage ou véhicule) de chaque image a été délimité manuellement à l'aide d'un cadre rectangulaire. Nous avons ensuite calculé le centre et la taille du cadre, et formé des paires d'images de visage et de véhicule appariées en fonction de la taille et du centre de ce cadre (visages : largeur moyenne \pm écart type = 212 ± 37 pixels; véhicules : largeur moyenne \pm écart type = 221 ± 33 pixels, $t(29) = 0.89, p = 0.38$; visages : hauteur moyenne \pm écart type; visages 247 ± 34 pixels; véhicules : hauteur moyenne \pm écart type = 230 ± 35 pixels, $t(29) = 1.67, p = 0.11$; visages : moyenne \pm écart-type du centre en X = 151 ± 8 pixels; véhicules : moyenne \pm écart-type du centre en X = 152 ± 10 pixels, $t(29) = 0.36, p = 0.72$; visages : moyenne \pm écart-type du centre en Y = 150 ± 7 pixels; véhicules: moyenne \pm écart-type du centre en Y = 152 ± 9 pixels, $t(29) = 1.07, p = 0.29$). La luminance moyenne et le contraste moyen ("root mean squared contrast" - RMS) de chaque image ont été égalisés à la luminance moyenne et au contraste RMS de l'ensemble des images (luminance moyenne des images = 0.25, contraste RMS moyen = 0.44 pour des valeurs de luminance comprises entre 0 et 1).

Matériel. L'expérience a été construite à l'aide de la Psychtoolbox (Brainard, 1997 ; Pelli, 1997) implantée dans le logiciel Matlab R2018a (MathWorks, Inc., MA, USA). Les stimuli étaient présentés sur un moniteur LCD 16/9 d'une résolution spatiale de 1360 x 768 pixels, un taux de rafraîchissement de 60 Hz et une luminance moyenne de 39 candelas par mètre carré. Les participants ont réalisé l'expérience dans une salle sombre et calme. La distance entre l'œil et l'écran était fixée à 60 cm à l'aide d'une mentonnière. Les mouvements oculaires ont été enregistrés par un oculomètre Eyelink 1000 Plus (SR Research) avec un taux d'échantillonnage de 1000 Hz et une précision spatiale de 0.01° d'angle visuel. Chez tous les participants, uniquement l'œil droit a été utilisé pour l'enregistrement du mouvement oculaire.

Design. Chaque essai se déroulait de la manière suivante: un point de fixation (0.73° d'angle visuel) apparaissait au centre de l'écran pendant une durée aléatoire de 800 ms à 1600 ms, suivi d'un écran gris (gap) de 200 ms. Une image de visage et une image de véhicule étaient ensuite présentées simultanément de chaque côté de la croix de fixation (le centre de chaque image était positionné à 8° du centre de l'écran) pendant 400 ms, suivies d'un écran gris pendant 1000 ms (Figure 2). Nous avons utilisé un paradigme de choix saccadique, adapté des études de Crouzet et al. (2010), Guyader et al. (2017) et Kauffmann et al. (2019). Tous les participants ont participé à deux sessions expérimentales, une session pour laquelle les cibles étaient des visages humains (les distracteurs étaient des images de véhicules) et l'autre pour laquelle les cibles étaient des images contenant des véhicules (les distracteurs étaient des images de visage humain). Les participants avaient pour consigne de diriger leur regard le plus rapidement possible vers l'image contenant le stimulus cible. L'ordre des sessions expérimentales était contrebalancé entre les participants. Pour chaque session expérimentale, les images étaient présentées à l'endroit dans la moitié des essais et à l'envers

dans l'autre moitié. L'ordre de présentation des images à l'endroit ou à l'envers était aléatoire. Chaque image de chaque condition (i.e. images à l'endroit et images à l'envers) était présentée deux fois au cours d'une session, une fois à gauche et une fois à droite de la croix de fixation. Chaque session contenait 120 essais (i.e. 30 paires d'images x 2 conditions de présentation x 2 côtés de présentation) et durait environ 10 minutes. Afin de se familiariser avec la tâche, tous les participants réalisaient avant l'expérience un court entraînement (8 essais) sur des images non présentées dans l'expérience.

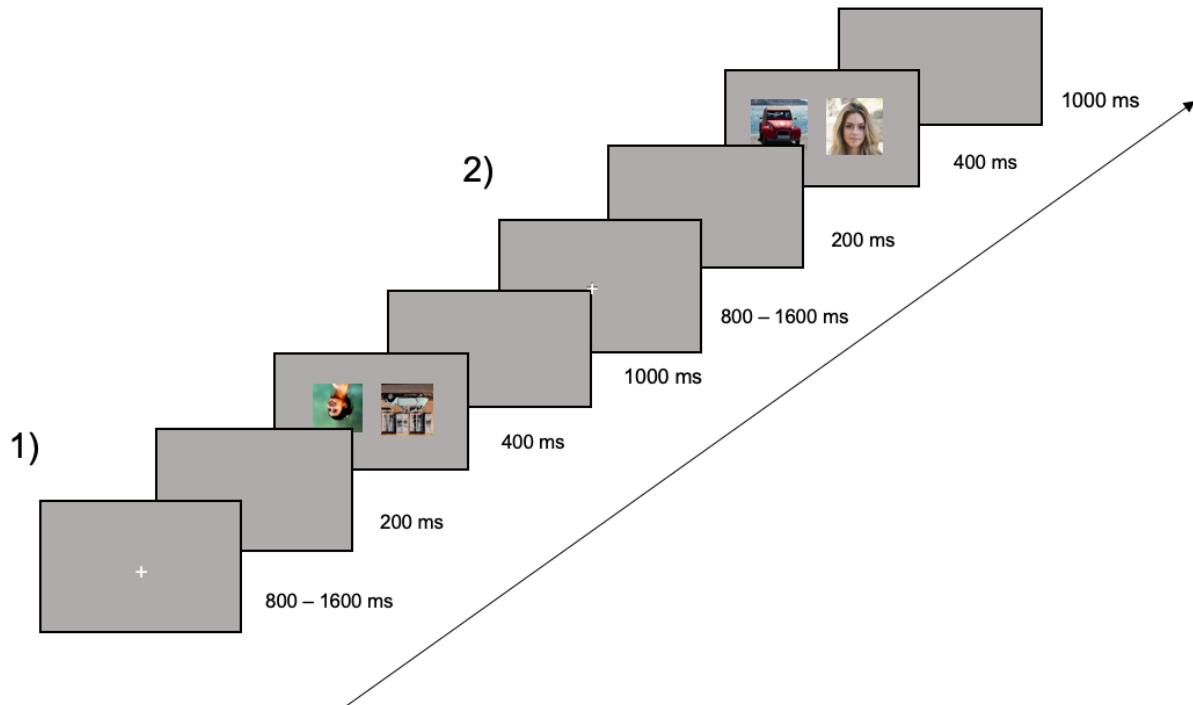


Figure 2. Tâche de choix saccadique. Après une période d'apparition pseudo-aléatoire de la croix de fixation et d'un écran gris (gap) de 200 ms. Les deux images étaient présentées simultanément à gauche et à droite du centre de l'écran (pendant 400 ms). Puis la présentation des images était suivie à nouveau d'un écran gris pendant 1000 ms. Les participants devaient initier une saccade vers le stimulus cible (visage ou véhicule) (1) pour la condition à l'envers ou (2) pour la condition à l'endroit.

Analyse des données. L'exactitude des réponses (taux d'erreur moyen - mER en %), les temps de réaction saccadique (TRS en ms) et les amplitudes verticale (X) et horizontale (Y) (en degrés d'angle visuel) des premières saccades correctes ont été analysés. Afin d'éviter d'inclure dans les analyses les saccades aux profils déviants en termes de latence, d'amplitude ou de durée et susceptibles de ne pas refléter la catégorisation effective des stimuli, seules les saccades satisfaisant certains critères ont été incluses dans les analyses. Les saccades étaient considérées comme valides si : 1 - elles étaient initiées au moins 50 ms après la présentation des images ; 2 - elles avaient une durée inférieure à 100 ms; 3 - leur position de départ se situait dans un carré de 2° d'angle visuel autour de la croix de fixation centrale ; 4 - leur amplitude totale était supérieure à 1° d'angle visuel. Au total 11.95 % des saccades ne satisfaisaient pas ces critères et ont été retirées des analyses.

Les analyses ont été réalisées sur STATISTICA 10.0 ® (Dell Inc., Round Rock, USA). Nous avons effectué des analyses de variance (ANOVA) pour mesures répétées de type 2x2 avec la Catégorie de la cible (visage ou véhicule) et le Type de présentation (à l'endroit ou à l'envers) comme facteurs intra-sujets sur les mER, les TRS et l'amplitude des

saccades en X et Y. Les tailles d'effet ont été calculées en utilisant un eta carré partiel (η_p^2). Le seuil de significativité des tests a été fixé à $\alpha = 0,05$ et les p-values reportées ont été corrigées par une correction de Bonferroni pour les comparaisons multiples.

Résultats.

Exactitude. L'ANOVA réalisée sur les mER indique un effet principal de la Catégorie de la cible ($F_{1,23} = 39.61, p < 0.0001, \eta_p^2 = 0.632$). Les participants faisaient en moyenne plus d'erreurs pour effectuer la tâche lorsque la cible était un véhicule (i.e. plus d'erreurs de saccades vers le distracteur visage ; $21.94 \pm 7.29\%$) que lorsque la cible était un visage (i.e. le distracteur était un véhicule ; $10.87 \pm 8.56\%$). Nous n'avons pas observé d'effet principal du Type de présentation ($F_{1,23} < 1$). Cependant, ce facteur interagissait significativement avec la Catégorie de la cible ($F_{1,23} = 6.72, p < 0.05, \eta_p^2 = 0.226$). Les comparaisons planifiées ont révélé un effet de la Catégorie de la cible quel que soit le Type de présentation. Plus précisément, les participants faisaient plus d'erreurs lorsque la cible était un véhicule que lorsque la cible était un visage pour les images inversées (cible visage $12.37 \pm 8.62\%$; cible véhicule $21.12 \pm 11.78\%$; $p < 0.001$) et pour les images à l'endroit (cible visage : $9.37 \pm 8.05\%$; cible véhicule : $22.75 \pm 11\%$; $p < 0.001$). Par ailleurs, les participants faisaient plus d'erreurs lorsque les images étaient à l'envers qu'à l'endroit lorsque la cible était un visage (à l'envers: $12.37 \pm 8.62\%$, à l'endroit: $9.37 \pm 8.05\%$, $p < 0.01$) mais cette différence n'était pas significative lorsque la cible était un véhicule (à l'envers: $21.12 \pm 11.78\%$, à l'endroit: $22.75 \pm 11.00\%$, $p = 0.74$). L'interaction suggère donc que l'inversion des stimuli impacte davantage les performances lorsque la cible est un visage que lorsque c'est un véhicule.

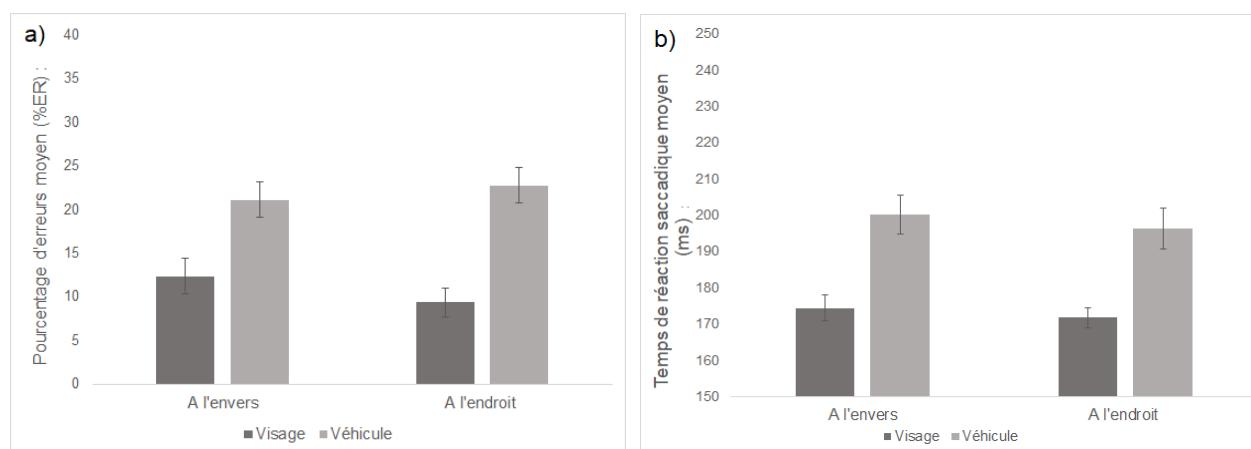


Figure 3. (a) Pourcentage d'erreur (%ER) moyen et (b) Temps de réaction saccadique moyen (ms) en fonction du Type de présentation (à l'envers vs. à l'endroit) et de la Catégorie de la cible (visage vs. véhicule). Les barres d'erreurs correspondent à l'erreur standard.

Temps de réaction saccadique moyen et temps de réaction saccadique minimum.

L'analyse des TRS révèle tout d'abord un effet principal de la Catégorie de la cible ($F_{1,23} = 40.77, p < 0.0001, \eta_p^2 = 0.639$). Les participants étaient plus rapides pour initier une saccade correcte vers les visages (173 ± 16 ms) que vers les véhicules (198 ± 27 ms). L'analyse révèle aussi un effet principal du Type de présentation ($F_{1,23} = 9.49, p < 0.01, \eta_p^2 = 0.292$). Les participants étaient plus rapides pour initier une saccade lorsque les images étaient présentées à l'endroit (184 ± 25 ms) plutôt que présentées qu'à l'envers (187 ± 25 ms). Cependant, nous n'avons pas observé d'interaction significative entre ces deux facteurs ($F_{1,23}$

$<1, p = 0.63$.

Comme dans l'étude de Crouzet et al. (2010), nous nous sommes également intéressés au temps de réaction saccadique minimum à partir duquel la proportion de saccades correctes était significativement plus importante que la proportion de saccades erreur, afin d'examiner si celui-ci différait en fonction de la condition d'inversion. Pour cela, nous avons regroupé les saccades correctes et erreurs en fonction de leur latence. Nous avons ainsi constitué 34 bins de 10 ms (e.g. le premier bin 50 contient toutes les saccades de 51 à 60 ms). Un χ^2 a été calculé pour chaque bin afin de comparer la proportion de saccades correctes et de saccades incorrectes. Si au moins 5 tests subséquents atteignaient le seuil de significativité au niveau $p < 0.05$, le premier de ces bins était considéré comme le TRS minimum. Le bin avec le TRS minimum devait par ailleurs contenir au minimum cinq saccades correctes (Figure 4), (voir Crouzet et al. 2010 ; et Kirchner & Thorpe, 2006 pour plus de détails sur cette analyse).

Ces analyses ont montré que lorsque la cible était un visage, le TRS minimum sur l'ensemble des participants était de 120 ms ($\chi^2 = 36.36, p < 0.01$) pour les images à l'endroit et de 110 ms ($\chi^2 = 41.12, p < 0.01$) pour les images à l'envers. Lorsque la cible était un véhicule, le TRS minimum était de 140 ms pour les deux Types de présentation des images (présentation à l'endroit : $\chi^2 = 10.81, p < 0.01$ et présentation à l'envers : $\chi^2 = 4.80, p < 0.01$).

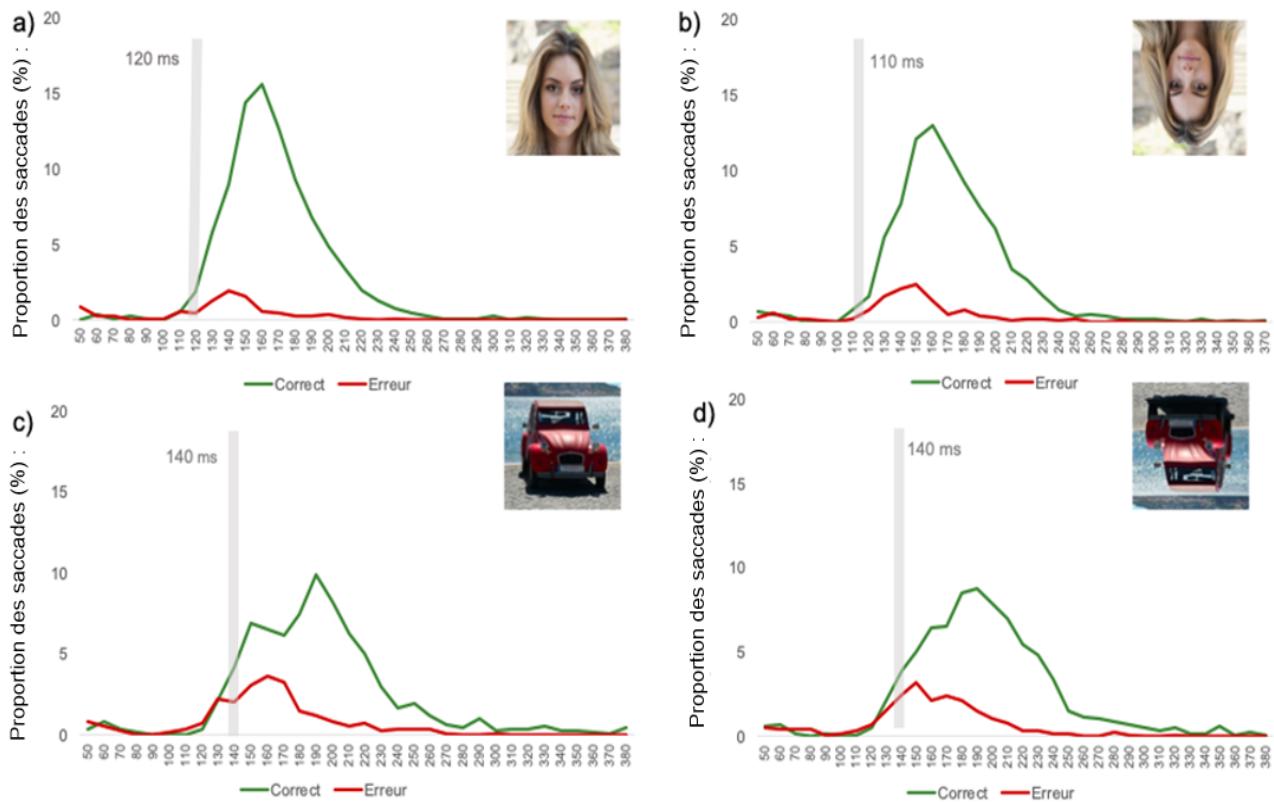


Figure 4. Distribution de la proportion de saccades correctes (en vert) et erreurs (en rouge) pour chaque bin de temps de réaction saccadique pour les cibles visages présentées (a) à l'endroit et (b) à l'envers, et pour les cibles véhicules présentées (c) à l'endroit et (d) à l'envers. Les barres verticales grises représentent le temps de réaction saccadique minimum.

En inversant l'orientation des visages nous avions fait l'hypothèse de perturber le traitement holistique des visages. On remarque cependant que chaque type de cible entraîne une répartition similaire de la proportion de saccades correctes et erreurs en fonction des TRS

entre leur présentation à l'endroit et à l'envers. Ainsi l'analyse des temps de réaction saccadique moyen et minimum ne vont pas totalement dans le sens de notre hypothèse: bien que l'inversion des visages impacte les performances en termes de taux d'erreurs et de temps de réaction saccadique moyen, elle ne perturbe pas significativement les latences minimum d'initiation de saccades.

Dans l'ensemble, ces résultats indiquent que si l'inversion des visages impacte légèrement leur catégorisation, en entraînant plus d'erreurs de saccades, elle n'annule pas totalement le biais vers ces stimuli, puisqu'on observe malgré tout de meilleures performances (en termes de TRS et de mER) lorsque la cible est un visage que lorsque c'est un véhicule, même lorsque les stimuli sont inversés. Ces résultats nous amènent donc à nous demander si les participants utilisent une stratégie pour détecter les visages plus aisément que les véhicules, qui ne serait pas basée sur leur traitement holistique mais sur le traitement d'éléments spécifiques des visages permettant de les détecter plus rapidement.

Afin de répondre à cette question, nous nous sommes donc ensuite intéressés à l'amplitude des saccades, c'est à dire la distance entre leur point de départ et d'arrivée sur l'axe vertical et horizontal. Cette amplitude sur l'axe horizontale (X) nous informe sur le poids relatif des visages par rapport au poids relatif aux véhicules lors de la programmation de la première saccade (Kauffmann et al. 2019). Ainsi, l'amplitude horizontale permet d'inférer si même lorsque les visages sont présentés à l'envers ils ont toujours plus de poids que les véhicules lors de la programmation des saccades.

L'étude de l'amplitude sur l'axe vertical (Y) nous indiquera si une stratégie est utilisée pour chercher un élément particulier dans les visages à présentés à l'endroit ou à l'envers. En effet, si les participants s'appuient plutôt sur la détection d'éléments particuliers des visages pour effectuer la tâche (e.g., les yeux), indépendamment de leur orientation, cela se devrait se refléter sur l'amplitude des saccades en Y, qui devrait différer en fonction de la condition d'inversion.

Amplitude axe horizontal et amplitude axe vertical.

L'ANOVA réalisée sur les amplitudes horizontales (X) moyennes des premières saccades correctes révèle un effet principal de la Catégorie de la cible ($F_{1,23} = 4.48, p < 0.05, \eta_p^2 = 0.163$). Plus précisément, les participants réalisaient une saccade plus petite lorsque la cible était un véhicule ($6.90 \pm 0.72^\circ$) que lorsque c'était un visage ($7.58 \pm 0.57^\circ$). Cependant nous n'avons pas observé d'effet significatif du Type de présentation ($F_{1,23} < 1, p = 0.97$). L'ANOVA révèle aussi un effet d'interaction entre les deux facteurs ($F_{1,23} = 4.48, p < 0.05, \eta_p^2 = 0.163$) (Figure 5). Les comparaisons planifiées ont révélé un effet de la Catégorie de la cible, quel que soit le Type de présentation. Plus précisément, les participants faisaient des saccades horizontales plus courtes vers les véhicules que vers les visages lorsque les images étaient inversées (cible visage $7.51 \pm 2.29^\circ$; cible véhicule $6.98 \pm 2.34^\circ$; $p < 0.001$) ou présentées à l'endroit (cible visage : $7.63 \pm 2.53^\circ$; cible véhicule : $6.88 \pm 2.68^\circ$; $p < 0.001$). De plus on observe un effet significatif du Type de présentation pour la Catégorie de la cible visage ($F_{1,23} = 5.85, p < 0.05$) mais pas pour les véhicules ($F_{1,23} = 1.51, p = 0.46$). Plus précisément, les participants effectuaient des saccades plus courtes quand ils devaient diriger leur regard vers le visage à l'envers ($7.51 \pm 2.29^\circ$), plutôt que vers le visage à l'endroit ($7.63 \pm 2.53^\circ$).

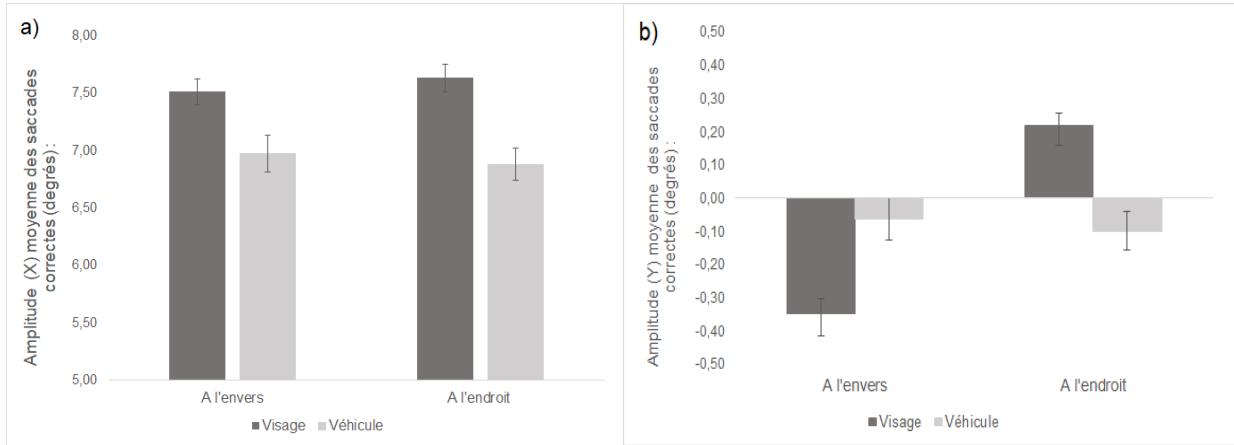


Figure 5. (a) Amplitude (X) moyenne (degrés) (b) amplitude (Y) moyenne (degrés), en fonction de la condition Type de présentation (à l'envers vs. à l'endroit) et la Catégorie de la cible (visage vs. véhicule). Les barres d'erreurs représentent l'erreur standard.

L'ANOVA réalisée sur les amplitudes verticales (Y) moyennes des premières saccades correctes n'a pas révélé d'effet significatif de la Catégorie de la cible ($F_{1,23} < 1, p = 0.71$), mais un effet principal du Type de présentation ($F_{1,23} = 71.13, p < 0.00001, \eta_p^2 = 0.756$). Plus précisément, les participants réalisaient une saccade plus petite lorsque les cibles étaient présentées à l'endroit ($0.06 \pm 0.23^\circ$) que lors d'une présentation à l'envers ($0.20 \pm 0.27^\circ$). L'ANOVA révèle aussi un effet d'interaction entre les deux facteurs ($F_{1,23} = 183.36, p < 0.00001, \eta_p^2 = 0.888$).

Les comparaisons planifiées ont révélé un effet de la Catégorie de la cible, quel que soit le Type de présentation. Plus précisément, les participants faisaient des saccades verticales plus courtes vers les véhicules que vers les visages lorsque les images étaient inversées (cible visage $0.35 \pm 0.22^\circ$; cible véhicule $0.06 \pm 0.32^\circ; p < 0.0001$) ou présentées à l'endroit (cible visage : $0.22 \pm 0.18^\circ$; cible véhicule : $0.10 \pm 0.28^\circ; p < 0.0001$). De plus on observe un effet significatif du Type de présentation pour la Catégorie de la cible visage ($F_{1,23} = 147.28, p < 0.0001$) mais pas pour les véhicules ($F_{1,23} = 1.51, p = 0.21$). Plus précisément, les participants effectuaient des saccades plus courtes quand ils devaient diriger leur regard vers le visage à l'endroit ($0.22 \pm 0.18^\circ$), plutôt que vers le visage à l'envers ($0.35 \pm 0.22^\circ$).

En d'autres termes, les participants effectuaient des saccades vers le haut lorsque les visages étaient présentés à l'endroit et des saccades vers le bas lorsque les visages étaient présentés à l'envers. Ces résultats suggèrent que dans les deux conditions d'inversion (à l'endroit et à l'envers) les participants dirigent leur regard vers le même élément du visage.

Position finale de la première saccade correcte.

Dans le but d'explorer quel était l'élément que les participants étaient susceptibles de fixer dans les visages présentés à l'endroit et à l'envers, nous avons tout d'abord relevé la localisation des points d'impact des saccades des participants (i.e. la position à la fin d'un mouvement oculaire, représentant l'endroit où le participant fixe l'image). Puis nous avons retracé ces points d'impact sur les différents stimuli, vus à gauche, à droite à l'endroit et à l'envers (Figure 6). Sur la base de ces figures on peut remarquer que les impacts des saccades se situent généralement autour de la zone des yeux dans les visages.

Afin de quantifier la nature du lien entre la position finale des saccades et la position des yeux dans les images de visages présentées, nous avons réalisé une corrélation de Spearman entre la distance des yeux dans les visages relativement au point de fixation et

l'amplitude en (Y) des saccades correctes.

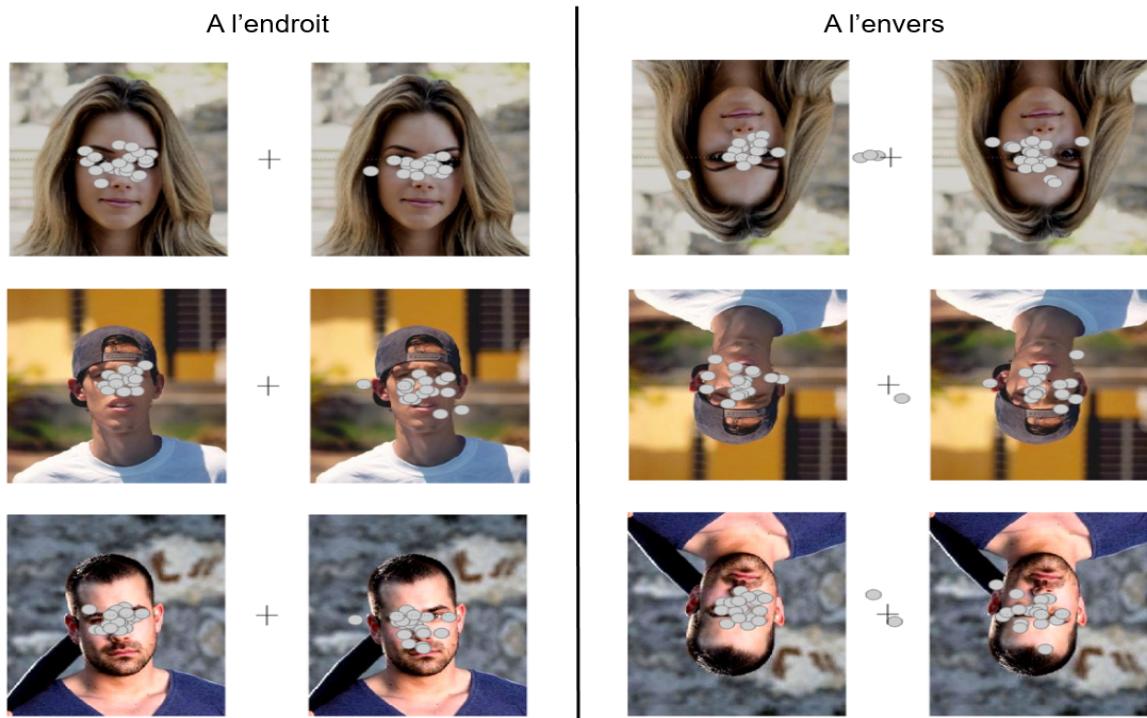


Figure 6. Exemple de positions finales des saccades correctes vers la Catégorie cible visage (présentée à gauche ou à droite) en fonction de la condition à l'endroit (gauche) et à l'envers (droite). Les points gris représentent le point d'impact des saccades correctes de chaque participant.

Corrélation entre la distance des yeux dans les visages relativement au point de fixation et l'amplitude en (Y) des saccades correctes.

Afin de déterminer la position des yeux dans les images de visages, nous avons placé sur chaque image un cadre qui englobait les yeux puis nous en avons calculé son centre (i.e. position du point moyen entre les deux yeux). On obtient ainsi la distance des yeux sur l'axe (Y) vertical, par rapport au centre de l'image (pour plus de clarté nous avons inversé la polarité des données pour la condition à l'envers). Ces calculs nous ont ainsi permis de tester les corrélations (corrélation de Spearman) entre la distance relative des yeux dans les visages par rapport au centre de l'écran et les amplitudes en Y des saccades correctes pour chaque type de présentation. Les résultats ont révélé une forte corrélation positive pour les visages présentés à l'envers ($r = 0.84, p < 0.0001$) pour les visages présentés à l'endroit ($r = 0.86, p < 0.0001$; Figure 7). Cette corrélation entre la distance des yeux sur l'axe (Y) vertical, par rapport au centre de l'image et la distance entre le centre de l'image et les l'amplitude (Y) des saccades correctes indique que l'amplitude en (Y) des saccades augmente avec la distance des yeux par rapport au centre de l'image. A titre exploratoire nous avons par ailleurs calculés ces mêmes corrélations en divisant l'analyse en fonction du côté de présentation des images de visages (Voir Annexe III).

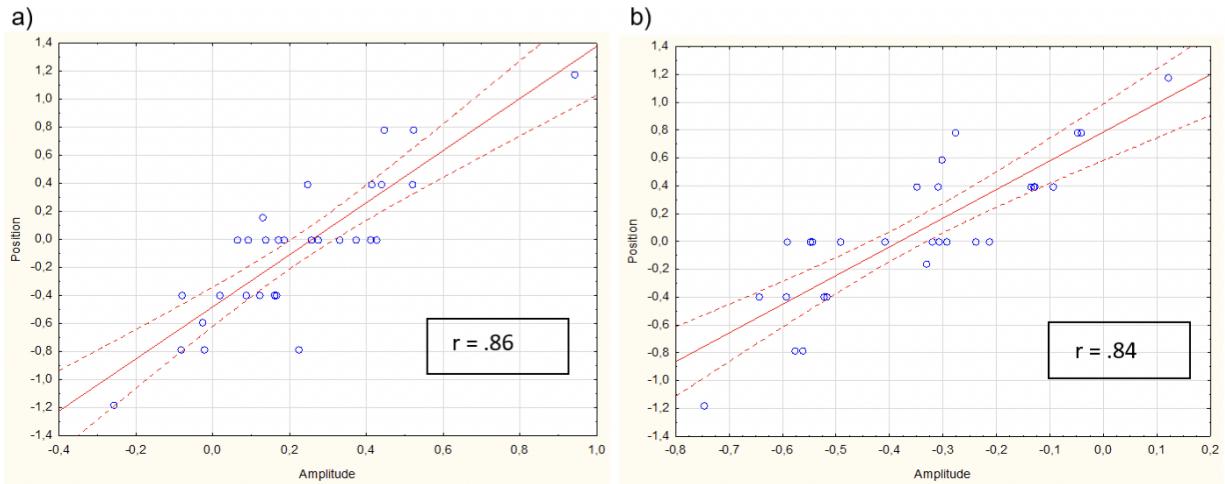


Figure 7. Corrélation entre la distance relative des yeux dans les visages et l'amplitude Y des saccades correctes (**a**) pour la condition Type de présentation à l'endroit et (**b**) pour la condition Type de présentation à l'envers. Nous avons par ailleurs calculé pour chaque participant le coefficient de corrélation pour les visages présentés à l'endroit et à l'envers qui révèle une consistance de ces corrélations entre les participants (Annexe I).

Discussion de l'Expérience 1

Cette première étude visait à tester si la détection ultra-rapide voire automatique des visages lors d'une tâche de choix saccadique observée dans de nombreux travaux (Crouzet et al., 2010 ; Crouzet & Thorpe, 2011 ; Boucart et al. 2016 ; Guyader, Chauvin, Boucart & Peyrin, 2017 ; Kauffmann et al., 2019) reposait sur un traitement holistique des visages, connu pour être perturbé par l'inversion de ceux-ci. Sous cette hypothèse, nous nous attendions à ce que l'inversion des visages lors d'une tâche de choix saccadique ait pour effet de réduire considérablement ce biais pour les visages. En accord avec les précédentes études nous avons tout d'abord observé que lorsque les stimuli étaient présentés à l'endroit, les participants étaient plus rapides et faisaient moins d'erreur pour initier leurs saccades vers la cible lorsque celle-ci était un visage que lorsque c'était un véhicule. Nos résultats ont également montré que l'inversion des stimuli avait pour effet d'impacter les performances, et notamment de réduire l'effet de la Catégorie de la cible (i.e. biais vers les visages) sur les mER. Cependant, nous n'avons pas observé d'interaction entre la Catégorie de la cible et le Type de présentation sur les SRT, suggérant que l'inversion des stimuli avait pour effet d'augmenter globalement les temps de réaction, indépendamment de la catégorie de la cible. De plus, l'analyse des SRT minimums suggère que l'inversion des stimuli n'impacte pas les temps de réactions saccadiques minimums vers les visages. Comme dans l'étude récente de Kauffmann et al. (2019), nous nous sommes également intéressés à l'amplitude horizontale des saccades, pouvant plus spécifiquement refléter des aspects de programmation. Nous avons pu répliquer les résultats de cette étude en montrant que lorsque les stimuli étaient à l'endroit, l'amplitude des saccades était plus petite lorsque la cible était un véhicule que lorsque c'était un visage. En accord avec l'interprétation proposée par Kauffmann et al. (2019), ce résultat pourrait refléter un poids plus important des visages lors d'une programmation en parallèle de deux saccades concurrentes. De façon intéressante, cette différence d'amplitude était réduite lorsque les stimuli étaient inversés, suggérant que le poids des visages était moindre lors de la mise en concurrence. Dans l'ensemble, les résultats de cette tâche de choix saccadique, supposent qu'un traitement holistique peut expliquer en partie le biais pour les visages

puisque l'inversion de ceux-ci réduit cet effet visible sur la mER et la moyenne des amplitudes (X) vers les véhicules.

Cependant ce biais vers les visages ne peut pas être entièrement expliqué par un traitement holistique puisque les saccades vers les visages demeurent bien plus rapides que vers les véhicules. Il est donc envisageable que d'autres facteurs expliquent ces saccades très courtes, tel que la détection d'éléments saillants qui pourraient subsister et attirer l'attention même quand les visages sont inversés. En effet, l'analyse de l'amplitude verticales (Y) a révélé que les saccades étaient dirigées de façon différente en fonction de l'orientation des stimuli et de la Catégorie de la cible. En particulier, lorsque la cible était un visage, nous avons observé une forte corrélation entre la distance relative des yeux dans les visages et ces amplitudes Y. Tout cela suggère que les saccades sont majoritairement programmées vers les yeux dans les visages, qu'ils soient à l'endroit ou à l'envers. La programmation des saccades résulterait peut-être d'une stratégie basée sur la saillance d'éléments du visage et plus particulièrement des yeux, plutôt que de se baser uniquement sur un traitement holistique (i.e. traitement du visage dans sa globalité).

Pour aller plus loin nous avons dans une seconde expérience voulu perturber encore plus fortement un traitement holistique des visages. Pour cela nous avons utilisé une technique modifiant la configuration spatiale des visages tout en préservant les éléments les plus importants tels que les yeux. Perturber la configuration spatiale des visages réduit considérablement leur rapidité de catégorisation et provoque plus d'erreurs de reconnaissance (Donnelly et al., 1994 ; Tanaka & Farah, 1993 ; Cassia, Kuefner, Westerlund & Nelson, 2006). Sous l'hypothèse que les participants ne puissent plus se baser sur un traitement holistique qui contient une configuration des éléments du visage, ce sont notamment les éléments isolés du visage qui devraient être utilisés pour réaliser la tâche de choix saccadique. Plus précisément, nous voulions tout d'abord observer à quel point l'altération de la configuration spatiale des images pouvait perturber les performances en termes d'erreur et de TRS. Nous voulions également quantifier à quel point la perturbation de la configuration spatiale des visages influencerait l'amplitude des saccades vers la cible, ainsi observer si le poids conféré aux visages lors de la programmation des saccades, par rapport aux véhicules, était impacté.

Expérience 2

Outils et méthode

Participants. Vingt-trois participants n'ayant pas participé à la première étude (20 femmes et 3 hommes ; 20.74 ± 2.75 ans, dont 2 gauchers) avec une vision normale ou corrigée de l'Université Grenoble Alpes ont participé à l'expérience pour des points d'examen. Tous les participants ont signé un consentement éclairé pour participer à l'expérience, qui a été conduite en accord avec le code éthique de l'association médicale mondiale pour les études impliquant l'humain (déclaration d'Helsinki).

Stimuli. Les stimuli utilisés ont été élaborés à partir des mêmes 60 images en couleur téléchargées de la base d'image Pixabay et sous licence libre Creative Commons (<https://pixabay.com/fr/>). Pour cette deuxième expérience, nous avons réalisé des scrambles à l'aide de Matlab R2018a (MathWorks, Inc., MA, USA). Chaque image était divisée en imagettes qui étaient ensuite repositionnées aléatoirement dans l'image, de façon à modifier complètement leur structure spatiale. Les scrambles d'images ont été construits selon trois conditions variant en terme de nombre d'imagettes (1×2 , 3×3 et 4×4 voir figure 8), permettant de plus ou moins préserver la structure spatiale des images tout en préservant leur contenu

sémantique. Comme dans l'expérience 1, la moitié des stimuli contenaient étaient construits à partir d'images de visages et l'autre moitié de véhicules (e.g. voiture, train, scooter). Toutes les images avaient une taille de 300 x 300 pixel soit 11.5° x 11.5° d'angle visuel pour une présentation à une distance de 60 cm.

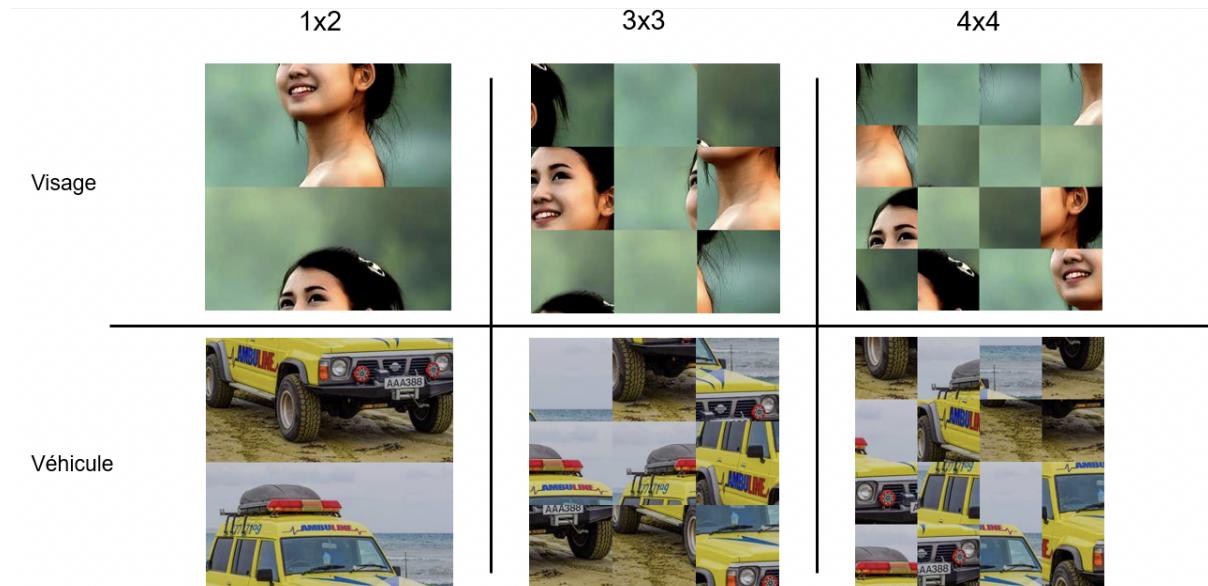


Figure 8. Exemples de scrambles, pour la Catégorie de la cible (e.i. visage et véhicule) selon le Type de scramble (e.i. 1×2, 3×3 et 4×4).

Design. Le paradigme expérimental était le même que dans l'Expérience 1. Pour chaque session expérimentale les participants ont tous passé les mêmes conditions (i.e. stimuli scramble 1×2, 3×3 et 4×4). Les différents types de scrambles étaient présentés par blocs de 10 essais en commençant par un bloc de la condition 1×2 puis des blocs des conditions 3×3 et 4×4 dans un ordre aléatoire. Les participants avaient pour consigne de diriger leur regard le plus rapidement possible vers le véhicule dans une session et vers le visage dans une autre session, l'ordre des sessions étant contrebalancée entre les participants. Chaque image de chaque condition (i.e. 1×2, 3×3 et 4×4) était présentée une fois à gauche et une fois à droite de la croix de fixation. Chaque session contenait 180 essais (i.e. 30 paires d'images x 3 conditions scramble x 2 côtés de présentation) et durait environ 20 minutes. Afin de se familiariser avec la tâche, tous les participants réalisaient avant l'expérience un court entraînement (12 essais) sur des images non présentées dans l'expérience.

Analyse des données. L'exactitude des réponses (taux d'erreur moyen - mER en %), les temps de réaction saccadique (TRS en ms) et les amplitudes verticale (X) et horizontale (Y) (en degrés d'angle visuel) des premières saccades correctes ont été analysés. Afin d'éviter d'inclure dans les analyses les saccades aux profils déviants en termes de latence, d'amplitude, de durée et susceptibles de ne pas refléter la catégorisation effective des stimuli, seules les saccades satisfaisant certains critères ont été incluses dans les analyses. Les saccades étaient considérées comme valides si : 1 - elles étaient initiées au moins 50 ms après la présentation des images ; 2 - elles avaient une durée inférieure à 100 ms ; 3 - leur position de départ se situait dans un carré de 2° d'angle visuel autour de la croix de fixation centrale ; 4 - leur amplitude totale était supérieure à 1° d'angle visuel. Au total 11.95 % des saccades ont été retirées des analyses. Les analyses ont été réalisées sur STATISTICA 10.0 ® (Dell Inc.,

Round Rock, USA). Nous avons effectué des analyses de variance (ANOVA) pour mesures répétées de type 2x2 avec la Catégorie de la cible (visage ou véhicule) et le Type de présentation (à l'endroit ou à l'envers) comme facteurs intra-sujets sur la mER, les TRS et l'amplitude des saccades en X et Y. Les tailles d'effet ont été calculées en utilisant un eta carré partiel (η_p^2). Le seuil de significativité des tests a été fixé à $\alpha = 0.05$ et les p-values reportées ont été corrigées par une correction de Bonferroni pour les comparaisons multiples.

Résultats

Exactitude. L'ANOVA réalisée indique un effet principal de la Catégorie de la cible ($F_{1, 22} = 53.82, p < 0.0001, \eta_p^2 = 0,710$). Les participants faisaient en moyenne plus d'erreurs pour réaliser la tâche lorsque la catégorie de la cible était un véhicule (i.e. plus d'erreurs de saccades vers le distracteur visage ; $29 \pm 12.4\%$) plutôt que lorsque la cible était un visage (i.e. le distracteur était un véhicule ; $14,51 \pm 8,85\%$). On observe également un effet principal du Type de scramble ($F_{1, 22} = 4.12, p < 0.05, \eta_p^2 = 0.158$). Ainsi, les participants faisaient d'autant plus d'erreurs que la structuration spatiale était altérée dans les scrambles (i.e., augmentation du nombre d'imagettes; $4\times 4 : 23.63 \pm 14.33\%, 3\times 3 : 21.73 \pm 13.23\% ; 1\times 2 : 19.91 \pm 11.80\%$). L'ANOVA indique de plus la présence d'une interaction significative entre les deux facteurs ($F_{1, 22} = 3.44, p < 0.05, \eta_p^2 = 0.135$), (Figure 9a). Les comparaisons planifiées ont révélé un effet du Type de scramble quel que soit la Catégorie de la cible. Plus précisément, les participants faisaient plus d'erreurs lorsque les images étaient présentées en scramble 4×4 (cible visage $15,17 \pm 9.36\%$; cible véhicule $32.09 \pm 13.52\%$; $F_{1, 22} = 20.64, p < 0.0001$) que lorsque les images étaient présentées en scramble 3×3 (cible visage : $13.89 \pm 10.25\%$; cible véhicule : $29.57 \pm 11.16\%$; $F_{1, 22} = 48.34, p < 0.0001$) ainsi qu'en scramble 1×2 (cible visage : $14.46 \pm 10.31\%$; cible véhicule : $25.35 \pm 10.79\%$; $F_{1, 22} = 42.32, p < 0.0001$). On observe par ailleurs, uniquement un effet linéaire significatif du Type de cible pour la catégorie cible véhicule ($F_{1, 22} = 14.73, p < 0.001$), mais non significatif pour les visages ($F_{1, 22} < 1, p = 0.01$). Plus précisément les participants faisaient d'autant plus d'erreurs que la structuration spatiale des véhicules était altérée dans les scrambles ($4\times 4 : 32.09 \pm 13.52\% ; 3\times 3 : 29.57 \pm 11.16\% ; 1\times 2 : 25.35 \pm 10.79\%$). Cette interaction suggère dans un premier temps que plus les images sont détériorées spatialement, plus les participants font d'erreurs. Mais aussi que l'effet attracteur des visages reste constant au travers les conditions et que l'altération spatiales des stimuli impacte surtout les saccades vers les véhicules.

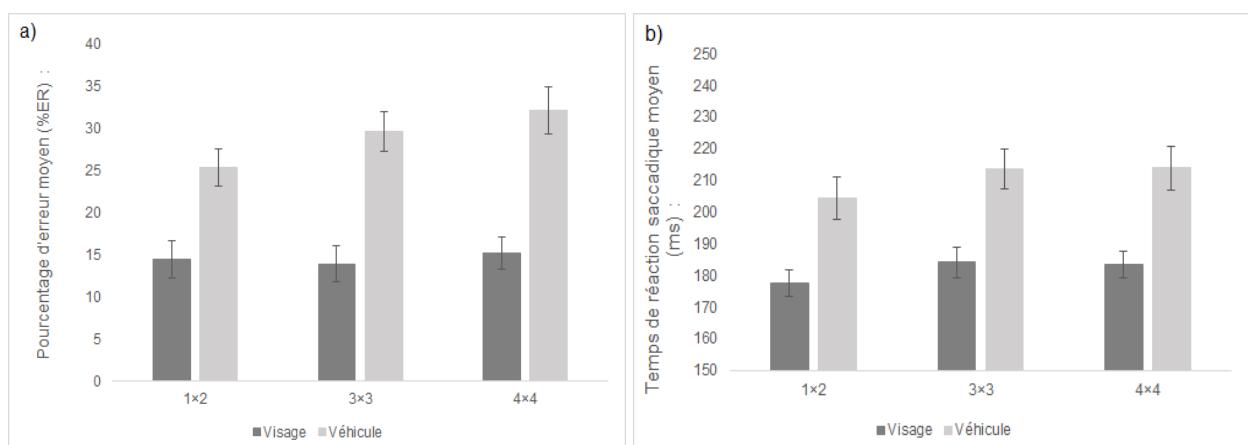


Figure 9. (a) Pourcentage d'erreur (%) moyen et (b) Temps de réaction saccadique moyen (ms) en

fonction de la condition Type de scramble (1×2 , 3×3 vs. 4×4) et de la Catégorie de la cible (visage vs. véhicule). Les barres d'erreurs représentent l'erreur standard.

Temps de réaction saccadique moyen et temps de réaction saccadique minimum.

L'analyse des TRS indique tout d'abord un effet principal de la Catégorie de la cible ($F_{1, 22} = 33.93, p < 0.0001, \eta_p^2 = 0.607$). Les participants étaient plus rapides pour initier une saccade correcte vers les visages (180 ± 21 ms) que vers les véhicules (209 ± 31 ms). L'analyse révèle aussi un effet principal du Type de scramble ($F_{1, 22} = 14.134, p < 0.0001, \eta_p^2 = 0.391$). Les participants étaient plus rapides pour initier une saccade lorsque les images étaient présentées en scramble 1×2 (189 ± 30 ms) et en scramble 3×3 ($197,42 \pm 30$ ms) plutôt que présentés en scramble 4×4 (197.03 ± 31 ms). Cependant, nous n'avons pas observé d'interaction significative entre ces deux facteurs ($F_{1, 22} < 1, p = 0.61$), (Figure 9b). De la même manière que pour l'expérience 1, nous avons voulu explorer si le temps minimum pour initier des saccades vers les différentes images cibles différait en fonction de la condition scramble (Figure 9). Lorsque la cible était un visage divisé en scramble 1×2 , le TRS minimum sur l'ensemble des participants était de 110 ms ($\chi^2 = 4.41, p < 0.01$), lorsque la cible était un visage divisé en scramble 3×3 , le TRS minimum était de 120 ms ($\chi^2 = 5.30, p < 0.05$) et enfin lorsque la cible était un visage divisé en scramble 4×4 , le TRS minimum était de 110 ms ($\chi^2 = 10.90, p < 0.01$). Pour les véhicules, le TRS minimum lorsque la cible était divisée en scramble 1×2 était de 140 ms ($\chi^2 = 8.41, p < 0.01$), était de 160 ms lorsque la cible était divisée en scramble 3×3 était de 160 ms ($\chi^2 = 8.41, p < 0.01$) et était de 190 ms lorsque la cible était divisée en scramble 4×4 était de 190 ms ($\chi^2 = 37.21, p < 0.01$). Le fait de modifier la configuration spatiale des véhicules perturbe le temps de réaction minimum pour initier des saccades vers les véhicules. Cependant la modification de la configuration spatiale a moins d'impact pour initier des saccades vers les visages sur les temps de réaction saccadiques minimums.

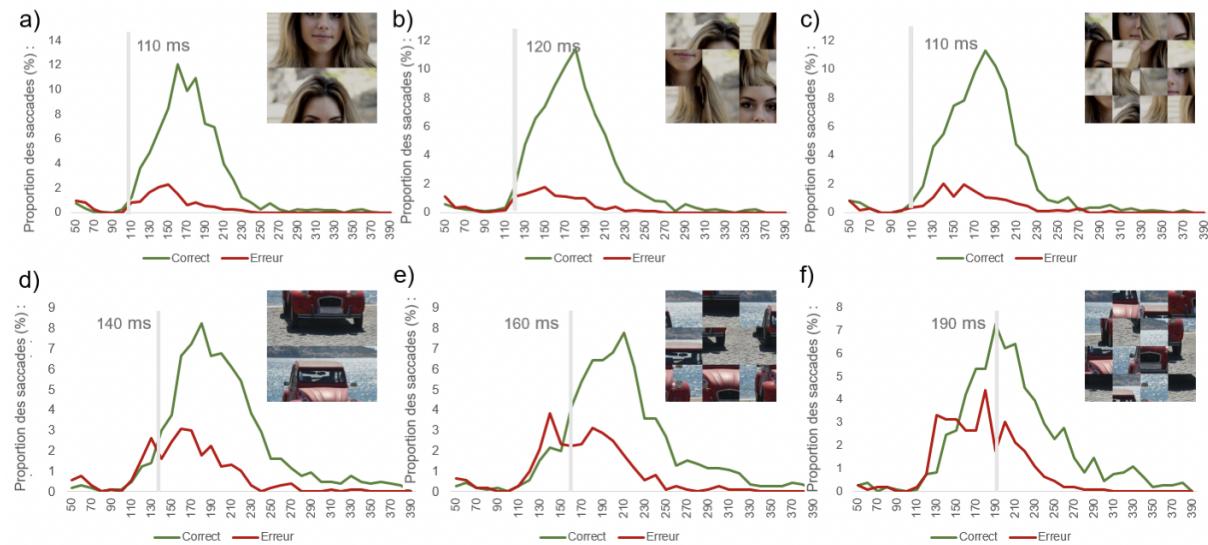


Figure 10. Distribution de la proportion de saccades correctes (en vert) et erreur (en rouge) pour chaque bin de temps de réaction saccadique (a) Pour le type de scramble 1×2 et la catégorie de la cible visage (b) pour le type de scramble 3×3 et la catégorie cible visage (c) pour le type de scramble 4×4 et la catégorie cible visage (d) pour le type de scramble 1×2 et la catégorie cible véhicule (e) pour le type de scramble 3×3 et la catégorie cible véhicule et (f) pour le type de scramble 4×4 et la catégorie cible véhicule. Les barres verticales grises représentent le temps de réaction saccadique minimum.

Amplitude axe horizontal (X) et amplitude axe vertical (Y).

L'ANOVA réalisée sur les amplitudes horizontales (X) moyennes des premières saccades correctes indique tout d'abord un effet principal de la Catégorie de la cible ($F_{1, 22} = 16,28, p < 0,001, \eta_p^2 = 0,425$). Les participants réalisaient des saccades plus petites lorsque la cible un véhicule ($5,98 \pm 0,93^\circ$) que lorsque la cible était un visage ($7,58 \pm 0,57^\circ$). L'ANOVA indique aussi un effet principal du Type de scramble ($F_{1, 22} = 166, p < 0.0001, \eta_p^2 = 0.883$), les participants réalisaient en moyenne des saccades plus petites lorsque les images étaient divisées en 3×3 ($6,33 \pm 1,13^\circ$) plutôt que lorsqu'elles étaient divisées en 1×2 ($7,11 \pm 0,82^\circ$) et encore plus petites quand les images étaient divisées en 4×4 ($5,64 \pm 0,74^\circ$). Un effet d'interaction entre la catégorie de la cible et le type de scramble est aussi présent ($F_{1, 22} = 26,14, p < 0.0001, \eta_p^2 = 0,543$) (Figure 11a). Les comparaisons planifiées ont révélé un effet du Type de scramble quel que soit la Catégorie de la cible. Plus précisément, l'effet significatif était observé pour les scrambles présentés en 1×2 ($F_{1, 22} = 62,73, p < 0.0001$) mais pas pour les scrambles 3×3 ($F_{1, 22} = 4,64, p = 0,089$), ainsi que les scrambles 4×4 ($F_{1, 22} < 1, p = 0,66$). Lorsque les scrambles étaient divisés en 1×2 les participants réalisaient des saccades plus petites vers les véhicules ($6,71 \pm 0,75^\circ$) que vers les visages ($7,52 \pm 0,71^\circ$). Les analyses ont par ailleurs révélé un effet quadratique significatif de la Catégorie de la cible pour la catégorie cible visage quel que soit le Type de scramble ($F_{1, 22} = 119,62, p < 0.0001$). Plus précisément les participants faisaient d'autant plus de saccades horizontales courtes que la structuration spatiale des visages était altérée dans les scrambles ($4 \times 4 : 5,69 \pm 0,77^\circ ; 3 \times 3 : 5,88 \pm 0,86^\circ ; 1 \times 2 : 7,52 \pm 0,71^\circ$). De la même manière pour les véhicules, les participants faisaient d'autant plus de saccades horizontales courtes que la structuration spatiale des véhicules était altérée dans les scrambles ($4 \times 4 : 5,61 \pm 0,73^\circ ; 3 \times 3 : 5,57 \pm 0,82^\circ ; 1 \times 2 : 6,71 \pm 0,75^\circ$). Cette interaction suggère que plus les cibles ont une configuration spatiale perturbée, plus les participants ont tendance à réaliser des saccades de courte amplitude sur l'axe horizontal (X).

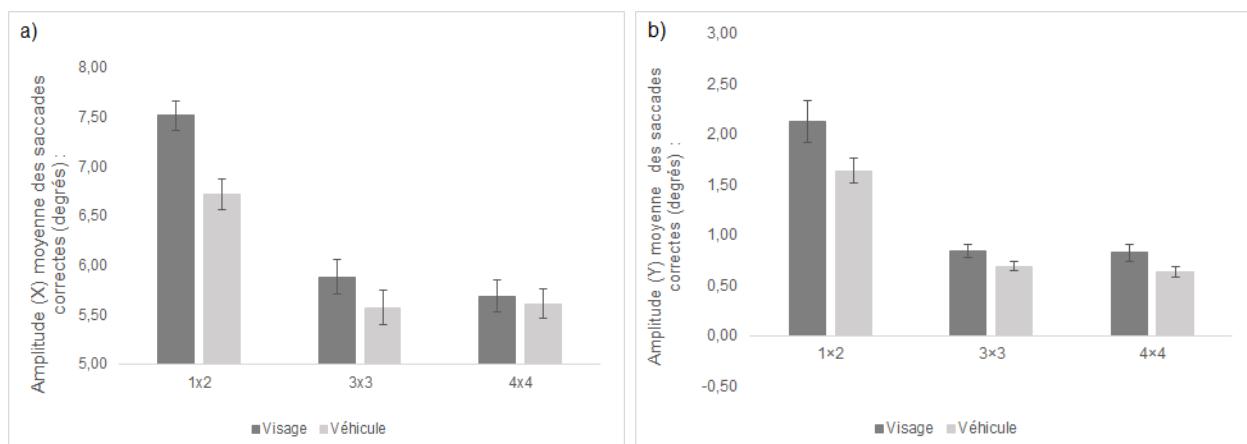


Figure 11. (a) Amplitude moyenne (X) (degrés)(b) Amplitude moyenne (Y) (degrés) en fonction de la condition Type de scramble ($1 \times 2, 3 \times 3$ vs. 4×4) et de la Catégorie de la cible (visage vs. véhicule). Les barres d'erreurs représentent l'erreur standard.

L'ANOVA réalisée sur les amplitudes verticales (Y) indique un effet principal de la Catégorie de la cible ($F_{1, 22} = 11,88, p < 0,01, \eta_p^2 = 0,350$). Plus précisément, en moyenne les participants réalisaient une saccade plus petite lorsque la cible était un véhicule ($0,99 \pm$

0.60°), que lorsque la cible était un visage ($1.25 \pm 0.87^\circ$). L'ANOVA révèle aussi un effet principal du Type de scramble ($F_{1, 22} = 82.81, p < 0.0001, \eta_p^2 = 0.790$), les participants réalisaient en moyenne des saccades plus courtes lorsque les images étaient divisées en 3×3 ($0.76 \pm 0.28^\circ$) plutôt que lorsqu'elles étaient divisées en 1×2 ($1.87 \pm 0.84^\circ$) et encore plus courtes quand les images étaient divisées en 4×4 ($0.73 \pm 0.34^\circ$). Un effet d'interaction entre la Catégorie de la cible et le Type de scramble était aussi présent ($F_{1, 22} = 6.50, p < 0.01, \eta_p^2 = 0.228$), (Figure 11b). Les comparaisons planifiées ont révélé un effet de la Catégorie de la cible, quel que soit le Type de scramble. Plus précisément, les participants faisaient des saccades verticales plus courtes vers les véhicules que vers les visages lorsque les images étaient divisées en scramble 4×4 (cible visage $0.82 \pm 0.39^\circ$; cible véhicule $0.63 \pm 0.25^\circ$; $F_{1, 22} = 9.81, p < 0.05$) que lorsque les images étaient divisées en scramble 1×2 (cible visage : $2.12 \pm 1^\circ$; cible véhicule : $1.64 \pm 0.57^\circ$; $F_{1, 22} = 10.79, p < 0.05$) et encore plus courte vers les scramble divisés en 3×3 (cible visage : $2.12 \pm 0.32^\circ$; cible véhicule : $1.64 \pm 0.22^\circ$; $F_{1, 22} = 16.15, p < 0.05$). De plus on observe un effet quadratique significatif du Type de scramble pour la Catégorie de la cible visage ($F_{1, 23} = 50.32, p < 0.0001$) et pour la Catégorie de la cible véhicule ($F_{1, 23} = 72.22, p < 0.0001$). Les participants faisaient d'autant plus de saccades verticales courtes que la structuration spatiale des visages était altérée dans les scrambles (4×4 : $0.82 \pm 0.39^\circ$; 3×3 : $0.85 \pm 0.32^\circ$; 1×2 : $2.12 \pm 1^\circ$). De la même manière pour les véhicules, les participants faisaient d'autant plus de saccades verticales courtes que la structuration spatiale des visages était altérée dans les scrambles (4×4 : $0.63 \pm 0.24^\circ$; 3×3 : $0.69 \pm 0.21^\circ$; 1×2 : $1.64 \pm 0.57^\circ$). De même pour les amplitudes horizontales, cette interaction suggère que plus les cibles ont une configuration spatiale perturbée, plus les participants ont tendance à réaliser des saccades plus courtes en termes d'amplitude sur l'axe vertical (Y).

Position finale de la première saccade correcte.

Nous avons dans cette deuxième expérience voulu explorer si le même élément que lors de l'expérience précédente (à savoir les yeux, comme suggéré par les résultats de l'expérience 1) était susceptible d'être fixé dans les images de visages par les participants. Nous avons de plus voulu observer si le même élément était fixé lorsque la configuration des visages était perturbée avec les différents Types de scramble. Nous avons ainsi retranscrit la position finale de la première saccade sur les différents stimuli, divisés en scramble 1×2 , 3×3 et 4×4 (Figure 12). Sur la base de ces figures, on observe que les scrambles divisés en 1×2 entraîne des saccades vers la partie haute de l'image (e.i. yeux, bouche et nez). Alors que les impacts des saccades sur les scrambles de 3×3 et 4×4 n'aboutissent pas sur le même élément du visage en fonction de leur côté de présentation et selon la position des éléments du visage dans le scramble.



Figure 12. Exemple de positions finales des saccades correctes vers la catégorie cible visage en fonction de la condition Type de présentation. Les points gris représentent le point d'impact de chaque participant

Discussion seconde expérience

A titre indicatif et pour une meilleure clarté de l'ensemble des résultats, un tableau récapitulatif résume les résultats des expériences 1 et 2 (Annexe IV). Dans la continuité de la première expérience, cette deuxième expérience avait pour projet d'observer si la perturbation du traitement holistique et plus particulièrement la perturbation de la configuration spatiale a pour effet de réduire ce biais des visages. Nous avions aussi émis l'hypothèse qu'en empêchant le traitement de la configuration des visages, les éléments du visage seraient privilégiés pour la tâche de catégorisation. En cohérence avec les précédentes études les participants étaient plus rapides et faisaient moins d'erreur pour initier leurs saccades vers la cible lorsque celle-ci était un visage que lorsque la cible était un véhicule. De façon étonnante, mélanger la configuration typique d'un visage n'entravait aucunement leur reconnaissance. Cependant la perturbation des éléments constituant les véhicules impacte considérablement les saccades vers ceux-ci, ainsi plus leur configuration spatiale était dégradée plus les saccades étaient dirigées vers le distracteur. De façon similaire à l'expérience précédente nous n'avons pas observé d'interaction entre l'identité de la cible et le type de scramble sur les SRT, suggérant que la dégradation de la configuration spatiale des images avait pour effet d'augmenter globalement les temps de réaction, indépendamment de la catégorie de la cible. De même, les SRT minimums révèlent que seulement les saccades vers les véhicules sont impactées par la modification de la configuration spatiale.

Similairement à l'expérience 1 et en accord avec l'interprétation de Kauffmann et al., (2019), les amplitudes horizontale (X) étaient toujours plus petites vers les véhicules lorsque que la configuration des visages était préservée dans les scrambles 1x2. Toutefois les saccades réalisées vers des cibles qui ne possédaient plus aucune configuration spatiale étaient considérablement réduites en termes d'amplitude (i.e., scrambles 3x3 et 4x4). En outre, à la différence des images possédant une configuration spatiale intacte (i.e., scramble 1x2), l'étendue de amplitudes horizontales vers les scrambles dégradés spatialement étaient

semblables quel que soit la cible. L'analyse de l'amplitude verticale (Y) suggère que lorsque la configuration des visages était préservée (scrambles 1×2), les participants effectuaient des saccades de plus grande amplitude et dirigées vers le haut de l'image. Cependant de la même manière que pour les amplitudes horizontales, lorsque la configuration était altérée on observe les saccades de plus petites amplitudes. En effet, même si les visages étaient toujours aussi bien reconnus dans cette tâche de paradigme saccadique, on observe un impact sur la programmation des saccades. On pourrait à première vue penser que cette réduction de l'amplitude horizontale serait expliquée par un poids particulièrement réduit des visages par rapport à celui des véhicules de par la dégradation de sa configuration. En revanche une explication alternative est possible, compte tenu du manque de contrôle de la position des éléments des visages dans les scrambles 3×3 et 4×4. Il serait possible que lorsque l'information configurale des visages n'était plus présente, les participants adoptaient une stratégie économique consistant à déplacer leur attention et programmer leurs saccades vers les éléments du visage les plus proches du point de fixation. Ainsi la réduction de l'amplitude des saccades horizontale et verticales reflèterait cette stratégie économique. Il s'agit ici d'une des limites de notre étude, qui ne nous permettent pas de quantifier cet effet et de conclure sur l'amplitude des saccades vers les visages divisés en scramble. Cependant la proportion des points d'impacts corrobore l'hypothèse qu'en absence d'une configuration spatiale cohérente, les participants programmaient des saccades vers les différents éléments du visage. Plus intéressant encore, l'initiation de ces saccades vers les éléments d'un visage ne semblait pas avoir de conséquence sur les SRT minimum.

Dans l'ensemble, cette deuxième expérience suggère que même sans cohérence spatiale entre les parties des visages, la saillance et la pertinence des éléments d'un visage sembleraient être les indices le plus porteur de poids pour programmer et initier les saccades vers la cible.

Discussion générale

Dans cette étude, nous nous sommes intéressés au rôle du traitement holistique pour la reconnaissance ultra-rapide de visages, à travers des mesures de saccades oculaires. Pour cela, nous avons limité l'accès au contenu holistique des visages en les présentant à l'endroit ou à l'envers lors d'une tâche de choix saccadique. Nous voulions examiner dans quelle mesure un traitement holistique des visages explique le biais pour ces stimuli observé dans plusieurs études (Crouzet et al., 2010 ; Crouzet et Thorpe, 2011 ; Boucart et al. 2016 ; Guyader, Chauvin, Boucart et Peyrin, 2017 ; Kauffmann et al., 2019). Dans l'ensemble, nos résultats permettent de conforter ces précédents travaux, en mettant en évidence un biais vers les visages qui se traduit par des saccades plus rapides vers ceux-ci, par plus de saccades erreurs lorsque les visages sont des distracteurs, mais aussi une différence en terme d'amplitude des saccades en fonction de leur cible, les saccades vers les véhicules étant plus courtes que les saccades vers les visages.

Notre étude a également permis de préciser ces études de choix saccadique. Dans une première expérience, nous avons observé que l'inversion des stimuli entraînait un taux d'erreur plus important lorsque la cible était un visage relativement à la condition où les visages étaient à l'endroit. De façon importante, l'inversion des stimuli avait également pour effet de réduire la différence d'amplitude horizontale entre les deux catégories de stimuli. Nous avons par ailleurs relevé des saccades dirigées dans la région des yeux pour les visages à l'endroit et à l'envers. Dans la seconde expérience nous avons observés un biais vers les visages même sans configuration spatiale typique.

Biais vers les visages en parti expliqués par un traitement holistique ?

L'avantage d'un traitement holistique maximiseraient la sensibilité à l'information configurale (Farah et al., 1993), ce qui permettrait une stratégie attentionnelle automatique vers les visages (Richler, Mack, et al., 2009) où l'attention serait déployée sur tout le visage au même moment (Richler, Wong et al., 2011). Ce traitement d'un visage dans sa globalité serait produit au cours des processus perceptifs précoce (Goffaux & Rossion, 2006). Ainsi, le traitement de la configuration d'un visage amènerait à diriger les saccades de façon automatique vers celui-ci. Par conséquent, lorsqu'un visage est présenté en tant que distracteur, il serait difficile de désengager l'attention vers le visage lors de la programmation d'une saccade vers un véhicule. Nos résultats indiquent que l'inversion des visages provoque plus d'erreurs (mais n'indiquent aucune augmentation des temps moyens et minimums lors de l'initiation des saccades). Nos résultats sont soutenus par ceux observés par Gilchrist et Proske (2006). Ces auteurs ont utilisé un paradigme d'anti-saccades (i.e. effectuer une saccade à l'opposé de l'apparition d'un distracteur) permettant d'évaluer à la fois la programmation d'une saccade volontaire mais aussi l'inhibition d'une saccade involontaire vers le distracteur. Cette étude a non seulement permis de révéler que les visages à l'endroit généraient plus de saccades involontaires vers eux que les visages à l'envers, mais aussi qu'en moyenne les anti-saccades correctes étaient initiées dans les mêmes laps de temps. Il serait ainsi possible de considérer que les visages à l'endroit et à l'envers soient traités de la même manière lorsqu'ils sont mis en concurrence avec les véhicules dans un paradigme de choix saccadique. Ainsi même si l'inversion des visages avait pour effet de réduire en partie le biais vers les visages, celui-ci persistait pour la programmation des saccades vers les véhicules. Cela suggère que d'autres éléments des visages, indépendamment de leur configuration pourraient également expliquer ces résultats.

Rôle de la saillance et de la pertinence des éléments du visage dans le biais vers les visages ?

La perception rapide d'un visage provoquerait ainsi une saccade vers l'élément le plus saillant ou pertinent ? Pour réaliser une saccade correcte lors d'une tâche de choix saccadique, le système saccadique doit attribuer le plus de poids à la saccade programmée vers la cible qu'à la saccade vers le distracteur. L'attribution du poids lors de la programmation des saccades se baserait sur une carte de priorité, qui serait conçue à partir des éléments visuels perçus dans les images. Ainsi le poids attribué à chaque saccade concurrente dépendrait de la saillance de chacun des éléments de l'image ciblée (i.e. information perceptive liée au signal), (Itti & Koch 2001 ; White, Kan, Levy, Itti, & Munoz, 2017) et de leur pertinence (pour une revue voir Fecteau, & Munoz, 2006). Dans nos deux expériences, nous avons observé que la majorité des saccades vers un visage aboutissait sur une fixation vers différents éléments du visage et plus particulièrement dans la région des yeux. Ces résultats sont partagés avec d'autres études (Peterson & Eckstein ,2012 ; Blais, Jack, Scheepers & Fiset, 2008 ; Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006 ; Janik, Wellens, Goldberg & Dell'Osso, 1978) qui montrent que la région des yeux est la région du visage majoritairement fixée. Cependant il est important de remarquer que les saccades dirigées vers la zone des yeux peuvent également être expliquées par un autre phénomène observé pour la programmation de saccades: Le biais du centre de gravité. Ce biais rendrait compte d'une première saccade programmée vers le centre de la cible (Findlay, 1982 ; Findlay & Gilchrist, 1998). Or le centre de gravité des visages serait la région des yeux (Bidemann, Scheepers, Burton, 2009).

Pour initier une saccade vers un visage qui a une configuration spatiale perturbée, la programmation de la première saccade pourrait refléter une stratégie pour optimiser les performances. Ainsi, le système saccadique se baserait sur les éléments les plus informatifs,

tels que les yeux en priorité pour différencier un visage d'un véhicule. Lorsque les yeux ne sont pas accessibles facilement et/ou que la saccade est trop coûteuse, le système saccadique pourrait programmer une saccade vers une autre élément du visage moins saillant que les yeux tel que la bouche. En accord avec cette interprétation, Lewis et Edmonds (2003) postulent l'utilisation de stratégies pour la détection des visages dans une scène. Ces auteurs ont pour cela comparé quels étaient les différences de temps de réactions selon l'élément du visage masqué. L'élément du visage dont le masquage affectait le plus la détection rapide d'un visage dans une scènes était les yeux. Ainsi les fixations dans la zone des yeux dans notre étude pourraient refléter la stratégie la plus rapide qui consiste à localiser les yeux quelle que soit l'orientation du visage.

Interaction entre le traitement holistique et la saillance des traits du visage impactent la programmation des saccades ?

Nous avons observé des résultats similaires à ceux de Kauffmann et al. (2019) pour les amplitudes des saccades en condition normale (i.e. stimuli présentés à l'endroit). Ces auteurs proposent que les mécanismes sous-jacents de la détection ultra rapide des visages impliqueraient le Colliculus Supérieur (CS) qui joue un rôle clé pour la programmation et l'exécution des saccades, mais qui serait également impliqué dans la détection très rapide de stimuli saillants importants pour le comportement tels que les visages. Les travaux de Nguyen et al. (2014) chez le macaque montrent par exemple que les neurones des couches superficielles du CS sont activés à partir de 50 ms lorsqu'ils sont exposés à des images de visages et sont activés à partir de 25 ms pour des patterns de formes géométriques agencées de la même manière qu'un visage à l'endroit, mais aussi à l'envers. Ainsi le système saccadique serait influencé par des informations visuelles très précoces et lié aux composantes de bas niveau des stimuli.

Le traitement holistique des visages pourrait apporter de nouveaux indices sur la perception-ultra rapide des visages. Outre la stratégie de programmer une saccade vers l'élément le plus saillant et pertinent, le traitement holistique pourrait renforcer l'attention vers les zones susceptibles de contenir une information pertinente pour la tâche. Rappelons que la théorie d'un traitement holistique comprend une représentation du visage dans sa globalité (i.e. les éléments d'un visage et leur relation) en mémoire. Ainsi la position relative des éléments d'un visage serait représentée implicitement en mémoire. On peut penser que cette représentation de la localisation des yeux pourrait avoir une influence lors de la programmation d'une saccade. Ainsi une information top-down (une information liée aux représentations, voir Richler, Cheung & Gauthier 2011a ; Richler et al. 2009) de cette représentation holistique pourrait ajouter un poids lors de la programmation en faveur de la zone où se trouve généralement les yeux, et ainsi participer à diriger le regard vers le distracteur. Ainsi les relations spatiales entre les éléments d'un visage pourraient être utilisés en addition à l'information conférée par les éléments eux-mêmes.

Il est possible d'émettre quelques limites sur notre méthodologie, notamment sur le choix de nos stimuli. Les visages font parti d'une catégorie d'objets très homogènes, tous les exemplaires de cette catégorie ont la même organisation spatiale et les mêmes éléments caractéristiques (i.e. yeux, nez et bouche) alors que véhicules non. Nous aurions pu utiliser des véhicules prototypiques tels que les voitures. De plus nous avons fait le choix de choisir des images écologiques, mais nous aurions pu aussi pré-traiter les images pour contrôler la position des éléments des visages dans nos différentes conditions de présentation. Il semble aussi important de préciser que le terme "traitement holistique" regroupe beaucoup de théories qui ne sont pas toujours en accords les unes avec les autres.

Cette recherche est le premier lien entre les études explorant le traitement holistique pour la reconnaissance des visages (i.e encodage d'un visage puis jugement de familiarité) et les études impliquant les processus de choix saccadique (i.e encodage simultané de deux images puis choix saccadique). Ainsi d'autres études sur la perception rapide des visages sont nécessaires pour évaluer les premières étapes précoce de cette théorie d'un traitement holistique des visages.

Conclusion

Cette étude avait pour but d'apporter des informations supplémentaires concernant la manière dont les visages sont perçus par rapport à d'autres autres objets. Elle avait aussi pour rôle d'ajouter des éléments en faveur d'un traitement holistique des visages contribuant à ce biais vers les visages. Ces résultats viennent confirmer les études précédentes, qui montrent qu'un visage provoque des réponses comportementales caractéristique par rapport à des non-visages. Ces résultats viennent aussi ajouter des indices à propos du caractère spécial d'un visage qui provoque des saccades vers ses caractéristiques élémentaires.

Références

- Allen, G., Peterson, M., & Rhodes, G. (2006). Seeking a Common Gestalt Approach to the Perception of Faces, Objects, and Scenes. *The American Journal Of Psychology*, 119(2), 311. doi: 10.2307/20445341
- Barton, J., Radcliffe, N., Cherkasova, M., Edelman, J., & Intriligator, J. (2006). Information Processing during Face Recognition: The Effects of Familiarity, Inversion, and Morphing on Scanning Fixations. *Perception*, 35(8), 1089-1105. doi: 10.1068/p5547
- Bentin, S., Allison, A., Puce, A., Perez, T., & McCarthy, G. (1996). Electro-physiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8, 551–565
- Bindemann, M., Scheepers, C., & Burton, A. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal Of Vision*, 9(2), 7-7. doi: 10.1167/9.2.7
- Boucart, M., Lenoble, Q., Quettelart, J., Szaffarczyk, S., Despretz, P., & Thorpe, S. (2016). Finding faces, animals, and vehicles in far peripheral vision. *Journal Of Vision*, 16(2), 10. doi: 10.1167/16.2.10
- Blais, C., Jack, R., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture Shapes How We Look at Faces. *Plos ONE*, 3(8), e3022. doi: 10.1371/journal.pone.0003022
- Brainard, D. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433-436. doi: 10.1163/156856897x00357
- Brants, M., Wagemans, J., & Op de Beeck, H. (2011). Activation of Fusiform Face Area by Greebles Is Related to Face Similarity but Not Expertise. *Journal Of Cognitive Neuroscience*, 23(12), 3949-3958. doi: 10.1162/jocn_a_00072
- Cassia, V., Kuefner, D., Westerlund, A., & Nelson, C. (2006). Modulation of Face-sensitive Event-related Potentials by Canonical and Distorted Human Faces: The Role of Vertical Symmetry and Up-Down Featural Arrangement. *Journal Of Cognitive Neuroscience*, 18(8), 1343-1358. doi: 10.1162/jocn.2006.18.8.1343
- Cheung, O., Richler, J., Palmeri, T., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal Of Experimental Psychology: Human Perception And Performance*, 34(6), 1327-1336. doi: 10.1037/a0011752
- Coutrot, A., & Guyader, N. (2014). How saliency, faces, and sound influence gaze in dynamic social scenes. *Journal Of Vision*, 14(8), 5-5. doi: 10.1167/14.8.5
- Crouzet, S. & Thorpe (2011). Low-level cues and ultra-fast face detection. *Frontiers In Psychology*, 2. doi: 10.3389/fpsyg.2011.00342
- Crouzet, S., Kirchner H., & Thope.J (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal Of Vision*, 10(4), 1-17. doi: 10.1167/10.4.16
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal Of Experimental Psychology: General*, 115(2), 107-117. doi: 10.1037//0096-3445.115.2.107

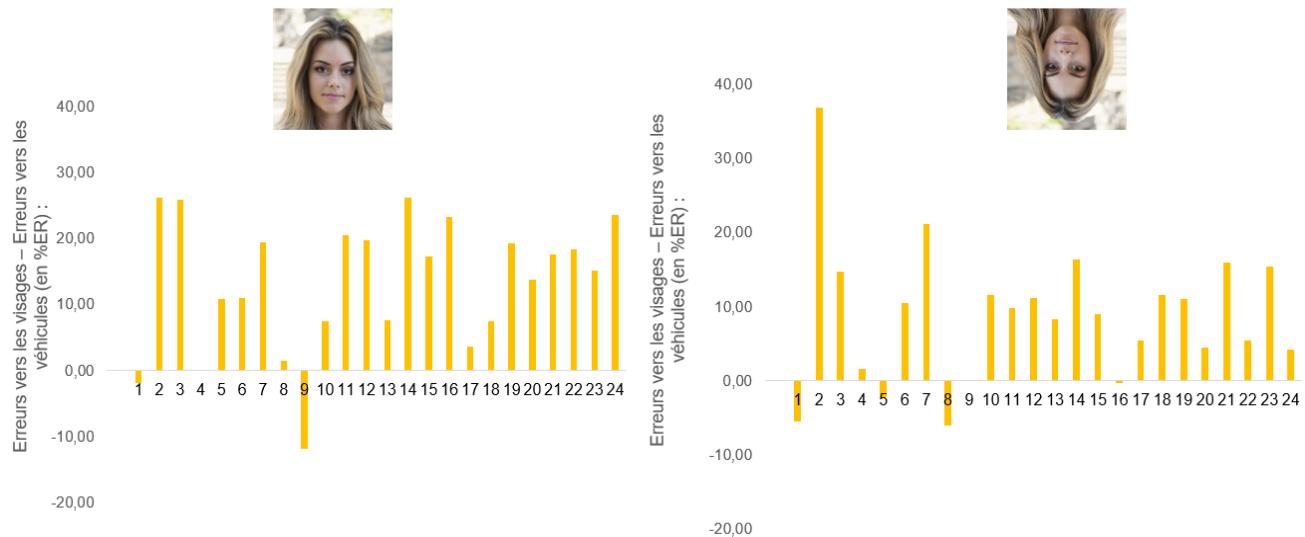
- Donnelly, N., Humphreys, G., & Sawyer, J. (1994). Stimulus factors affecting the categorisation of faces and scrambled faces. *Acta Psychologica*, 85(3), 219-234. doi: 10.1016/0001-6918(94)90036-1
- Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review Of Vision Science*, 1(1), 393-416. doi: 10.1146/annurev-vision-082114-035518
- Fecteau, J., & Munoz, D. (2006). Salience, relevance, and firing: a priority map for target selection. *Trends In Cognitive Sciences*, 10(8), 382-390. doi: 10.1016/j.tics.2006.06.011
- Findlay, J. (1982). Global visual processing for saccadic eye movements. *Vision Research*, 22(8), 1033-1045. doi: 10.1016/0042-6989(82)90040-2
- Findlay, J., & Gilchrist, I. (1997). Spatial Scale and Saccade Programming. *Perception*, 26(9), 1159-1167. doi: 10.1088/p261159
- Foulsham, T., Cheng, J., Tracy, J., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition*, 117(3), 319-331. doi: 10.1016/j.cognition.2010.09.003
- Galton, F. R. S. (1883). Recent Literature Inquiries into Human Faculty and its Development. New York: Macmillan & Co. 1883. *The Boston Medical And Surgical Journal*, 109(22), 520-520. doi: 10.1056/nejm188311291092206
- Gilaie-Dotan, S., Gelbard-Sagiv, H., & Malach, R. (2010). Perceptual shape sensitivity to upright and inverted faces is reflected in neuronal adaptation. *Neuroimage*, 50(2), 383-395. doi: 10.1016/j.neuroimage.2009.12.077
- Gilchrist, I., & Proske, H. (2006). Anti-saccades away from faces: evidence for an influence of high-level visual processes on saccade programming. *Experimental Brain Research*, 173(4), 708-712. doi: 10.1007/s00221-006-0455-1
- Goffaux, V. (2009). Spatial interactions in upright and inverted faces: Re-exploration of spatial scale influence. *Vision Research*, 49(7), 774-781. doi: 10.1016/j.visres.2009.02.009
- Goffaux, V., & Rossion, B. (2006). Faces are “spatial”–holistic face perception is supported by low spatial frequencies. *Journal of experimental psychology. Human perception and performance*, 32, 1023–1039. <http://dx.doi.org/10.1037/0096-1523.32.4.1023>.
- Goffaux, V., Gauthier, I., & Rossion, B. (2003). Spatial scale contribution to early visual differences between face and object processing. *Cognitive Brain Research*, 16(3), 416-424. doi: 10.1016/s0926-6410(03)00056-9
- Goffaux, V., Hault, B., Michel, C., Vuong, Q., & Rossion, B. (2005). The Respective Role of Low and High Spatial Frequencies in Supporting Configural and Featural Processing of Faces. *Perception*, 34(1), 77-86. doi: 10.1088/p5370

- Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., & Goebel, R. (2010). From Coarse to Fine? Spatial and Temporal Dynamics of Cortical Face Processing. *Cerebral Cortex*, 21(2), 467-476. doi: 10.1093/cercor/bhq112
- Goffaux, V., van Zon, J., & Schiltz, C. (2011). The horizontal tuning of face perception relies on the processing of intermediate and high spatial frequencies. *Journal Of Vision*, 11(10), 1-1. doi: 10.1167/11.10.
- Guyader, N., Chauvin, A., Boucart, M., & Peyrin, C. (2017). Do low spatial frequencies explain the extremely fast saccades towards human faces?. *Vision Research*, 133, 100-111. doi: 10.1016/j.visres.2016.12.019
- Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends In Cognitive Sciences*, 4(6), 223-233. doi: 10.1016/s1364-6613(00)01482-0
- Hirvenkari, L., Ruusuvuori, J., Saarinen, V., Kivioja, M., Peräkylä, A., & Hari, R. (2013). Influence of Turn-Taking in a Two-Person Conversation on the Gaze of a Viewer. *Plos ONE*, 8(8), e71569. doi: 10.1371/journal.pone.0071569
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194-203. doi: 10.1038/35058500
- Jacques, C., & Rossion, B. (2004). Concurrent processing reveals competition between visual representations of faces. *Neuroreport*, 15(15), 2417-2421. doi: 10.1097/00001756-200410250-00023
- Janik, S., Wellens, A., Goldberg, M., & Dell'Osso, L. (1978). Eyes as the Center of Focus in the Visual Examination of Human Faces. *Perceptual And Motor Skills*, 47(3), 857-858. doi: 10.2466/pms.1978.47.3.857
- Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). Learning to predict where humans look. *2009 IEEE 12Th International Conference On Computer Vision*. doi: 10.1109/iccv.2009.5459462
- Kanwisher, N., McDermott, J., & Chun, M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal Of Neuroscience*, 17(11), 4302-4311. doi: 10.1523/jneurosci.17-11-04302.1997
- Kauffmann, L., Peyrin, C., Chauvin, A., Entzmann, L., Breuil, C., & Guyader, N. (2019). Face perception influences the programming of eye movements. *Scientific Reports*, 9(1). doi: 10.1038/s41598-018-36510-0
- Kirchner, H., & Thorpe, S. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, 46(11), 1762-1776. doi: 10.1016/j.visres.2005.10.002
- Lewis, M., & Edmonds, A. (2003). Face Detection: Mapping Human Performance. *Perception*, 32(8), 903-920. doi: 10.1068/p5007

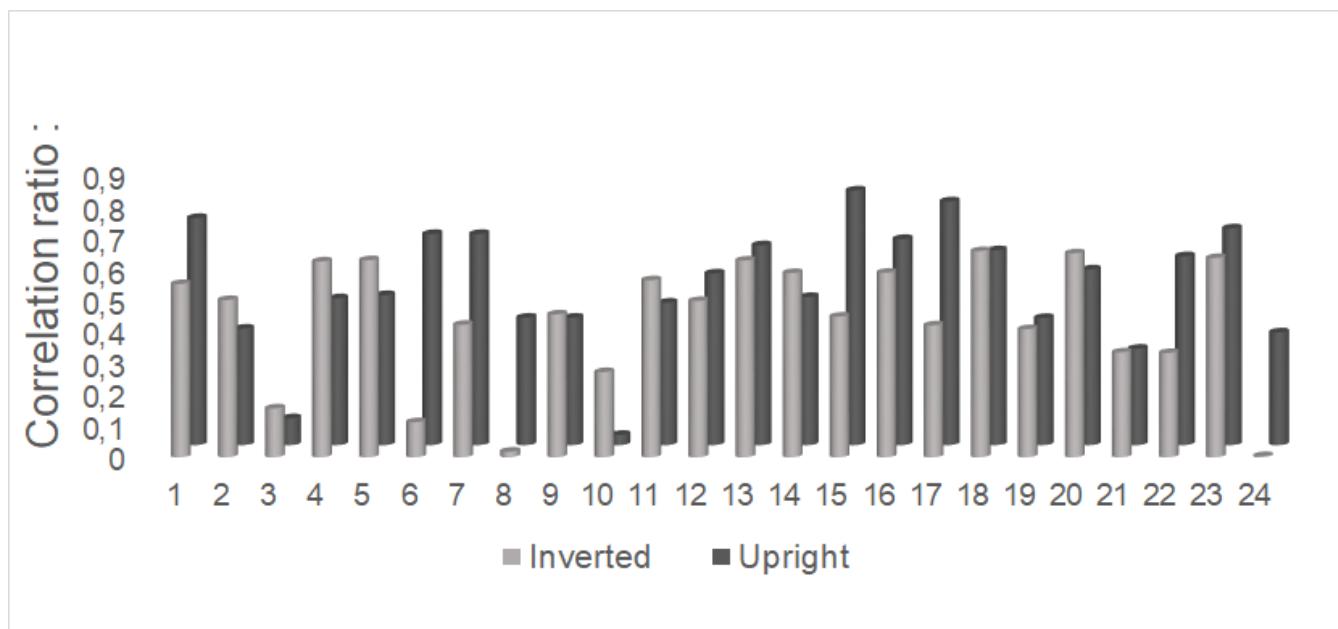
- Marat, S., Rahman, A., Pellerin, D., Guyader, N., & Houzet, D. (2012). Improving Visual Saliency by Adding ‘Face Feature Map’ and ‘Center Bias’. *Cognitive Computation*, 5(1), 63-75. doi: 10.1007/s12559-012-9146-3
- Maurer, D., Grand, R., & Mondloch, C. (2002). The many faces of configural processing. *Trends In Cognitive Sciences*, 6(6), 255-260. doi: 10.1016/s1364-6613(02)01903-4
- Mazard, A., Schiltz, C., & Rossion, B. (2006). Recovery from adaptation to facial identity is larger for upright than inverted faces in the human occipito-temporal cortex. *Neuropsychologia*, 44(6), 912-922. doi: 10.1016/j.neuropsychologia.2005.08.015
- McKone, E., Davies, A., Darke, H., Crookes, K., Wickramariyaratne, T., & Zappia, S. et al. (2013). Importance of the Inverted Control in Measuring Holistic Face Processing with the Composite Effect and Part-Whole Effect. *Frontiers In Psychology*, 4. doi: 10.3389/fpsyg.2013.00033
- McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review*, 16(5), 778-797. doi: 10.3758/pbr.16.5.778
- Mondloch, C., Le Grand, R., & Maurer, D. (2002). Configural Face Processing Develops more Slowly than Featural Face Processing. *Perception*, 31(5), 553-566. doi: 10.1068/p3339
- Morton, J., & Johnson, M. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, 98(2), 164-181. doi: 10.1037//0033-295x.98.2.164
- Nelson, C.A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, 10, 3–18.
- Nguyen, M., Matsumoto, J., Hori, E., Maior, R., Tomaz, C., & Tran, A. et al. (2014). Neuronal responses to face-like and facial stimuli in the monkey superior colliculus. *Frontiers In Behavioral Neuroscience*, 8. doi: 10.3389/fnbeh.2014.00085
- Peterson, M., Rhodes, G. (2006). Oxford University Press Inc. (2006). Perception of Faces, Objects, and Scenes. [Erscheinungsort nicht ermittelbar].
- Peterson, M., & Eckstein, M. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings Of The National Academy Of Sciences*, 109(48), E3314-E3323. doi: 10.1073/pnas.1214269109
- Pelli, D. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437-442. doi: 10.1163/156856897x00366
- Rakover, S., & Teucher, B. (1997). Facial inversion effects: Parts and whole relationship. *Perception & Psychophysics*, 59(5), 752-761. doi: 10.3758/bf03206021
- Richler, J., Cheung, O., & Gauthier, I. (2011). Holistic Processing Predicts Face Recognition. *Psychological Science*, 22(4), 464-471. doi: 10.1177/0956797611401753

- Richler, J., Mack, M., Gauthier, I., & Palmeri, T. (2009). Holistic processing of faces happens at a glance. *Vision Research*, 49(23), 2856-2861. doi: 10.1016/j.visres.2009.08.025
- Richler, J., Wong, Y., & Gauthier, I. (2011). Perceptual Expertise as a Shift From Strategic Interference to Automatic Holistic Processing. *Current Directions In Psychological Science*, 20(2), 129-134. doi: 10.1177/0963721411402472
- Rosson, B., & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: Evidence from the face composite illusion. *Journal Of Vision*, 8(4), 3. doi: 10.1167/8.4.3
- Rosson, B., & Gauthier, I. (2002). How Does the Brain Process Upright and Inverted Faces?. *Behavioral And Cognitive Neuroscience Reviews*, 1(1), 63-75. doi: 10.1177/1534582302001001004
- Rosson, B., Gauthier, I., Tarr, M., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects. *Neuroreport*, 11(1), 69-72. doi: 10.1097/00001756-200001170-00014
- Rosson, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage*, 39(4), 1959-1979. doi: 10.1016/j.neuroimage.2007.10.011
- Silson, E., Groen, I., Kravitz, D., & Baker, C. (2016). Evaluating the correspondence between face-, scene-, and object-selectivity and retinotopic organization within lateral occipitotemporal cortex. *Journal Of Vision*, 16(6), 14. doi: 10.1167/16.6.14
- Tanaka, J., & Farah, M. (1993). Parts and Wholes in Face Recognition. *The Quarterly Journal Of Experimental Psychology Section A*, 46(2), 225-245. doi: 10.1080/14640749308401045
- Thompson, P. (1980). Margaret Thatcher: A New Illusion. *Perception*, 9(4), 483-484. doi: 10.1068/p090483
- White, B., Kan, J., Levy, R., Itti, L., & Munoz, D. (2017). Superior colliculus encodes visual saliency before the primary visual cortex. *Proceedings Of The National Academy Of Sciences*, 114(35), 9451-9456. doi: 10.1073/pnas.1701003114
- Xu, Y., Liu, J., & Kanwisher, N. (2005). The M170 is selective for faces, not for expertise. *Neuropsychologia*, 43(4), 588-597. doi: 10.1016/j.neuropsychologia.2004.07.016
- Yin, R. (1969). Looking at upside-down faces. *Journal Of Experimental Psychology*, 81(1), 141-145. doi: 10.1037/h0027474
- Yovel, G., & Kanwisher, N. (2004). Face Perception Domain Specific, Not Process Specific. *Neuron*, 44(5), 889-898. doi: 10.1016/s0896-6273(04)00728-7
- Yovel, G., & Kanwisher, N. (2005). The Neural Basis of the Behavioral Face-Inversion Effect. *Current Biology*, 15(24), 2256-2262. doi: 10.1016/j.cub.2005.10.072

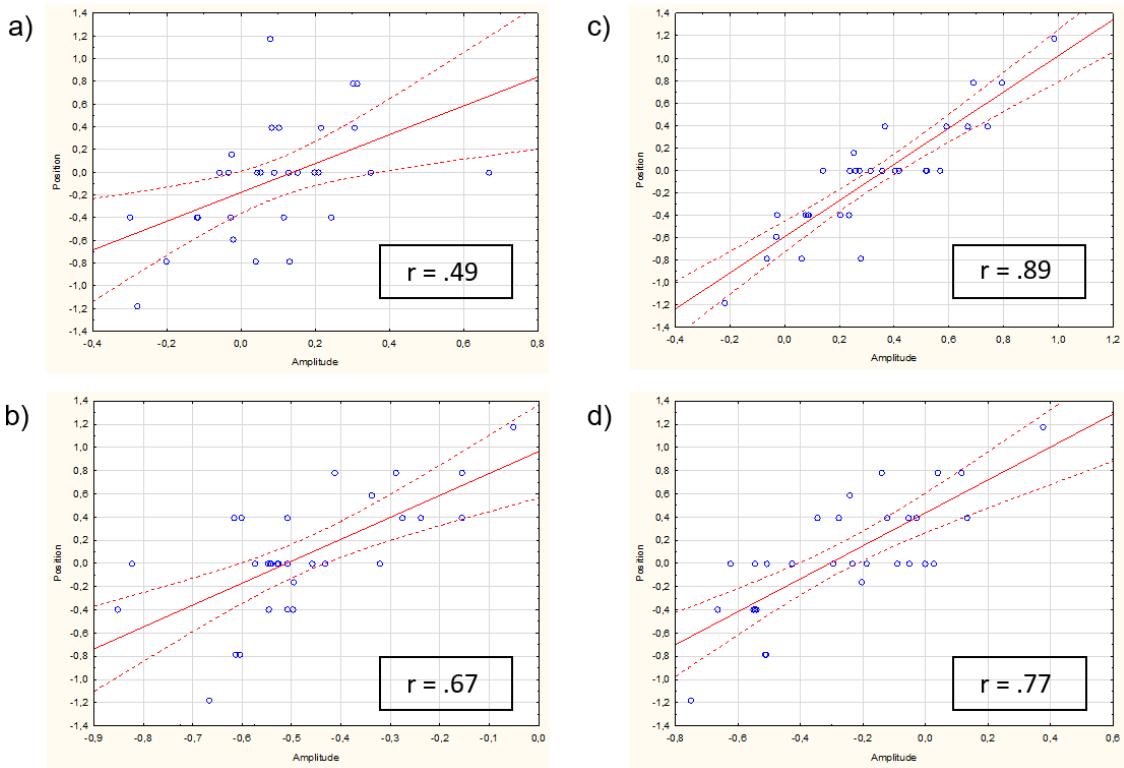
Annexes



ANNEXE I. Sensibilité (d') des participants au biais des visages. Les valeurs positives indique un défaut d'autant plus important que l'inhibition des saccades vers les visages a échoué.



ANNEXE II. Corrélations entre la distance relative des yeux dans les visage et l'amplitude Y des saccades correctes pour chaque participant, en fonction du type de présentation.



ANNEXE III. Corrélation entre la position relative des yeux sur les images et les amplitudes moyennes réalisées sur chaque image (en degrés) (a) et (b) pour une présentation dans l'hémichamp gauche et (c) et (d) pour une présentation dans l'hémichamp droit. (a) et (c) pour une présentation à l'endroit et (b) et (d) pour une présentation à l'envers.

Tableau 2

Récapitulatif des résultats de l'expérience 1 et 2

Condition expérimentale	Visages				Véhicules			
	À l'endroit	À l'envers	Scramble 1×2	Scramble 3×3	À l'endroit	À l'envers	Scramble 1×2	Scramble 3×3
Pourcentage d'erreur moyen (%ER)	9,37 ± 8,05	12,37 ± 8,62	14,46 ± 10,31	13,89 ± 10,25	15,17 ± 9,36	22,75 ± 11	21,12 ± 11,78	25,35 ± 10,79
								29,57 ± 11,16
								32,09 ± 13,52
Temps de réponse saccadique moyen en milliseconde (TRS)	172 ± 13	175 ± 17	177 ± 20	184 ± 23	183 ± 21	198 ± 27 m	200 ± 26	205 ± 32
								214 ± 30
								214 ± 23
Amplitude (X) moyenne en degrés	7,63 ± 2,53	7,51 ± 2,29	7,52 ± 0,71	5,88 ± 0,86	5,69 ± 0,77	6,88 ± 2,68	6,98 ± 2,34	6,71 ± 0,75
								5,57 ± 0,82
								5,61 ± 0,73

ANNEXE IV. Tableau récapitulatif de l'expérience 1 et 2