

Université de Montréal

Information utile à la catégorisation du sexe des visages

par

Nicolas Dupuis-Roy

Département de Psychologie

Faculté des Arts et Sciences

Thèse présentée à la Faculté des Arts et Sciences

en vue de l'obtention du grade de Ph.D

en Psychologie

option Sciences Cognitives et Neuropsychologie

Septembre, 2014

© Nicolas Dupuis-Roy, 2014

Résumé

La présente thèse avait pour mandat d'examiner la question suivante : quels sont les indices visuels utilisés pour catégoriser le sexe d'un visage et comment sont-ils traités par le cerveau humain? La plupart des études examinant l'importance de certaines régions faciales pour la catégorisation du sexe des visages présentaient des limites quant à leur validité externe. L'article 1 visait à investiguer l'utilisation des indices achromatiques et chromatiques (sur l'axe xy) dans un contexte de plus grande validité externe. Pour ce faire, nous avons utilisé la technique Bubbles afin d'échantillonner l'espace xy de visages en couleurs n'ayant subi aucune transformation. Afin d'éviter les problèmes liés à la grande répétition des mêmes visages, nous avons utilisé un grand nombre de visages (c.-à-d. 300 visages caucasiens d'hommes et de femmes) et chaque visage n'a été présenté qu'une seule fois à chacun des 30 participants. Les résultats indiquent que la région des yeux et des sourcils—probablement dans le canal blanc-noir—est l'indice le plus important pour discriminer correctement le genre des visages; et que la région de la bouche—probablement dans le canal rouge-vert—est l'indice le plus important pour discriminer rapidement et correctement le genre des visages.

Plusieurs études suggèrent qu'un indice facial que nous n'avons pas étudié dans l'article 1—les distances interattributs—est crucial à la catégorisation du sexe. L'étude de Taschereau et al. (2010) présente toutefois des données allant à l'encontre de cette hypothèse : les performances d'identification des visages étaient beaucoup plus faibles lorsque seules les distances interattributs réalistes étaient disponibles que lorsque toutes les autres informations faciales à l'exception des distances interattributs réalistes étaient disponibles. Quoi qu'il en soit, il est possible que la faible performance observée dans la condition où seules les distances interattributs étaient disponibles soit explicable non par une incapacité d'utiliser ces indices efficacement, mais plutôt par le peu d'information contenue dans ces indices. L'article 2 avait donc comme objectif principal d'évaluer l'efficacité—une mesure de performance qui compense pour la faiblesse de l'information disponible—des distances interattributs réalistes pour la catégorisation du sexe des visages chez 60 participants. Afin de maximiser la validité externe, les distances interattributs manipulées respectaient la distribution et la matrice de

covariance observées dans un large échantillon de visages (N=515). Les résultats indiquent que les efficacités associées aux visages ne possédant que de l'information au niveau des distances interattributs sont un ordre de magnitude plus faibles que celles associées aux visages possédant toute l'information que possèdent normalement les visages sauf les distances interattributs et donnent le coup de grâce à l'hypothèse selon laquelle les distances interattributs seraient cruciales à la discrimination du sexe des visages.

L'article 3 avait pour objectif principal de tester l'hypothèse formulée à la fin de l'article 1 suivant laquelle l'information chromatique dans la région de la bouche serait extraite très rapidement par le système visuel lors de la discrimination du sexe. Cent douze participants ont chacun complété 900 essais d'une tâche de discrimination du genre pendant laquelle l'information achromatique et chromatique des visages était échantillonnée spatiotemporellement avec la technique Bubbles. Les résultats d'une analyse présentée en Discussion seulement confirment l'utilisation rapide de l'information chromatique dans la région de la bouche. De plus, l'utilisation d'un échantillonnage spatiotemporel nous a permis de faire des analyses temps-fréquences desquelles a découlé une découverte intéressante quant aux mécanismes d'encodage des informations spatiales dans le temps. Il semblerait que l'information achromatique et chromatique à l'intérieur d'une même région faciale est échantillonnée à la même fréquence par le cerveau alors que les différentes parties du visage sont échantillonnées à des fréquences différentes (entre 6 et 10 Hz). Ce code fréquentiel est compatible avec certaines évidences électrophysiologiques récentes qui suggèrent que les parties de visages sont « multiplexées » par la fréquence d'oscillations transitoires synchronisées dans le cerveau.

Mots-clés : Perception visuelle, stratégies d'extraction de l'information, discrimination du genre des visages, catégorisation du sexe, Bulles, efficacité, analyse de l'observateur idéal, couleur.

Abstract

In this thesis, we asked which face cues are used for gender discrimination and how they are processed by the human brain. The few studies that have tried to answer this question used only a small set of grayscale stimuli, often distorted and presented a large number of times. In Article 1, we reassessed the importance of chromatic and achromatic facial cues for gender categorisation in a more realistic setting. We used the Bubbles technique to sample the xy plane of a set of 300 color photographs of Caucasian faces, each presented only once to 30 participants. Results show that the region of the eyes and the eyebrows—probably in the light-dark channel—is the most important facial cue for accurate gender discrimination; and that the mouth region is also driving fast correct responses (but not fast incorrect responses)—the gender discrimination information in the mouth region is concentrated in the red-green color channel.

Several studies suggest that a cue which we did not examine in Article 1—interattribute distances (e.g. the interpupil distance)—is critical for gender discrimination. A recent study by Taschereau-Dumouchel et al. (2010) challenged this idea by showing that participants were nearly at chance when asked to identify faces on the sole basis of real-world interattribute distances, while they were nearly perfect when all other facial cues were shown. However, it remains possible that humans are highly tuned to interattribute distances but that the information conveyed by these cues is scarce. In Article 2, we tested this hypothesis by contrasting the efficiencies—a measure of performance that factors out task difficulty—of 60 observers in six face-gender categorization tasks. Most importantly, efficiencies for faces that varied only in terms of their interattribute distances were an order of magnitude lower than efficiencies for faces that varied in all respects except their interattribute distances or in all respects. These results provide a definitive blow to the idea that real-world interattribute distances are critical for face gender categorization.

In Article 3, we tested and confirmed the hypothesis formulated at the end of Article 1 according to which observers would rapidly extract chromatic information in the mouth region

during sex categorization. One hundred and twelve participants each performed 900 trials of a face gender categorization task in which the achromatic and isoluminant chromatic content of faces were partially revealed in space and time with Bubbles. This experiment also allowed us to perform time-frequency analyses, which showed that achromatic and isoluminant chromatic information within the same facial part were sampled by the brain at the same frequency, whereas different facial parts were sampled at distinct frequencies (ranging from 6 to 10 Hz). This sampling code is consistent with recent electrophysiological evidence suggesting that facial features are multiplexed by the frequency of transient synchronized oscillations in the brain.

Keywords : Visual perception, Information extraction strategies, face gender discrimination, sex categorization, Bubbles, Efficiency, Ideal observer analysis, Color.

Table des matières

RÉSUMÉ	III
ABSTRACT	VI
TABLE DES MATIÈRES	VIII
LISTE DES FIGURES	XI
LISTE DES ABBRÉVIATIONS	XIV
REMERCIEMENTS	XV
CHAPITRE 1: INTRODUCTION GÉNÉRALE	19
1.1. AVANT-PROPOS	21
1.1.2. <i>Catégorisation du sexe des visages: Importance et caractéristiques neurocognitives</i>	22
1.1.3. <i>Question principale et enjeux méthodologiques généraux</i>	25
1.1.4. <i>Présentation générale de la thèse</i>	31
1.2. DIMORPHISMES SEXUELS FACIAUX ET LEUR IMPORTANCE PERCEPTIVE	32
1.2.1. <i>Indices chromatiques et achromatiques</i>	33
1.2.2. <i>Indices faciaux relationnels</i>	51
1.3. PRÉSENTATION DES ARTICLES DE LA THÈSE	62
CHAPITRE 2: ARTICLE 1	65
2.1. ABSTRACT	67
2.2. INTRODUCTION	68
2.3. METHODS	69
2.3.1. <i>Participants</i>	69
2.3.2. <i>Stimuli</i>	69
2.3.3. <i>Apparatus</i>	70
2.3.3. <i>Procedure</i>	71
2.4.1. <i>Linear classification image analysis</i>	71
2.4.1. <i>Beyond linear classification Images</i>	72
2.4. CONCLUSION	75
2.9. FIGURE AND TABLES	82

CHAPITRE 3: ARTICLE 2	87
3.1. ABSTRACT.....	90
3.2. INTRODUCTION.....	91
3.3. METHODS.....	93
3.3.1 <i>Participants</i>	93
3.3.2 <i>Stimuli</i>	93
3.3.3 <i>Apparatus</i>	94
3.3.4 <i>Signal and noise adjustments</i>	94
3.3.5 <i>Procedure</i>	95
3.4. RESULTS.....	95
3.4.1 <i>Efficiencies</i>	96
3.5. DISCUSSION.....	97
3.6. REFERENCES.....	99
3.7. FOOTNOTES.....	105
3.8. FIGURE.....	106
3.9. ACKNOWLEDGEMENTS.....	107
3.10. SUPPLEMENTARY MATERIAL.....	108
3.10.1 <i>Preliminary experiments for signal and noise adjustments</i>	108
CHAPITRE 4: ARTICLE 3	113
4.1. ABSTRACT.....	116
4.2. INTRODUCTION.....	117
4.3. METHODS.....	119
4.3.1 <i>Participants</i>	119
4.3.2 <i>Face database</i>	119
4.3.3 <i>Stimuli</i>	119
4.3.4 <i>Apparatus</i>	120
4.3.5 <i>Procedure</i>	121
4.4. RESULTS AND DISCUSSION.....	121
4.4.1 <i>Latency and frequency of transient synchronized information sampling</i>	121
4.6. AKNOWLEDGEMENTS.....	125
4.7. FIGURE LEGENDS.....	126
4.8. REFERENCES.....	130

CHAPITRE 5: DISCUSSION GÉNÉRALE	135
5.1. AVANT-PROPOS.....	137
5.2. UTILISATION DES INDICES FACIAUX CHROMATIQUES ET ACHROMATIQUES	137
5.3. IMPORTANCE RELATIVE ET RÔLE DES INDICES CHROMATIQUES	144
5.3.1. <i>Hypothèse sur l'utilisation rapide des informations chromatiques dans la région orale.</i> ..	146
5.3.2. <i>Rôle de la couleur et rapidité de traitement</i>	151
5.3.3. <i>Hypothèses sur l'importance de la région orale</i>	152
5.3.4. <i>Chromatique vs achromatique</i>	154
5.4. MÉCANISME D'ENCODAGE DES ATTRIBUTS FACIAUX	154
5.5. IMPORTANCE RELATIVE DES INDICES RELATIONNELS DE SECOND ORDRE	156
5.5.1 <i>Importance relative des distances interattributs</i>	157
5.5.2 <i>Importance des autres indices relationnels de deuxième ordre</i>	159
5.6. LIMITES ET PERSPECTIVES FUTURES	165
5.6.1. <i>Critiques et limites de la technique Bubbles</i>	166
5.6.2. <i>Limite à la validité externe</i>	170
5.6.3. <i>Pistes de recherches futures</i>	172
5.7 CONCLUSION.....	174
RÉFÉRENCES.....	179

Liste des figures

Chapitre 1- Introduction générale

Figure 1. Exemples de dimorphismes sexuels.....	22
Figure 2. Schéma de la théorie d'Ellis et Bruce & Young.....	25
Figure 3. Schéma du cadre théorique <i>RAP</i>	27
Figure 4. Distribution spatiale des informations sexuelles	36
Figure 5. Différence de luminance locale entre les visages des deux sexes	38
Figure 6. Transformation des contrastes locaux	39
Figure 7. Stimuli employés par les études sur l'utilisation des indices achromatiques chromatiques	41
Figure 8. Principaux résultats des études sur l'utilisation des régions faciales	48
Figure 9. Exemples de visages composites	52
Figure 10. Repères craniofaciaux utilisés par Farkas et al. (1994, 1987)	55
Figure 11. Repères craniofaciaux de surface utilisés par Burton et al. (1993)	57
Figure 12. Exemples de stimuli utilisés par Taschereau-Dumouchel et al. (2010) ...	58
Figure 13. Tailles d'effet associée aux DIA et quatre autres dimorphismes sexuels ..	60

Chapitre 2- Article 1

Figure 1. Examples of faces sampled or not with Bubbles.....	84
Figure 2. Regions significantly used to discriminate the sex of faces.....	85

Figure 3. Distribution of surface sexual information on faces	86
---	----

Chapitre 3- Article 2

Figure 1. Examples of face stimuli that were used and main results of the article....	109
---	-----

Chapitre 4- Article 3

Figure 1. Stimulus generation process and events in a trial	131
Figure 2. Time-frequency spectrum for chromatic and achromatic information in each facial part	132
Figure 3. Peak amplitudes from time-frequency spectra	133

Chapitre 5- Discussion générale

Figure 1. Sommaire des résultats des Articles 1 et 2	144
Figure 2. Utilité de 9 régions faciales d'intérêt	146
Figure 3. Volume de classification chromatique et achromatique	152
Figure 4. Régions faciales significatives et profile temporel de ces régions	153
Figure 5. Résultats principaux de Schyns, Bonnar et Gosselin (2002) et stimuli de Gold, Mundy et Tjan (2012)	164
Figure 6. Schéma du design expérimental de Richlet, Cheung et Gauthier (2011) ...	169
Figure 7. Exemple de Bubble porn.	171
Figure 8. Profil spectral moyens de visages filtrés et non-filtrés par les Bulles	172

Figure 9. Profil spectral moyen de séries temporelles filtrées et non-filtrées par des bulles.....174

Liste des abréviations

CI : Classification image

IC : Image de classification

DIA : Distances interattributs

IAD : Interattribute distances

Remerciements

Cette thèse ne se serait jamais achevée sans l'encouragement, le soutien et l'apport multiforme de nombreuses personnes qui me sont chères. D'abord, j'aimerais souligner le rôle important de plusieurs de mes ami(e)s. À Stephen, à Christiane et à toute la famille Lemieux, j'aimerais dire merci pour tous ces succulents repas dominicaux et ces rendez-vous sportifs. Vous m'avez fait sourire dans les moments de découragement et n'avez cessé de m'aimer malgré mes nombreux défauts. J'aimerais témoigner toute ma reconnaissance à Dr Dugs et ma sœur de cheveux pour m'avoir écouté, soutenu, compris, et surtout pour avoir alimenté ma folie tout en nourrissant cette source inépuisable et essentielle de procrastination que fut la musique. J'aimerais également exprimer la chance que j'ai eu de faire ce stage au BRAMS où j'ai fait la rencontre d'êtres singuliers et irrésistibles : Mona, Pauline et Régis pour n'en nommer que quelques-uns. Vous avez pu me voir évoluer de l'intérieur et m'avez très bien accompagné dans ce long processus. J'aimerais remercier tout particulièrement Mona pour nos pauses café hebdomadaires et ton amitié; Pauline pour ces nombreux « papotages », ces pérégrinations au Renaissance, et le ski de soirée sur le Mont-Royal; et Régis pour sa présence féline, musicale et indescriptible. Par la même occasion, j'aimerais remercier Sébastien pour son authenticité et son amitié, et Maïlis pour son soutien, sa compréhension, son intelligence, sa sensibilité et sa gourmandise... J'aimerais également remercier Pablo de m'avoir enduré toutes ces années, de m'avoir prodigué tous ces conseils en matière de design et d'avoir enjolivé toutes mes affiches scientifiques. Je remercie Steve et Cristina qui m'ont « offert » l'îlot de paix et de tranquillité dont j'avais besoin pour commencer la rédaction de cette thèse. Je veux aussi remercier toute ma famille, mes parents Normand, Lise, Yvette et Yvon, ainsi que mes frères Christian, Jérémie, Jean-François et Dominic qui m'ont candidement exprimé leurs encouragements, leurs impatiences, leurs questionnements, leurs découragements, et qui n'ont jamais cessé de croire en moi et en la finalité possible de cette aventure sans fin. Également, un merci tout spécial à Lise et à Normand qui m'ont offert l'accès à ce havre de paix au Lac Cayamant. C'est là que tout a débuté et c'est là que s'est achevée cette thèse.

J'aimerais remercier Marc Schönwiesner qui a su m'influencer au bon moment et m'ouvrir à d'autres visions scientifiques ainsi que Mélanie Vachon pour son support idéologique et psychologique.

Comme mon passage au laboratoire de perception fut plutôt long, j'ai pu y fréquenter plusieurs êtres très attachants avec qui j'ai partagé des tranches de vie importantes et inoubliables. Alan, Alexandre, Colin, Cynthia, Caroline, Catherine, Daniel, Éric, Frédéric JAM, Isabelle, Jude, Karine, Kim, Laurent, Lucia, Mélanie, Simon, Solène, Sylvain, Verena, Xavier et Zakia, c'est à vous que je pense. J'aimerais d'abord remercier mes collègues de bureau, Caroline, Catherine, Cynthia et Sylvain qui ont su non seulement endurer mes réflexions absurdes, mes flatulences incessantes et mon humour douteux, mais qui ont également pu me remettre sur le droit chemin pendant mes nombreux moments d'égarement mental. Merci Caro de m'avoir suivi dans mes réflexions absurdes et d'avoir pu y trouver source d'amusement; merci Catherine de m'avoir écouté sans jugement et de m'avoir soutenu pendant ces nombreuses épreuves; merci Cynthia pour ton humour et ton amitié durant toutes ces belles années; merci Karine pour ton soutien psychologique, ces potins juteux, ces conversations rocambolesques, ta personnalité enjouée et... ces nombreux nains de jardin. J'aimerais aussi remercier Daniel qui fut présent au tout début de cette longue et interminable aventure, et sans les encouragements duquel je n'aurais probablement jamais commencé le doctorat. Merci d'avoir cru en mes capacités plus que moi, et de m'avoir assisté et conseillé durant la gestation de nombreux projets. Merci à Xavier pour ton amitié et pour ces nombreuses marches au café La Retenue durant lesquelles tu auras fait l'essai de milliers (voire de millions ?) de calembours aussi douteux qu'inventifs. Merci à Alex, Simon et Solène pour tous ces instants de procrastination savoureux, et à Laurent pour ces discussions théoritico-potineuses. J'aimerais aussi remercier Véronique d'avoir pris le temps de lire, de commenter et de corriger cette thèse.

Finalement, j'aimerais souligner l'apport distinctif de mon directeur de thèse, Frédéric Gosselin. D'abord, je dois te remercier de m'avoir convaincu d'entreprendre cette longue

aventure. La liberté que tu m'as offerte tant sur le plan idéologique que terrestre a été déterminante dans mon choix de poursuivre au doctorat et d'y rester. Tu as su voir mon potentiel et le cultiver pendant ces longues années, et pour cela seulement, je te serai éternellement reconnaissant. Tes idées singulières, ta créativité scientifique, ton côté irrévérencieux et marginal font de toi un libre penseur dont l'influence s'est exercée sur moi au-delà du cadre académique. C'est ainsi qu'au fil des années tu es passé de directeur de thèse à associé et ami. Même si parfois ces multiples rôles ont pu mener à des prises de bec, au final, elles se sont soldées par un approfondissement de notre collaboration. En conclusion, j'aimerais te remercier de ton soutien psychologique et de tes encouragements durant mes nombreuses pannes d'inspiration et de motivation. Je dois aussi te remercier du support financier et de m'avoir envoyé un peu partout sur la planète pour disséminer la bonne nouvelle scientifique que sont les Bulles.

Enfin, cette thèse n'aurait jamais été possible sans le soutien financier du Fonds de Recherche en Nature et Technologies, du Conseil de Recherches en Sciences Humaines et du Centre de Recherche en Neuropsychologie et Cognition.

Chapitre 1: Introduction générale

1.1. Avant-propos

« Je m'appuie de tout mon poids sur le rebord de faïence, j'approche mon visage de la glace jusqu'à la toucher. Les yeux, le nez et la bouche disparaissent : il ne reste plus rien d'humain. Des rides brunes de chaque côté du gonflement fiévreux des lèvres, des crevasses, des taupinières. Un soyeux duvet blanc court sur les grandes pentes des joues, deux poils sortent des narines : c'est une carte géologique en relief. Et malgré tout, ce monde lunaire m'est familier. Je ne peux pas dire que j'en reconnaisse les détails. Mais l'ensemble me fait une impression de déjà-vu qui m'engourdit : je glisse doucement dans le sommeil. » Sartre, (1938)

Non seulement peut-il faire office de carte, comme l'évoquait si bien Sartre, mais le visage représente aussi le véhicule des informations sociales. En effet, il suffit d'un simple coup d'oeil sur un visage pour en saisir l'émotion, l'âge, le sexe, l'ethnie ou pour déterminer s'il est digne de confiance. Cette aptitude est d'autant plus remarquable qu'elle s'applique à des conditions visuelles variables, dynamiques et souvent bruitées, que ce soit à la pénombre ou lorsque le visage est partiellement voilé; sous la pluie ou à des distances de plusieurs mètres; ou encore dans une foule et à travers le feuillage. Le système visuel humain semble donc avoir évolué de sorte à extraire efficacement les informations faciales à travers une multitude de conditions visuelles, et ce, malgré une série de contraintes biologiques et environnementales.

L'une des informations faciales pour laquelle l'expression phénotypique et le traitement visuel ont dû être priorisés très tôt dans l'évolution humaine, concerne les dimorphismes sexuels. Ces derniers désignent toutes les différences morphologiques entre les mâles et les femelles. Chez l'*homo sapiens*, le mâle adulte possède une taille et une masse en moyenne supérieure à celle de la femelle, ce qui nécessite des capacités cardiorespiratoires accrues. De telles différences se répercutent sur la morphologie du visage qui, en plus d'être de plus grande taille chez l'homme, présente plusieurs traits distinctifs (voir Figure 1). Par exemple, le nez et le nasopharynx sont plus larges chez l'homme, ce qui a pour effet de rehausser la racine du nez et les zones adjacentes du visage telles que les arcades sourcilières. C'est pourquoi, selon (Enlow, 1990), les joues des hommes sembleraient moins saillantes et que leurs yeux paraîtraient plus profondément enchâssés dans la boîte crânienne (voir Figure 1b et c). En revanche, la femme prototypique possède un plus petit nez, ce qui lui procure des yeux d'apparence plus grands et des joues plus saillantes. Cette caractéristique des joues va de pair avec une plus grande concentration

de tissus adipeux faciaux (Shepherd, 1981). La plus faible pilosité faciale chez la femme se répercute aussi au niveau de ses sourcils qui sont généralement plus fins, clairsemés et circonflexes. La Figure 1a illustre plusieurs de ces dimorphismes sexuels. La Figure 1b présente un visage mâle (droite) et femelle (gauche) prototypique dont la forme tridimensionnelle a été isolée, tandis que la Figure 1c présente les différences statistiques entre de tels visages. Les images de la Figure 1c montrent bien la protubérance des arcades sourcilières, du nez et du pharynx.

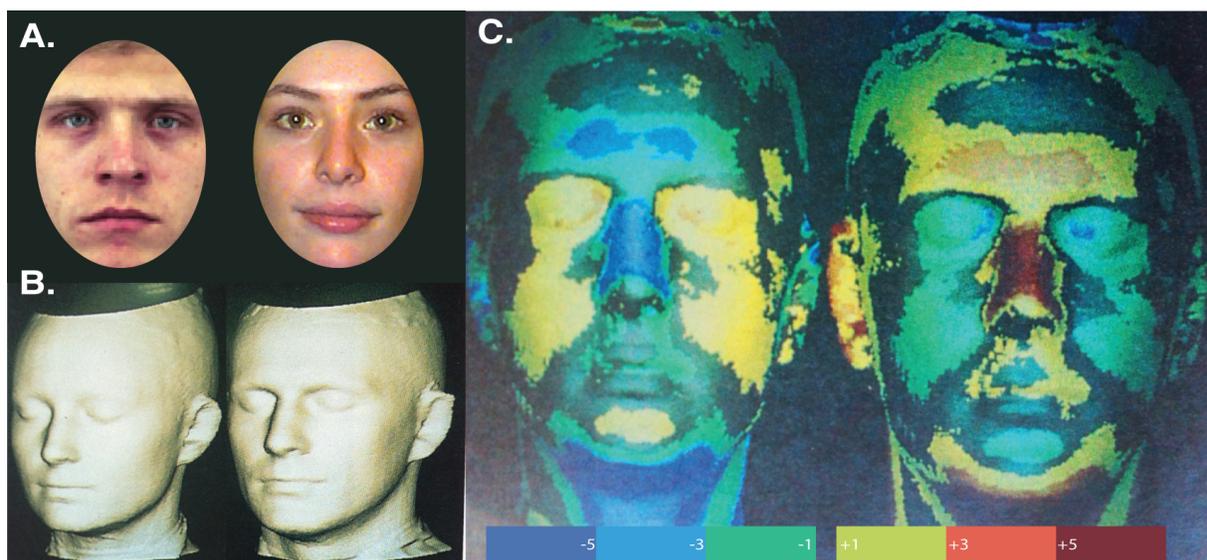


Figure 1. A) Visages caucasiens inexpressifs des deux sexes présentés en plan frontal et alignés sur 12 repères craniofaciaux. B) Forme tridimensionnelle moyenne des visages féminins (gauche) et masculins (droit) créée à l'aide d'un scanner laser (tiré de Bruce et al., 1993). C) Différences statistiques entre la forme tridimensionnelle des visages féminins et masculins présentés en B (tiré de Bruce et al., 1998). Le code de couleur, exprimé en valeurs arbitraires, qualifie le niveau de protubérance (jaune-rouge) et de récession (bleu-violet). Ces statistiques sont juxtaposées au visage féminin moyen (gauche) et au visage masculin moyen (droite).

1.1.2. Catégorisation du sexe des visages: Importance et caractéristiques neurocognitives

La capacité à traiter les dimorphismes sexuels est cruciale à la survie humaine, si ce n'est qu'au niveau de la sélection du partenaire sexuel. En fait, cette capacité perceptive détient un rôle

important dans le développement humain. Par exemple, elle renvoie à la formation des représentations du genre chez l'enfant et occupe, en ce sens, une place centrale dans son développement affectif, cognitif, social et identitaire. Cette capacité est aussi sollicitée dans l'acquisition du langage, et se répercute sur l'ensemble des interactions humaines par le truchement des représentations sociales.

Par son importance biologique et développementale, on pourrait s'attendre à ce le système visuel humain ait évolué de sorte à développer une certaine expertise dans le traitement des dimorphismes sexuels, et ce, précocement dans son ontogenèse. C'est effectivement ce que suggèrent certains travaux. Par exemple, dans une étude menée par Cornell (1974), des enfants âgés de 4.5 à 6 mois ont d'abord été exposés à une série de visages d'un même sexe et ont ensuite été testés sur de nouveaux visages du même sexe et de sexe opposé. Sur la base du regard préférentiel pour le sexe opposé, les chercheurs ont conclu qu'à partir de 5.5 mois, les enfants pouvaient discriminer visuellement les visages d'hommes des visages de femmes. Leinbach et Fagot (1993) ont également observé que dès sept mois, les nouveau-nés présentaient une habituation sélective à la catégorie sexuelle des visages. Wild et Barrett (2000) ont confirmé, chez des enfants de sept ans, la capacité à discriminer le sexe de visages dénués de signes culturels faciaux et non-faciaux ostensibles tels que les bijoux, la coiffure ou le maquillage. À l'âge adulte, les observateurs humains atteignent des performances frôlant la perfection. À titre d'exemple, Bruce et al. (1993) ont enregistré un taux moyen de bonne réponse de 0.96 dans des conditions visuelles restreintes, c.-à-d. avec des visages inexpressifs, présentés de face, dévoilant un minimum de pilosité (c.-à-d. les sourcils) et n'affichant aucun signe sexuel accessoire (par ex. des boucles d'oreille).

D'autres évidences suggèrent qu'un tel jugement est implicite et automatique, c.-à-d. que de bonnes performances peuvent être atteintes sans que l'observateur n'en connaisse la stratégie visuelle sous-jacente et sans qu'il n'ait à déployer d'efforts attentionnels ou mnémoniques particuliers. Une première observation allant en ce sens est que ce jugement facial peut être exécuté très rapidement: Barragan-Jason, Lachat et Barbeau (2012) ont montré, à l'aide d'un paradigme de catégorisation rapide, que la discrimination du sexe des visages était plus rapide ($M_d = 379$ ms, $SD = 34$ ms) et juste ($M = 93.6\%$, $SD = 4.1\%$) que plusieurs autres jugements

faciaux dichotomiques tels que le jugement de familiarité ($Md = 488ms$, $SD = 54ms$; $M = 89.6\%$, $SD = 4.7\%$). D'autres chercheurs ont aussi observé de bonnes performances de discrimination dans des conditions où les visages étaient présentés très brièvement (p. ex. 75 ms) et suivis d'un masque visuel (p. ex. 200ms; voir O'Toole, Peterson & Deffenbacher, 1996). Plus encore, Reddy, Wilken et Koch (2004) ont montré que des participants à qui l'on avait demandé d'effectuer une tâche exigeant que l'attention soit focalisée au centre du champ visuel, étaient encore capables de catégoriser avec justesse le sexe de visages présentés simultanément en périphérie du champ visuel et pendant seulement 26 ms (suivis d'un masque visuel de 200 ms).

Parallèlement à ces observations comportementales, des études utilisant l'électroencéphalographie (EEG) ont rapporté des potentiels évoqués significatifs liés à la catégorisation du sexe des visages à de très faibles latences (voir Mouchetant-Rostaing et al., 2000 ; Mouchetant-Rostaing & Giard, 2003). Par exemple, Mouchetant-Rostaing et Giard (2003) ont enregistré les potentiels évoqués dans deux conditions de visionnement passif où les participants devaient détecter la présence de lunette sur les visages (i.e. un « oddball »). Dans la condition “gender incidental”, on y présentait des visages masculins et féminins, tandis que dans la condition “no-gender” on y affichait des visages d'un seul sexe. Une analyse de la différence “gender incidental” vs “no-gender” indique des différences significatives de potentiels entre 40-90 ms sur des électrodes occipitopariétales gauches. D'autres études en neuroimagerie révèlent que les aires corticales activées par la catégorisation du sexe des visages sont plus postérieures que celles activées lors de l'identification des mêmes visages (Sergent, Ohta & McDonald, 1992). Ces régions incluaient le cunéus et les gyri lingual, occipitotemporal inférieur et occipital latéral, médial et inférieur. En conclusion, l'ensemble de ces résultats suggère que les informations dimorphiques sont traitées tôt dans le système visuel, et ce, de façon quasi automatique et implicite.

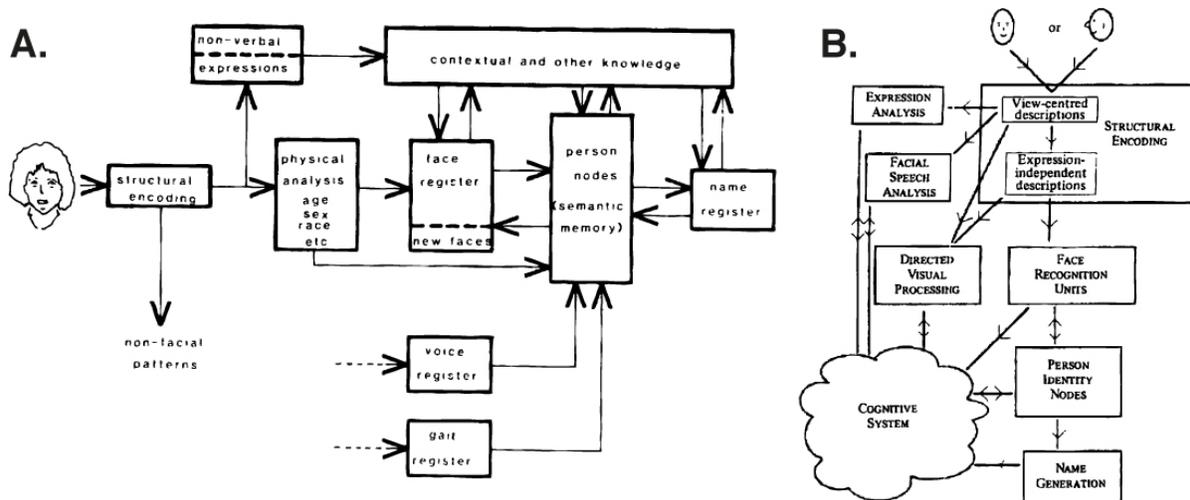


Figure 2. Représentations schématiques de la théorie A) d'Ellis (1986) et de B) Bruce et Young (1986).

Ces conclusions vont dans le sens de modèles importants de la perception des visages qui s'appuient principalement sur l'étude clinique des prosopagnosiques. Par exemple, le modèle d'Ellis (1986; voir Figure 2A) propose qu'après avoir reconnu un objet visuel comme étant un visage, certaines de ses caractéristiques physiques seraient analysées. Cette analyse serait indépendante de l'analyse de l'expression faciale et surviendrait avant le registre voué aux processus de reconnaissance faciale. Le modèle de Bruce et Young (1986) soutient aussi l'idée d'une certaine indépendance fonctionnelle entre le traitement des informations sémantiques dérivées des caractéristiques physiques des visages (p. ex. le sexe, l'âge, la race) et le traitement des informations liées aux caractéristiques stables (p. ex. l'identité) ou dynamiques (p. ex. l'expression) des visages. Selon ce modèle (voir Figure 2B), le traitement des caractéristiques physiques des visages non familiers renverrait aux processus visuels dirigés (c.-à-d. impliquant l'attention sélective) qui se déploieraient de façon indépendante et parallèlement aux autres processus faciaux voués à l'analyse des expressions faciales, des gestes faciaux langagiers et de l'identité faciale. Ces modèles proposent donc tous deux l'indépendance du traitement des informations faciales liées aux catégories physiques telle que le sexe, et le modèle d'Ellis met de l'avant la précocité de ce même traitement.

1.1.3. Question principale et enjeux méthodologiques généraux

Les dimorphismes sexuels décrits en avant-propos (section 1.1.1) ne représentent qu'une infime fraction des différences morphologiques visibles entre les visages adultes féminins et masculins. En fait, il existe aussi des différences subtiles au niveau du positionnement relatif des principaux attributs faciaux (p. ex. nez, bouche, yeux et sourcils), au niveau de la pigmentation de la peau, sa texture, etc. Compte tenu des ressources cognitives et perceptives limitées du cerveau humain, la question qui s'ensuit naturellement pour le chercheur en perception visuelle est de savoir lesquels, parmi l'éventail de ces signaux visuels, sont utilisés par l'observateur humain afin de qualifier le sexe d'un visage. Ce questionnement, central à la présente thèse, renvoie à des enjeux méthodologiques et épistémologiques qu'il convient d'expliquer avant de poursuivre la présentation des thèmes particuliers. Afin de structurer la présentation de ces enjeux, nous aurons recours au cadre théorique $R \otimes A \approx P$ (Gosselin & Schyns, 2002).

En 2002, Gosselin et Schyns proposent un cadre théorique simple clarifiant le problème informationnel en catégorisation visuelle. Sa formulation est la suivante: $R \otimes A \approx P$, où R correspond aux représentations visuelles internes, A signifie l'information disponible à l'observateur pour faire une tâche de catégorisation donnée, et P représente l'information *puissante*, soit la portion des informations disponibles A incluse dans les représentations internes R (voir Figure 3A). Les opérateurs \otimes et \approx symbolisent une interaction et une similarité respectivement.

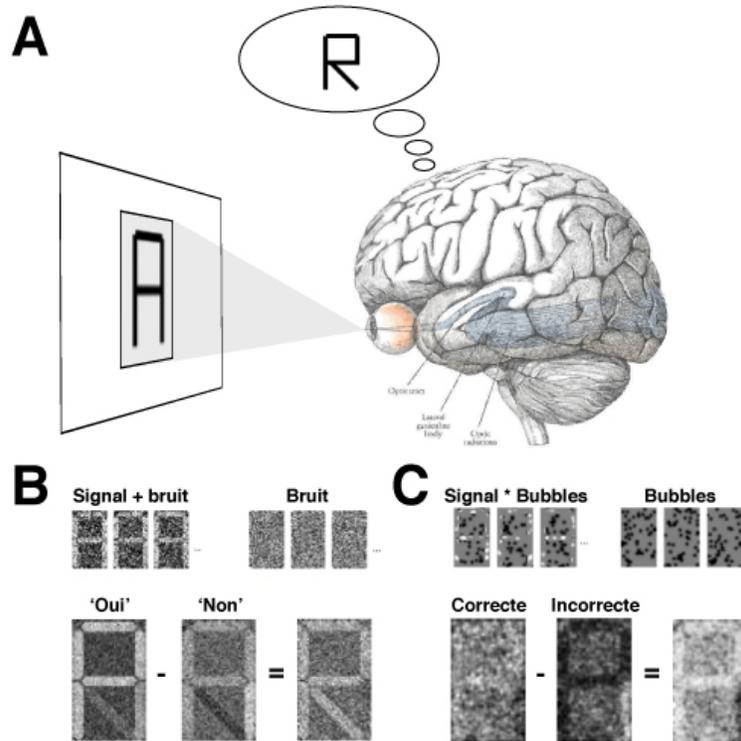


Figure 3. A) Schéma illustrant la nature distincte des informations impliquées dans une tâche de catégorisation. A symbolise l'information physique disponible pour effectuer la tâche tandis que R est la représentation interne à laquelle renvoie l'information A . B) et C) Illustration du fonctionnement de la "reverse correlation" (panneau B) et de la méthode Bubbles (panneau C).

Reprenons cette formulation dans le cadre de notre questionnaire sur la perception du sexe des visages. Dans ce cas, A représente l'ensemble des informations dimorphiques sexuelles disponibles. Cette information est habituellement estimée à l'aide de techniques statistiques que l'on applique sur des mesures physiques telles que la luminance et la chrominance. Nous verrons d'ici peu certaines méthodes d'estimation de A . La représentation R quant à elle, est un concept avant tout psychologique. Elle symbolise l'ensemble des connaissances perceptives et cognitives d'un observateur humain ayant trait au sexe d'un visage — ou à tout autre concept. Dans le contexte de $R \otimes A \approx P$, R inclurait toutes les informations regroupées dans A si l'humain n'était limité par aucune contrainte sensorielle, perceptive ou cognitive et donc que sa connaissance du sexe des visages était elle aussi complète. En réalité, les systèmes visuel et cognitif sont limités, que ce soit par des processus stochastiques comme le bruit neuronal ou des processus

déterministes de bas niveau tels que la sensibilité au contraste, ou de plus haut niveau tels que la mémoire et l'attention. Ainsi, en réalité, R ne contiendra vraisemblablement pas toutes les informations détenues dans A . Cependant, dans la mesure où l'observateur humain est adapté à son environnement visuel, R devrait se rapprocher de A . C'est précisément ce rapprochement, c.-à-d. le chevauchement entre A et R , qui est symbolisé par P . En fait, P est la portion des dimorphismes sexuels visibles qui est incluse dans la représentation de l'observateur humain. C'est principalement cette information qui sera étudiée dans la présente thèse. Voyons maintenant les méthodologies permettant d'accéder à chacune de ces informations.

L'information disponible A est, comme nous l'avons suggéré, accessible par des techniques statistiques. Prenons l'exemple d'un dimorphisme simple, soit la taille du visage (sur la verticale). Supposons que la tâche soumise à un observateur était de catégoriser le sexe d'un ensemble de 100 visages d'hommes et de femmes ne se distinguant que par leur taille — tout autre aspect du visage étant sexuellement neutre. Dans une telle situation, il suffirait, pour estimer A , de mesurer la taille de chacun des visages et de comparer la taille des visages d'hommes et de femmes à l'aide d'une statistique tel un score Z ou une distance de Cohen—advenant que la distribution des tailles soit normale. Évidemment, la plupart des situations de catégorisation visuelle qui se présentent dans la vie réelle ainsi que celles qui sont testées expérimentalement comportent un niveau de complexité plus élevé. Ceci implique la présence simultanée de plusieurs informations sexuelles, ce qui évoque la question de l'intégration des diverses informations. Dans pareilles situations, l'estimation statistique de A requerrait donc une technique comme la régression linéaire multiple qui permet d'intégrer optimalement plusieurs sources de variation. Dans les conditions expérimentales complexes où l'emploi des techniques statistiques standards s'avère trop complexe ou impossible, les psychophysiciens ont habituellement recours à l'analyse de l'*observateur idéal* — voir Tjan et al., (1995), Sekuler et al. (2004) et Schyns, Gosselin et Smith (2009) pour quelques exemples d'application de l'analyse. Cette analyse consiste à soumettre un observateur idéal à la même tâche et aux mêmes stimuli que les participants humains. Cet observateur n'est contraint par aucune limite perceptive, attentionnelle ou mnémonique : il connaît donc tous les visages de l'expérience et accomplit la tâche optimalement. Formellement, un tel observateur est un modèle mathématique maximisant la probabilité de classer correctement les stimuli étant donné les informations qu'ils

contiennent. Une formulation complète d'un tel observateur est présentée dans l'Article 2 (voir section 3.4.1) et dans la discussion (voir section 5.3). En conclusion, il appert que l'estimation de A est limitée par l'échantillon utilisé, l'outil de mesure utilisé et les contraintes du modèle mathématique employé (p. ex. la linéarité).

Si A se réfère à des mesures physiques et peut être estimé mathématiquement, qu'en est-il de R et de P ? R est un concept d'abord psychologique qui renvoie à un construit théorique qui n'est pas directement observable. Certaines méthodes permettent toutefois d'y avoir partiellement accès en le projetant dans le même espace multidimensionnel que A . Une de ces méthodes est la "reverse correlation" (voir Ahumada & Lovell, 1971; Gosselin & Schyns, 2002; Gosselin & Schyns, 2003; Gosselin & Schyns, 2004; Mangini & Biederman, 2004; Gold et al., 2000). Prenons l'exemple d'une expérience sur la catégorisation du sexe. D'abord, des stimuli neutres, c.-à-d. des visages androgynes, sont présentés aux participants et on leur demande de les classer selon leur sexe. Sur ces images sont additionnées des plages de bruit visuel gaussien. Étant donné que les stimuli sont androgynes, il n'y a aucune bonne ou mauvaise réponse, et donc $A = 0$. En fait, le seul élément changeant d'un essai à l'autre est le bruit visuel. Inévitablement, si l'intensité du bruit visuel est suffisante, le visage androgyne bruité semblera plus masculin dans certains essais, et plus féminin dans d'autres. Étant donné que ces différences de perception ne seraient imputables qu'à la présence du bruit et la représentation de l'observateur, il suffirait d'analyser les plages de bruit ayant mené aux réponses 'homme' et de les comparer à celles ayant mené aux réponses 'femme' pour révéler les éléments masculins ou féminins de cette représentation. La "reverse correlation" vise justement à saisir l'impact du bruit sur les réponses des participants. Cette technique fut appliquée avec succès afin de révéler les propriétés des représentations auditives (p. ex. Ahumada & Lovell, 1971) ou visuelles (p. ex. Gosselin & Schyns, 2003; Gold et al., 2000; Mangini & Biederman, 2004) ainsi que les propriétés du champ récepteur de neurones (p. ex. Ringach & Shapley, 2004; DeAngelis, Ohzawa & Freeman, 1993).

Voyons maintenant à quoi ressembleraient les résultats d'une telle analyse dans une tâche plus simple. Supposons que l'on demande à un observateur de détecter la lettre 'A' camouflée dans du bruit visuel additif gaussien. Des exemples de tels stimuli sont présentés à la Figure 3B (en haut à gauche). Compte tenu qu'il soit fort peu probable que les représentations internes ne

soient en parfaite adéquation avec l'information disponible, on pourrait postuler que $R \neq A$. Supposons, pour illustrer cette inégalité, que l'observateur se représente la lettre 'A' plutôt que la lettre 'R' (voir Figure 3A). Simplifions encore d'avantage l'analyse de la "reverse correlation", et disons qu'elle revient à soustraire les plages de bruit ayant mené aux réponses 'non' (voir Figure 3B en bas au centre) aux plages de bruit ayant mené aux réponses 'oui' (voir Figure 3B en bas à gauche). La résultante est une image de classification qui illustre à quel point chaque pixel de l'image est corrélé avec les réponses 'oui'. Donc plus le pixel de cette image est clair, plus il est corrélé avec les réponses positives (voir Figure 3B en bas à droite). Notez que seules les plages de bruit sont utilisées dans cette analyse et donc que la structure contenue dans l'image de classification n'est explicable que par les réponses faites sur la base de la représentation R de l'observateur. Nous verrons à la section 1.2.1, un exemple d'application de cette méthode dans le cadre de la catégorisation du genre des visages. Les limites de la méthode y seront aussi abordées. Notons finalement que R pourrait contenir des informations qui ne sont pas incluses dans A , mais qui pourraient être estimées par la "reverse correlation". Ceci renvoie, entre autres choses, au concept des stimuli supranormaux et à l'idée popularisée par le biologiste Niko Tinbergen selon laquelle la représentation innée pourrait surpasser quelquefois les variations observables dans la nature (p. ex. ten Cate et al. 2009). Un exemple d'une telle représentation dans le contexte de la catégorie faciale du sexe sera présenté à la section 1.2.1. Ainsi, en plus des contraintes liées à l'emploi de la "reverse correlation", il appert que la projection de R dans l'espace A est nécessairement une opération réductionniste.

Comme ce fut mentionné précédemment, P est l'intersection entre R et A . Plus spécifiquement, il s'agit de l'ensemble des informations contenues dans A qui sont également contenues dans R . En ce sens, l'information P suppose déjà la projection de R dans l'espace A . La méthode Bubbles (voir Gosselin & Schyns, 2001; Gosselin & Schyns, 2004), une autre technique de classification d'images, permet d'estimer cette information. Prenons encore l'exemple d'une tâche de catégorisation du sexe des visages : des visages bruités d'hommes et de femmes sont présentés aux participants et ceux-ci doivent en déterminer le sexe. À la différence de la "reverse correlation", le bruit utilisé dans une telle expérience sera multiplicatif et non additif. La méthode Bubbles suppose plus spécifiquement l'utilisation d'un bruit multiplicatif binomial convolué par un filtre gaussien (c.-à-d. une *bulle*). Ce bruit aura pour effet de révéler certaines régions du

stimulus et d'en masquer d'autres, et ce, de façon totalement aléatoire. Ainsi, dans certains essais le bruit révélera certaines régions utiles à la tâche de catégorisation, tandis que dans d'autres essais aucune région utile ne sera révélée ce qui mènera vraisemblablement à des réponses moins exactes. Encore une fois, le bruit sera révélateur des représentations de l'observateur. En fait, la comparaison des plages de bruit menant à des réponses correctes avec les plages de bruit menant à des réponses incorrectes permettra d'identifier les régions qui, une fois dévoilées, mènent à des réponses plus justes.

Pour des fins de démonstration, simplifions maintenant l'analyse et considérons-la encore dans le cadre d'une tâche d'identification de la lettre 'A'. Supposons aussi que la représentation de l'observateur testé, une image de la lettre 'R', est imparfaite (voir Figure 3). En simplifiant à outrance, on pourrait dire que l'analyse typique dans une tâche utilisant la méthode Bubbles consiste à soustraire les plages de bruit ayant mené aux réponses incorrectes (voir Figure 3C en bas au centre) aux plages de bruit ayant mené aux réponses correctes (voir Figure 3C en bas à gauche). La résultante est une image de classification qui illustre à quel point chaque pixel de l'image est corrélée avec les réponses correctes. Donc plus le pixel de cette image est clair, plus il est corrélé avec les réponses justes (voir Figure 3C en bas à droite). Ainsi, l'image de classification résultant d'une expérience Bubbles exhibera la portion des informations *A* incluse dans la représentation interne *R*, c.-à-d. les informations *puissantes*. Noter que la technique Bubbles fut utilisée avec succès dans des domaines divers allant de la reconnaissance faciale (p. ex. Vinette, Gosselin & Schyns, 2004; Blais et al., 2012) à la lecture (p. ex. Fiset et al., 2009; Tadros et al., 2013), en passant par l'électrophysiologie (p. ex. Schyns et al., 2009; Schyns et al. 2011; Smith, Gosselin & Schyns, 2006) et la neuropsychologie clinique (p. ex. Gosselin et al., 2011; Adolphs et al., 2005). Cette méthode fut utilisée dans les articles 1 et 3 de cette thèse et sera donc présentée et discutée en plus amples détails aux sections 1.2.1, 2.3 et 4.3.

1.1.4. Présentation générale de la thèse

La question principale, à savoir *lesquelles des informations faciales sexuelles sont utilisées par l'observateur humain afin de qualifier le sexe d'un visage*, a d'abord servi de point de départ au premier article de cette thèse (voir Chapitre 2). De nouvelles questions, cette fois plus

précises, quant au rôle et à l'importance relative des indices achromatiques, chromatiques et configuraux ont par la suite découlé de ce premier article. Ces questions ont fait l'objet du deuxième article (voir Chapitre 3). Suite aux résultats obtenus dans les deux premiers articles, nous nous sommes questionnés sur les mécanismes perceptifs d'extraction des différents indices chromatiques et achromatiques. Le troisième article (voir Chapitre 4) visait donc à analyser les dynamiques spatiotemporelles d'extraction de ces indices visuels. Le contenu des trois articles formant le corpus de cette thèse progresse donc du général au particulier : on y examine d'abord le rôle fonctionnel de certains indices visuels, pour finalement conclure avec une hypothèse précise quant au mécanisme cérébral d'encodage de ces indices.

L'introduction des thématiques associées aux deux premiers articles de la thèse procèdera en suivant le cadre théorique $R \otimes A \approx P$, c.-à-d. que nous présenterons d'abord des analyses estimant les informations A pour la tâche de catégorisation faciale du sexe (voir sections 1.2.1.1 et 1.2.2.1); ensuite, nous recenserons, discuterons et critiquerons les études tentant d'évaluer les informations P et R (voir sections 1.2.1.2 et 1.2.2.2).

1.2. Dimorphismes sexuels faciaux et leur importance perceptive

Comme nous l'avons vu précédemment, les dimorphismes sexuels faciaux regroupent tous les signes morphologiques du visage présentant une différence entre les hommes et les femmes. Ces dimorphismes sont accessibles à l'observateur humain dans la mesure où ils se répercutent sur le spectre visible de la lumière réfléchiée par un visage. Nous parlerons d'indices visuels pour désigner ces signaux optiques visibles émanant des dimorphismes sexuels. Par exemple, les ombrages dans la région nasale seraient des indices achromatiques rattachés à la protrusion tridimensionnelle du nez.

Étant donné leur grand nombre, nous ne répertorierons pas l'ensemble des dimorphismes et des indices visuels qui s'y rattachent. Nous nous concentrerons plutôt sur ceux abordés dans la littérature sur la catégorisation du genre des visages. Notons également que pour être cohérents avec les trois articles de la thèse, tous les indices décrits et analysés dans les prochaines sections

se baseront, à moins d'avis contraire, sur des visages caucasiens adultes inexpressifs vus frontalement.

1.2.1. Indices chromatiques et achromatiques

Théoriquement, les indices achromatiques et chromatiques locaux regroupent tous les dimorphismes qui ont une incidence sur la lumière réfléchi par une constituante locale du visage. Ils regroupent principalement les dimorphismes de pigmentation, de texture et de forme. Notez qu'il sera question de réflectance relative plutôt que de réflectance absolue puisque les visages sont perçus dans des conditions d'illumination très variables et que le système visuel présente des mécanismes d'adaptation aux changements d'illumination tels que la constance des couleurs et de la clarté (Palmer, 1999).

Chez les Caucasiens, les dimorphismes de pigmentation renvoient au spectre de la lumière absorbée par différents pigments de l'épiderme, des yeux, des poils faciaux et des lèvres. En ce sens, il y a un certain recoupement entre ces dimorphismes et les dimorphismes de forme. Un exemple de dimorphisme pigmentaire concerne l'épiderme féminin. Chez la femme, l'épiderme est généralement plus pâle, possiblement pour optimiser, par la synthèse de la vitamine D, l'absorption du calcium nécessaire à la grossesse et la lactation (Jablonski, 2000; Frost, 1988; Russell, 2009). Les dimorphismes texturaux concernent la porosité de la peau, les rides, les taches et autres imperfections cutanées. Ceux-ci sont tributaires de plusieurs facteurs, dont l'âge et l'alimentation. Les dimorphismes de forme renvoient d'une part à la largeur et la hauteur des différentes constituantes du visage comme la bouche, le nez, les yeux et les sourcils, mais aussi aux ombrages provenant des structures tridimensionnelles du visage qui font obstacle à la lumière incidente. Le nez et les arcades sourcilières en sont des exemples.

1.2.1.1. Dispersion et importance statistique des indices chromatiques et achromatiques

Afin de mieux saisir la dispersion moyenne des indices achromatiques et chromatiques sur le visage ainsi que leur importance statistique respective, considérez la Figure 4A. Cette figure fait le sommaire de la distribution spatiale des informations sexuelles disponibles dans le canal L

(ou clair-sombre), a (ou rouge-vert) et b (ou jaune-bleu), soit chacun des canaux de l'espace des couleurs antagonistes CIE Lab. Le canal L englobe les indices achromatiques tandis que les canaux A et B regroupent les indices chromatiques. L'analyse statistique dont les résultats sont présentés dans cette figure consiste à évaluer la différence statistique (exprimée en d') entre la distribution des valeurs d'un pixel chez 150 visages de femmes et la distribution des valeurs de ce même pixel chez 150 visages d'hommes. Cette comparaison a été accomplie pour chaque canal de l'espace CIE Lab et pour chaque pixel des images. Une telle comparaison a d'abord nécessité l'alignement préalable de ces 300 visages sur 12 repères faciaux. Les détails concernant ces visages, leur alignement et cette analyse statistique sont fournis à la section 2.4.1. Ainsi, chaque courbe de cette figure détient un code de couleur qui exprime à quel point les valeurs de luminance relative contenues dans un pixel distinguent les visages d'hommes et de femmes. Pour chaque canal (chaque rangée), il y a six courbes d'isovaleurs qui délimitent les régions dont les pixels correspondent au 95^e, 85^e, 75^e, 25^e, 15^e et 5^e percentile des d' . Plus la valeur du d' s'éloigne de zéro plus le pixel sous-jacent est diagnostique du sexe des visages. Pour des fins d'interprétation, les courbes sont juxtaposées à un visage androgyne. Également, une légende jointe à l'échelle de couleur permet d'interpréter ces maps statistiques. Par exemple, les d' calculés dans la région de la bouche pour le canal rouge-vert (2^e rangée) se situe au 5^e percentile, ce qui indique que cette région est plus verte chez l'homme que chez la femme (ou inversement plus rouge chez la femme que chez l'homme). Globalement, cette figure révèle que les plus fortes variations se situent au niveau des canaux L (ou clair-foncé) et A (rouge-vert). Les principales différences achromatiques se localisent dans les régions orales, oculaires et sur le pourtour du visage tandis que les variations chromatiques (canaux A et B) se concentrent davantage dans la région de l'os maxillaire et dans les zones orale et oculaires.

À des fins de comparaison, considérez les résultats d'une analyse similaire réalisée par Nestor et Tarr (2008a) sur 200 visages Caucasiens (50% femmes) alignés en plan frontal et inexpressifs (Figure 4B). Les images de gauche présentent la différence entre la moyenne (point par point) des visages de femmes et la moyenne des visages d'hommes tandis que les images de droite présentent les pixels comportant des différences significatives (zones magenta et cyan) entre les hommes et les femmes selon de multiples tests-t bilatéraux. Une correction de Bonferroni a été appliquée ($p < 0.05$). Les zones magenta et cyan désignent des valeurs plus

élevées chez la femme et l'homme respectivement. Bien que l'analyse de Nestor et Tarr (2008) ne soit pas directement comparable à celle que nous avons menée étant donné l'utilisation de tests statistiques, elle confirme globalement nos résultats. Entre autres, elle révèle des différences plus marquées dans les canaux clair-foncé (L) et rouge-vert (A). Cependant, elle ne présente aucune différence significative dans le canal rouge-vert dans la région labiale.

Ces évaluations statistiques comportent plusieurs limites. Premièrement, elles ne tiennent pas compte de la covariance entre les pixels. Deuxièmement, elles se limitent à des échantillons de visages caucasiens inexpressifs ayant tous été alignés sur un plan frontal. Notez qu'un tel échantillon est par contre assez représentatif des échantillons utilisés dans les études évaluant l'importance perceptive de ces indices. Troisièmement, ces analyses n'évaluent pas l'importance d'autres indices achromatiques et chromatiques tels que les contrastes locaux, c.-à-d. des différences de luminance entre différentes zones contigües du visage. Par exemple, Russell (2003, 2009) a déjà montré que les visages de femmes présentaient un contraste plus prononcé entre les yeux, la bouche et la peau environnante que les visages d'hommes, et que les jugements de genre et de beauté interagissaient avec ces indices.

Nous avons présenté ces maps statistiques dans le but d'évaluer l'importance et la distribution des indices achromatiques et chromatiques dans le visage. Il s'agit maintenant de se questionner sur ce qu'elles signifient sur le plan perceptif. Compte tenu des limites évoquées ci-dessus, nous dirons d'abord que ces maps se veulent une approximation des informations achromatiques et chromatiques disponibles (c.-à-d. les informations A) à l'observateur humain dans un contexte restreint. Étant donné les limites cognitives et perceptives de l'observateur humain, celui-ci ne pourra épuiser l'ensemble des informations dimorphiques disponibles, misant plutôt sur l'utilisation efficace d'un sous-ensemble de celles-ci (ou sur l'utilisation moins efficace de toutes ces informations). Le besoin d'extraire des régularités dans un environnement complexe et dynamique pourrait constituer une autre raison l'incitant à sous-utiliser les informations disponibles. En effet, les visages que l'observateur humain perçoit varient en fonction de l'expression, de l'âge, de l'ethnie et se présentent sous des conditions d'illumination très variables et selon des points de vue changeants. Dans un tel environnement, une stratégie pourrait être de miser sur des indices qui présentent une certaine régularité à travers les différents

contextes, ainsi que sur les indices servant à d'autres fins (p. ex. d'autres jugements faciaux). Dans un tel contexte, on pourrait se demander quels indices achromatiques et chromatiques visuels seront privilégiés par l'observateur humain et quelles zones faciales contenant ces indices seront utilisées. La prochaine section examinera les études portant sur ces questions.

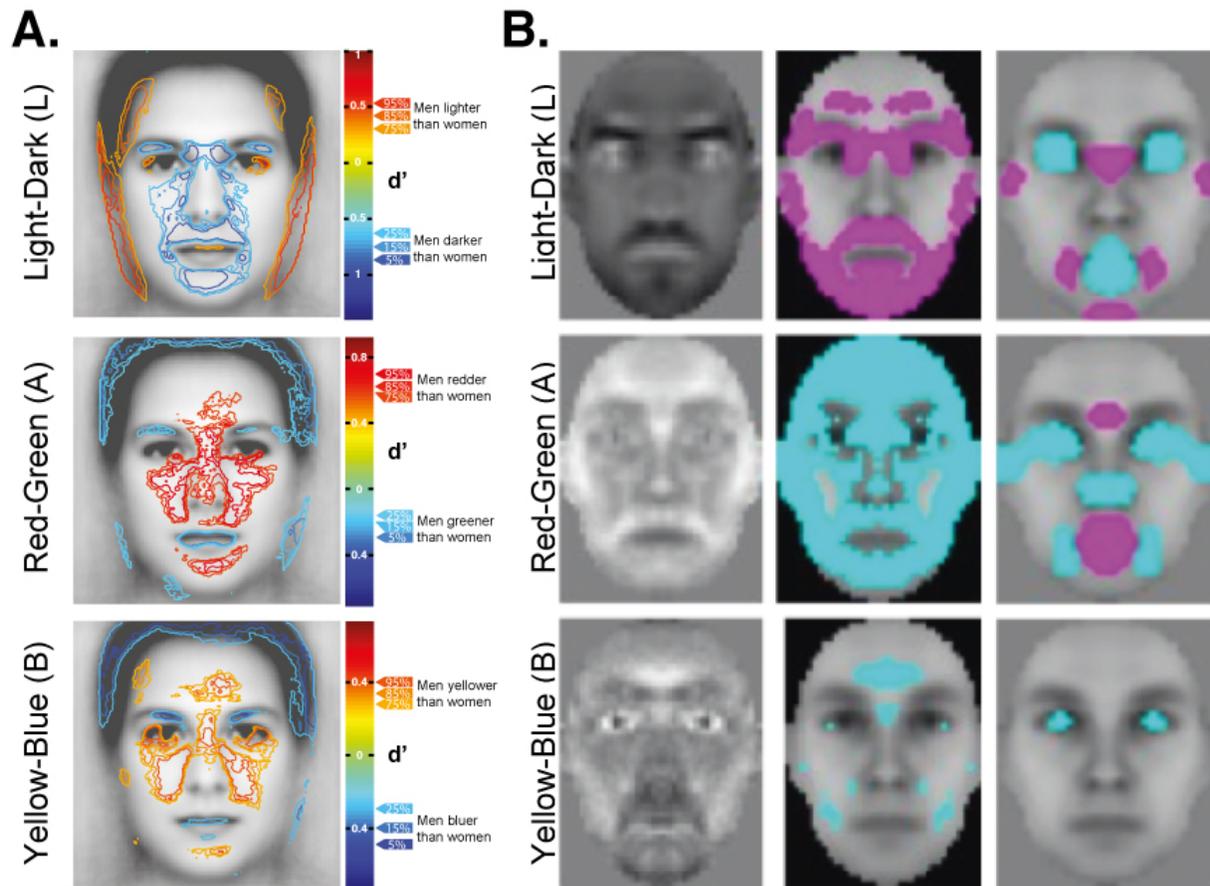


Figure 4. A) Courbes de contour faisant le sommaire de la distribution spatiale des informations sexuelles disponibles (exprimées en d') dans chaque pixel de l'image et dans le canal L (ou light-dark), A (ou Red-Green) et B (ou Yellow-Blue). B) Colonne de gauche : Sommaire des informations sexuelles (exprimées en scores Z) dans chaque pixel de l'image. Colonne du centre : Pixels de la colonne de gauche significatifs ($p < 0.05$). Les zones magenta et cyan désignent des valeurs plus élevées chez la femme et l'homme respectivement. Colonne de droite : Pixels significativement et positivement corrélés avec les réponses 'homme' (cyan) et 'femme' (magenta) des participants (tiré de Nestor & Tarr, 2008b).

1.2.1.2. Contraste entre les principaux attributs faciaux et la peau

Les analyses précédentes évaluent les indices achromatiques et chromatiques au niveau de chaque pixel d'une image. Certains chercheurs ont proposé que le contraste entre les lèvres et la peau avoisinante ainsi que le contraste local entre les yeux et la peau avoisinante contenaient également des informations sexuelles achromatiques (Russell, 2003, 2009). Plus spécifiquement, Russell (2009) a comparé la réflectance moyenne des yeux et des lèvres à la réflectance moyenne de la peau avoisinant ces attributs (voir Figure 5A, image de droite), et ce, sur 118 visages d'hommes et de femmes asiatiques (N = 51) et caucasiens (N = 67). Les résultats présentés à la Figure 5A ainsi que des comparaisons statistiques montrent que la peau entourant la bouche et les yeux est statistiquement plus lumineuse chez les femmes que chez les hommes, et ce, autant chez les visages caucasiens qu'asiatiques. En d'autres termes, le contraste local entre les yeux et la peau environnante et le contraste local entre la bouche et la peau environnante sont significativement plus élevés sur les visages de femmes.

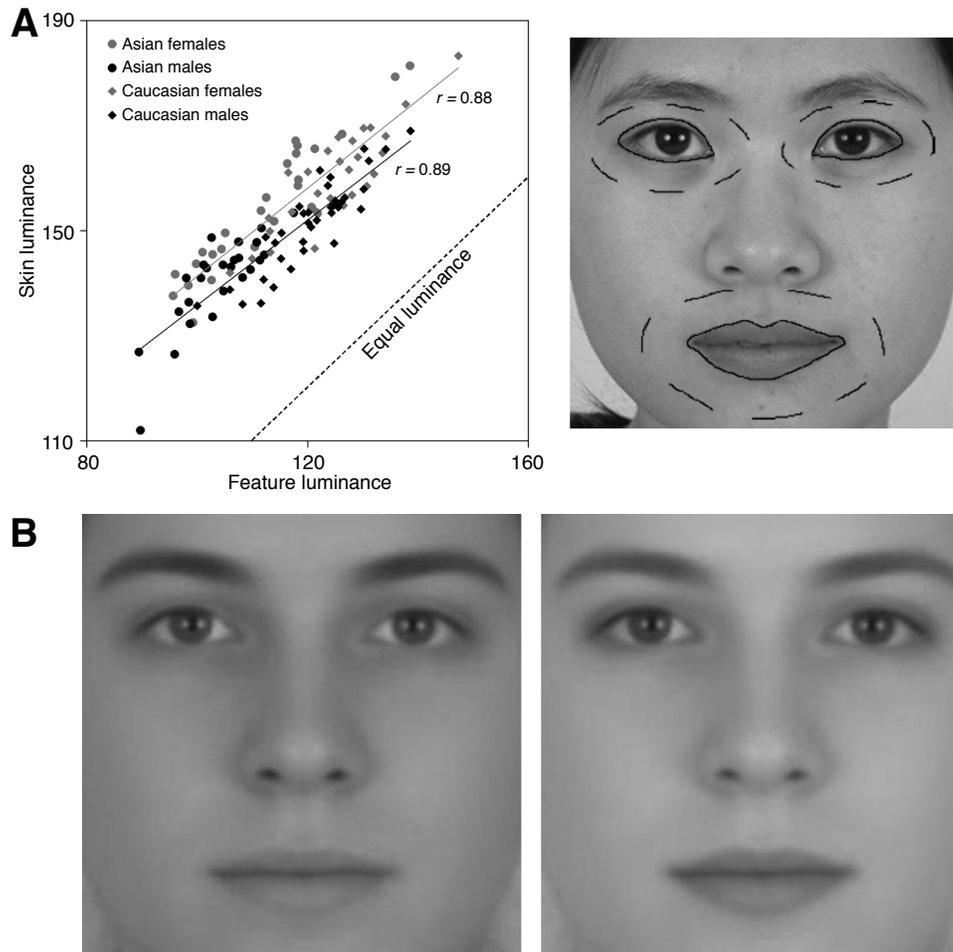


Figure 5. A) Luminance prélevée sur 118 visages d’hommes et de femmes asiatiques (N = 51) et caucasiens (N = 67). Le graphique de gauche montre la luminance moyenne des yeux et des lèvres (abscisse) pour chacun des visages ainsi que la luminance moyenne de la peau avoisinant ces attributs (ordonnée). L’image de gauche indique les régions dans lesquelles ces mesures furent prélevées (images tirées de Russell, 2009). B) Illusion du Sexe tirée de Russell (2009). Il s’agit de deux visages androgynes dont la luminance des yeux et de la bouche est identique. La luminance de la peau est plus élevée dans le visage de droite que celui de gauche.

Afin de connaître l’importance perceptive des contrastes locaux achromatiques mesurés sur ces 118 visages, Russell (2009) a demandé à 29 participants de juger le degré de masculinité des visages d’hommes et le degré de féminité des visages de femmes. Les résultats montrent une corrélation positive significative entre l’intensité des contrastes locaux et les jugements de

féminité de ces mêmes visages de femmes, et une corrélation négative significative entre l'insensibilité des contrastes locaux et les jugements de masculinité des visages d'hommes. Ces résultats sont aussi cohérents à ceux de Russell (2003) qui avait observé que l'augmentation du contraste local entre les yeux et tout le reste de la peau du visage ainsi que l'augmentation du contraste local entre la bouche et tout le reste de la peau du visage augmentaient le degré de féminité des visages de femmes et diminuaient le degré de masculinité des visages d'hommes.

Bien que les participants de Russell (2003, 2009) n'aient pas catégorisé le sexe des visages, l'illusion du Sexe qu'il a créée en 2009 (voir Figure 5B) suggère tout de même que ces contrastes locaux jouent un rôle sur le plan de la catégorisation sexuelle. Cette illusion comporte deux visages androgynes dont les yeux et la bouche sont parfaitement identiques sur le plan de la luminance. Pourtant, le visage de droite semble plus masculin que celui de gauche. En fait, la seule différence entre ces deux visages est que la luminance de la peau du reste du visage (i.e. excluant les yeux et la bouche) est plus élevée dans celui de droite que celui de gauche, ce qui a pour effet d'augmenter (visage de droite) ou de diminuer (visage de gauche) le contraste local entre les attributs (i.e. yeux et bouche) et le reste du visage.

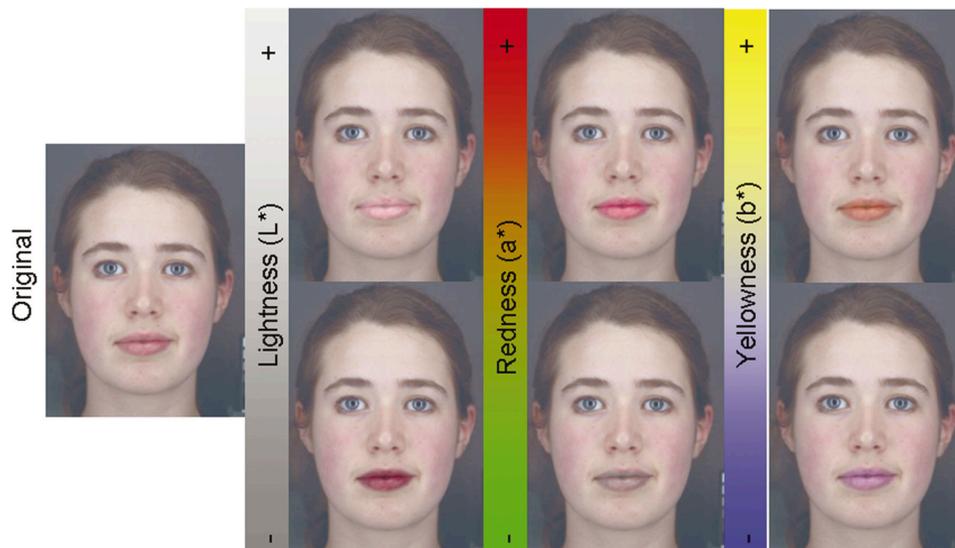


Figure 6. Transformation de contrastes locaux d'un visage de femme suivant les dimension de l'espace des couleurs CIE Lab (tiré de Stephen et McKeegan, 2010).

Qu'en est-il de l'importance des contrastes locaux chromatiques? Stephen et McKeegan (2010) ont testé l'effet de la manipulation du contraste local entre les lèvres et le reste du visage sur l'attractivité et la masculinité/féminité des visages d'hommes et de femmes. Pour ce faire, ils ont demandé à 48 participants caucasiens d'ajuster indépendamment l'intensité du canal de luminance (L^*), du canal rouge-vert (a^*) et du canal jaune-bleu (b^*) des lèvres de 48 visages de femmes et d'hommes de sorte à maximiser leur attractivité et leur féminité/masculinité respectivement (Figure 6). Chacun des canaux de l'espace CIE Lab était manipulé dans des conditions indépendantes. Les résultats ont montré que l'augmentation du contraste local des lèvres dans l'axe rouge-vert et la réduction du contraste local des lèvres dans l'axe b^* maximisaient l'attractivité des visages d'hommes et de femmes. De plus, cette augmentation (ou réduction) était plus prononcée pour les visages de femmes que pour les visages d'hommes. Également, la masculinité des visages d'hommes et la féminité des visages de femmes étaient liées à une diminution du contraste local dans l'axe bleu-jaune et une légère augmentation du contraste local dans l'axe L^* . Sur l'axe rouge-vert, la masculinité des visages d'hommes était lié à une diminution du contraste local au niveau labial tandis que la féminité des visages était liée à une augmentation de ce même contraste. Il est à noter qu'un seul canal de l'espace CIE Lab était ajusté à la fois, et que la valeur des deux autres canaux étaient maintenus constant à leur valeur moyenne respective. Dans ce contexte, les valeurs de contraste obtenues pour chacun des canaux doivent être interprétées en interaction avec la moyenne des deux autres canaux. De plus, aucune précaution ne semble avoir été prise afin que l'étendue des valeurs de contraste ajustées respecte celles observées sur de vrais visages de femmes et d'hommes Caucasiennes. Les valeurs de contraste enregistrées sont davantage en lien avec les représentations R que l'information puissante P .

1.2.1.3. Examen des études sur l'utilisation des indices achromatiques et chromatiques

Afin d'examiner l'importance perceptive des indices achromatiques et chromatiques ainsi que leur utilisation locale, les chercheurs ont eu recours à diverses méthodologies, dont le masquage (Roberts et Bruce, 1988; Bruce et al., 1993), l'isolement (Roberts et Bruce, 1988; Brown et Perret, 1993; Hill, Bruce & Akamatsu, 1995), la substitution (Brown et Perret, 1993 ; Yamaguchi, Hirukawa et Kanazawa, 1995; Hill, Bruce & Akamatsu, 1995), l'oculométrie

(Sæther et al., 2009) et les techniques de classification d'images (p. ex. Netor & Tarr, 2008b; Gosselin & Schyns, 2001). Le tableau 1 résume les principaux résultats de ces études. On y décrit le type de visages utilisés (colonne 2), la méthodologie utilisée (colonne 3), les zones faciales testées (colonne 4), les mesures prélevées (colonne 4) et les zones faciales en ordre descendant d'effet expérimental ou simplement celles qui ont été démontrées comme étant significatives (colonne 5). Notons que pour les études ayant utilisé des visages expressifs et/ou présentés selon d'autres angles (p. ex., profile 3/4), seuls les résultats ayant trait à la condition inexpressive et en plan frontal seront rapportés. Des exemples de visages ayant subi la méthode de l'isolement, du masquage, de la substitution d'attributs et de la technique de classification d'images sont présentés à la Figure 7.

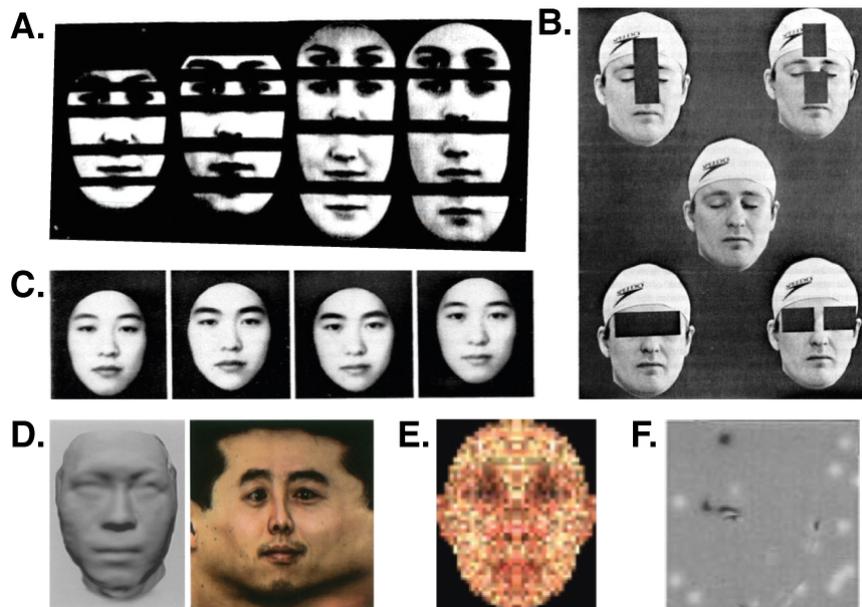


Figure 7. A) Stimuli employés avec la méthode de l'isolement (tiré de Brown & Perrett, 1993). B) Stimuli employés avec la méthode du masquage (tiré de Bruce et al., 1993). C) Visages prototypiques employés avec la méthode de la substitution d'attributs (tiré de Yamaguchi, Hirukawa & Kanazawa, 1995). De la gauche vers la droite sont présentés: un homme avec des sourcils de femme, une femme avec un contour de visage d'homme, une femme avec des sourcils d'homme et un hyper-mâle avec des sourcils de femme. D) Visage dont les indices d'ombrage (gauche) et les indices de couleur (droite) ont été isolés (tiré de Hill, Bruce & Akamatsu, 1995). E) Stimuli utilisés suivant la technique de "reverse correlation" (tiré de

Nestor & Tarr, 2008b). F) Visages échantillonnés à l'aide de la technique Bubbles (tiré de Gosselin & Schyns, 2001).

Tableau 1. Sommaire des résultats des études sur l'importance des zones et indices faciaux pour la catégorisation du sexe des visages

Étude	Type de visages	Méthode	Zones anatomiques testées	Mesure	Résultats significatifs (*) ou en ordre descendant d'effet expérimental (numéros)
Roberts et Bruce, 1988	- 64 visages (32 Caucasiens non familiers) - Tons de gris	Masquage d'attribut(s)	- Yeux - Nez+sourcil central - Bouche - Rien	Différence de temps de réponse entre la condition où rien n'est masqué et les autres conditions	1- Nez+sourcil central (311 ms) 2- Bouche (77 ms) 3- Yeux (43 ms)
		Isolement d'attribut(s)	- Yeux - Nez - Bouche	Justesse de réponse	1- Yeux (78%) 2- Nez (59%) 3- Bouche (67%)
Brown et Perret, 1993	- Un prototype mâle et un prototype femelle - Caucasiens - Alignés sur la distance interpupillaire - Tons de gris - Inexpressifs	Isolement d'attribut(s)	- Sourcils - Yeux - Nez - Bouche - Mâchoire - Menton - Sourcils+yeux - Yeux+nez - Nez+bouche - Bouche+mâchoire	Index de justesse de classification	1- Sourcils+yeux 2- Sourcils 3- Yeux 4- Mâchoire 5- Menton 6- Nez+bouche 7- Bouche
		Substitution d'attribut(s)	- Sourcils - Yeux - Nez - Bouche - Mâchoire - Sourcils+Yeux - Yeux+nez - Nez+bouche - Bouche+mâchoire	Pourcentage de classification incorrecte des prototypes ayant des attributs du sexe opposé	1- Mâchoire 2- Sourcils+yeux 3- Menton 4- Sourcils
Yamaguchi, Hirukawa et Kanazawa, 1995	- Visages japonais prototypiques - Alignés - Inexpressifs - Plan frontal - Tons de gris	Substitution d'attribut(s)	- Sourcils - Yeux - Nez - Bouche - Contour du visage - 10 paires possibles de ces attributs	Changement significatif de la justesse de réponse	* Sourcils * Contour * Toutes paires incluant ces deux attributs
Bruce, Burton, Hanna, Healey, Mason, Coombes, Fright et Linney, 1993	- 32 visages (16 Caucasiens) - Non alignés - Inexpressifs - Plan frontal et autres - Tons de gris	Masquage d'attribut(s)	- Yeux+sourcils+pont nasal - Yeux+sourcils - Nez+milieu du front+yeux fermés - Nez+sourcil central+milieu du front+yeux fermés	Justesse de réponse	1- Yeux+sourcils (76.6%) 2- Yeux+sourcils+pont nasal (79.4%) 3- Nez+milieu du front+yeux fermés (82.6%) 4- Nez+sourcil central+milieu du front+yeux fermés (84.7%)
Hill, Bruce et Akamatsu, 1995	- 16 visages (8 Caucasiens) - Non alignés - Plan frontal et autres - Tons de gris et couleurs - Inexpressifs	Isolement d'indices	Tous les pixels de l'image, mais des indices de surface isolés (texture, pigment, forme 2D) et des indices d'ombrage résultant de la forme 3D	Différence de justesse de réponse entre la tâche de discrimination de la race et la tâche de discrimination du genre	* Indices d'ombrage plus importants pour la race (88%) que le sexe (72%) * Indices de surface plus importants pour le sexe (97%) que la race (90%)
Nestor et	- Un androgyne	Technique de la	Tous les pixels de la	Pixels	(voir Figure 4)

Tarr, 2008b	caucasien - Inexpressif - Plan frontal - Tons de gris	classification d'images (Bruit additif Gaussien dans l'espace des couleurs CIE Lab).	moitié latérale gauche ou droite du visage.	significativement corrélés avec la réponse 'homme' ou la réponse 'femme'	* Yeux plus pâles, plus rouges et plus jaunes chez les hommes. * Bouche plus claire chez les hommes et plus rouge chez les femmes * Pont nasal plus clair et plus rouge chez les femmes * Base du nez plus rouge chez les hommes * Contour de la bouche plus clair chez les femmes et plus rouge chez les hommes.
Gosselin et Schyns, 2001	- 32 visages caucasiens - Inexpressifs - Plan frontal - Alignés à la main - Tons de gris	Technique de la classification d'images (Bruit binomial multiplicatif filtré par un passe-bas) - Échantillonnage de l'axe <i>xy</i>	Tous les pixels de l'image	Pixels significativement corrélés avec la classification correcte du sexe	(voir Figure 8B) - Zone orbitale (yeux+sourcils) - Zone orale
Schyns, Bonnar et Gosselin, 2002	- 10 visages caucasiens - Inexpressifs - Plan frontal - Alignés à la main - Tons de gris	Technique de la classification d'images (Bruit binomial multiplicatif filtré pas un passe-bas) - Échantillonnage de l'axe <i>xy</i> et des fréquences spatiales	Tous les pixels de l'image	Pixels et fréquences spatiales significativement corrélés avec la classification correcte du sexe	(voir Figure 8C) - Oeil à la gauche de l'image en haute fréquence spatiale (FS) - Zones orbitale, nasale et orale en moyenne haute FS - Contour du visage, zones orbitale, nasale et orale en moyenne basse FS
Sæther, Van Belle, Laeng, Brennen et Øvervoll, 2009	- 24 visages caucasiens en couleurs - Inexpressifs - Plan frontal et autres - Non alignés	Oculométrie	- Cheveux - Front - Sourcils - Yeux - Nez - Joues - Bouche - Menton/mâchoire - Oreilles	Fréquence significative des fixations par région	* Nez * Yeux * Région des joues
			Tous les pixels de l'image	Densité de probabilité associée au temps passé sur chaque pixel (seules les régions présentant une probabilité supérieure à 95% sont présentées)	(voir Figure 8A) - Zone infraorbitale - Pont nasal

Afin de suivre l'ordre de présentation des articles de thèse, nous examinerons d'abord les études portant sur l'utilisation des différentes zones faciales et poursuivrons avec les études testant l'importance relative des indices achromatiques et chromatiques.

1.2.1.3.1. Études sur l'utilisation des régions faciales

Considérons d'abord les résultats des six études ayant testé des zones faciales sélectionnées *a priori*, c.-à-d. les études de Roberts et Bruce (1988), Bruce et al. (1993), Brown et Perret (1993) et Yamaguchi, Hirukawa et Kanazawa (1995). Toutes ces expériences ont obtenu un résultat significatif ou un plus grand effet expérimental dans la région des sourcils (ou la grande région yeux-sourcils). Ces mêmes études soulignent aussi l'importance de la mâchoire et du menton, soit le contour du visage. Ces méthodologies comportent toutefois de sérieuses limites de validité. Premièrement, ces méthodologies requièrent la sélection *a priori* de zones faciales. En ce sens, les résultats quant à l'importance d'une zone deviennent donc tributaires de l'échantillon de zones testées. Deux régions ayant plus souvent été laissées de côté, mais qui semblent revêtir une certaine importance (donc qui auraient pu influencer les résultats globaux) sont la mâchoire et le menton. Deuxièmement, le découpage des zones anatomiques est généralement assez grossier. Par exemple, comme le font remarquer Bruce et al. (1993), les bandes habituellement apposées sur le nez masquent souvent le sourcil central (voir Figure 7B). De plus, les zones isolées comprennent souvent plus d'attributs que les auteurs le laissent entendre (voir Figure 7C). Troisièmement, il est probable que la manipulation systématique de certaines régions du visage ait incité les participants à adopter des stratégies locales. Par exemple, les études de masquage ont pu favoriser l'utilisation des zones ne subissant aucun masquage dans l'expérience. Dans la même veine, ces méthodes altèrent le réalisme des visages, en particulier la méthode de substitution, ce qui pourrait diminuer la validité externe des résultats.

Les techniques de classification d'images et d'oculométrie présentent quant à elles des échantillonnages sans *a priori*, c.-à-d. que tous les pixels de l'image ou presque y sont testés. Elles présentent donc un grand avantage sur le plan de la validité externe. Pour cette raison, nous les détaillerons et critiquerons une à une afin d'en évaluer la portée. D'abord, dans l'expérience de Nestor et Tarr (2008b), les participants devaient classifier le sexe d'un visage androgyne (le même à chaque essai) sur lequel une plage de bruit chromatique gaussien était ajoutée. Ce bruit chromatique était formé dans l'espace CIE Lab et se caractérisait par une variance égale à chacun des canaux. La Figure 7E montre un exemple de stimulus. Les participants devaient ainsi classifier, un à un, 20 000 androgynes bruités. Comme le seul élément changeant d'un essai à l'autre était le bruit visuel, les variations de réponse des participants ne pouvaient être imputables qu'à ce bruit. Afin de saisir l'impact de ce bruit sur les réponses des participants, les auteurs ont

utilisé une régression linéaire multiple où la valeur des pixels des plages de bruit utilisées agissait comme variable exploratoire et les réponses ('homme' ou 'femme') des participants comme variable dépendante. Un test statistique a ensuite été appliqué pour évaluer la probabilité d'observer le coefficient de régression associé à chaque pixel de l'image. La dernière colonne de la Figure 4B présente, pour chaque canal de l'espace des couleurs CIE Lab, les pixels significativement et positivement corrélés avec les réponses 'homme' (cyan) et les réponses 'femme' (magenta). Ces résultats soulignent particulièrement l'importance des pixels dans les zones orbitales, nasales et orales. Il faut toutefois noter ici que les jugements faciaux portaient sur un visage androgyne bruité, qu'aucune information dimorphique n'était réellement présente, et donc qu'aucune mesure de performance ne pouvait être utilisée. Ceci implique que les jugements faciaux étaient plus tributaires des représentations visuelles internes des participants que de l'utilisation des différents dimorphismes physiques (voir Gosselin & Schyns, 2002 pour une bonne illustration de ce principe). Comme le précisent les auteurs, le bruit dans les canaux a et b détenait des variations plus élevées que celles réellement présentes dans les visages. À titre d'exemple, les participants se représentaient la bouche comme plus claire chez l'homme et plus rouge chez la femme, même si les dimorphismes dans cette zone étaient, selon les statistiques d'images calculées par les auteurs (voir section 1.2.1), possiblement très faibles. Même si les représentations visuelles des participants devaient idéalement se baser sur des variations physiques réelles – la représentation de la bouche comme étant plus rouge chez la femme pourrait découler de l'utilisation répandue du rouge à lèvres — il est possible que ces résultats contiennent des biais représentationnels. Notez également que ces résultats sont entièrement dépendants de l'androgyne utilisé, en ce sens que la représentation visuelle révélée est à la fois fonction du bruit visuel additif et de l'androgyne utilisé. Ainsi, une même représentation visuelle aurait mené à des coefficients de régression différents si un autre visage androgyne avait été utilisé.

Une autre méthode faisant partie des techniques de classification d'images est la méthode Bubbles (voir Gosselin & Schyns, 2001; Schyns, Bonnar & Gosselin, 2002). Cette méthode consiste à échantillonner de façon aléatoire certaines dimensions d'un stimulus (p. ex. la dimension spatiale) à l'aide d'un bruit binomial multiplicatif ayant subi un passe-bas. Dans l'expérience de 2001, les participants devaient déterminer le sexe de visages échantillonnés sur la dimension spatiale (c.-à-d. l'axe xy). Plus spécifiquement, les visages étaient révélés à travers des

masques de bulles, c.-à-d. des masques gris opaques perforés de petites fenêtres gaussiennes (ci-après *bulles*)—un peu comme si on avait juxtaposé une tranche de fromage Gruyère sur un portrait. Les participants devaient ainsi classer, un à un, 512 visages partiellement révélés par ces masques de bulles. Dans l’expérience de 2002, les participants devaient classer 1000 visages échantillonnés à la fois sur l’axe *xy* et sur la dimension des fréquences spatiales. Chaque bulle révélait donc, pour une zone précise du visage, une portion de son contenu fréquentiel. Afin de connaître l’impact du bruit sur l’exactitude des réponses, les auteurs ont réalisé l’équivalent d’une régression multiple entre les masques de bulles (variables exploratoires) et les réponses correctes/incorrectes des participants (variables dépendantes). La résultante est une (ou des) *image(s) de classification* dont chaque pixel représente un coefficient de régression qui détermine, en quelque sorte, le lien entre l’information faciale révélée par le pixel et l’exactitude de la réponse. Plus la valeur des pixels est élevée, plus les informations faciales qu’ils révèlent sont importantes pour la tâche. Finalement, un intervalle de confiance ($p < 0.05$) a été utilisé afin d’identifier les coefficients de régression étant les plus importants statistiquement. Les détails de cette méthode et analyse sont exposés aux Chapitres 2 et 4. Les résultats de l’expérience de Gosselin et Schyns (2001) échantillonnant l’axe *xy* seulement sont exposés à la Figure 8B (image de gauche), tandis que les résultats de l’expérience de Schyns, Bonnar et Gosselin (2002) examinant l’axe *xy* et les fréquences spatiales sont illustrés à la Figure 8C (image de gauche). La Figure 8B (gauche) montre que les régions orale et orbitale (comprenant les yeux et les sourcils) sont significativement *corrélées* avec les réponses correctes. La Figure 8C (gauche) dévoile plusieurs régions significatives: l’oeil gauche en haute fréquence spatiale (FS), les zones orbitale, nasale et orale en moyennes hautes et moyennes basses FS et tout le contour du visage en moyennes basses FS (voir Schyns, Bonnar & Gosselin, 2002, pour des images de classification par bande de fréquences spatiales). La Figure 8C (droite) est le résultat d’une analyse de régression multiple menée sur les temps de réponse plutôt que sur l’exactitude de réponse. Les zones faciales qu’on y révèle sont donc celles étant significativement et positivement corrélées avec les réponses rapides des participants.

Contrairement à la méthode de “reverse correlation”, la technique Bubbles n’ajoute aucune information aux visages. Ce qu’on y révèle fait nécessairement partie des dimorphismes faciaux captés dans les photographies utilisées. Donc en ce sens, cette technique permet de cerner la

portion des informations physiques disponibles que l'observateur humain utilise pour catégoriser le sexe des visages. Sur ce point, Shycn, Bonnar et Gosselin (2002) ont également réalisé une analyse statistique afin d'évaluer les zones faciales utilisées par un *observateur idéal* ayant complété les mêmes essais que les observateurs humains. Cet observateur est dit idéal puisqu'il n'est contraint par aucune limite perceptive, attentionnelle ou mnémonique (c.-à-d. qu'il connaît tous les visages de l'expérience), et il accomplit la tâche optimalement (c.-à-d. qu'il maximise la probabilité de catégoriser correctement le sexe étant donné les l'information présentée à chaque essai). Cet observateur a d'abord été soumis à tous les essais de l'expérience, puis une analyse de régression multiple a été appliquée entre les plages de bruit et l'exactitude de ses réponses. Cette analyse permet d'identifier les zones faciales ayant maximisé ses bonnes réponses. L'image de droite à la Figure 8B illustre les régions du visage comportant les pixels significatifs, soit le contour et les régions orbitales et orales. Étant donné que cet observateur emploie toute l'information disponible pour la tâche, les régions faciales identifiées par analyse sont donc celles qui sont les plus diagnostiques du sexe des visages. Une critique à l'égard de ces études est la grande répétition d'un nombre assez restreint de visages. Une telle répétition peut promouvoir l'utilisation des particularités idiosyncrasiques des visages au détriment de leurs dimorphismes sexuels. Quoique possible, l'échantillonnage aléatoire propre à cette technique a nécessairement modéré l'utilisation d'une telle stratégie. En effet, même si un visage était présenté à répétition, les parties du visage révélées par les bulles variaient d'un essai à l'autre.

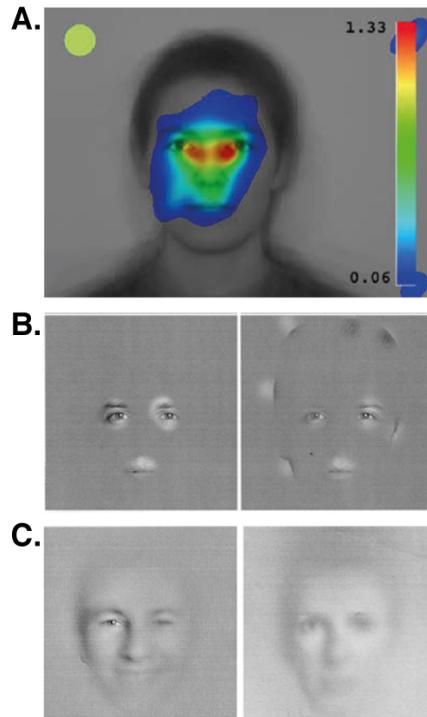


Figure 8. A) Probabilité qu'un pixel ait été balayé par la fovéa des participants (tiré de Saether et al., 2009). Les résultats sont exprimés selon un code de couleur de type « heat map » allant de la plus forte densité (rouge) à la plus faible (bleu). Le tout est juxtaposé à un visage moyen. B) Les zones du visage révélées sont celles dont les coefficients de régression étaient significatifs ($p < 0.05$ non corrigé pour comparaison multiple). À gauche figurent les résultats des observateurs humains et à droite ceux de l'observateur idéal. C) Les zones faciales et fréquences spatiales significativement ($p < 0.05$ non corrigé) corrélées avec l'exactitude de réponse (gauche) et la rapidité de réponse (droite).

Dans l'étude de Saether et al. (2009), des visages en couleurs d'hommes et de femmes étaient présentés un à un et les participants devaient en catégoriser le sexe pendant qu'étaient enregistrés leurs mouvements oculaires à l'aide d'un « eye-tracker ». Les auteurs ont d'abord estimé la fréquence des fixations dans neuf zones anatomiques déterminées *a priori* (voir Tableau 1). Les résultats soulignent l'importance du nez, des yeux et des joues. Les auteurs ont également réalisé une analyse sans *a priori* visant à estimer la probabilité qu'un pixel donné de l'image soit fixé ou balayé par la fovéa des participants. Les résultats, résumés au Tableau 1 et illustrés à la Figure 8A, suggèrent que la région infraorbitale et celle du pont nasal ont été les plus

fréquemment balayées par la fovéa. Une fois de plus, il est difficile d'interpréter ces données du point de vue de l'utilisation des informations dimorphiques parce qu'elles tiennent pour acquis, comme le soulignent les auteurs, que les zones fixées (ou les plus fréquemment survolées) ont bel et bien été utilisées pour faire la tâche. Les auteurs estiment que ce postulat a été respecté dans la mesure où les zones les plus fixées (ou survolées), c.-à-d. le nez, les yeux et les joues sont également diagnostiques du sexe des visages. Nonobstant le débat sur la signification réelle des données de cette technique (voir Saether et al., 2009), il demeure que les zones identifiées ne sont peut-être pas les seules utilisées dans cette tâche. En effet, les mesures oculaires utilisées (c.-à-d. les fixations et le balayage de la fovéa) mettent l'accent sur la région fovéale et font donc abstraction de la possibilité d'extraire les plus basses fréquences spatiales en vision périphérique. Il est possible, par exemple, que les participants aient traité l'information de la bouche en basses moyennes FS pendant qu'ils fixaient les régions infraorbitales.

En conclusion, il ressort de toutes ces études que la région orbitale (comprenant les sourcils et les yeux) est la plus importante pour la classification du sexe des visages. En plus des limites particulières de ces études, des problèmes méthodologiques minent leur validité externe. Premièrement, seulement trois de ces études ont testé des visages réalistes n'ayant pas subi de transformations majeures (excepté pour Saether et al., 2009; Gosselin & Schyns, 2001; Schyns, Bonnar & Gosselin, 2002). Qui plus est, toutes les études ont contrôlé l'apparence des stimuli (p. ex. en cachant les cheveux, les oreilles, etc.), les rendant ainsi moins représentatifs. De plus, les visages utilisés dans toutes les expériences — à l'exception de Saether et al. (2009), Nestor et Tarr (2008b) et Hill, Bruce & Akamatsu (1995)— étaient en tons de gris. Pourtant, des preuves ont démontré (1) l'existence d'information faciale dimorphique d'ordre purement chromatique et (2) l'intégration de ces informations dans la représentation des observateurs humains (p. ex. Nestor & Tarr, 2008b, Hill, Bruce & Akamatsu, 1995). Deuxièmement, toutes les études à l'exception de quatre (c.-à-d. Nestor & Tarr, 2008b; Gosselin & Schyns, 2001; Schyns, Bonnar & Gosselin, 2002; Saether et al., 2009) utilisent des méthodologies comme l'isolement, le masquage et la substitution qui requièrent la sélection *a priori* de certaines zones faciales. Et parmi les études testant tous les pixels des visages, seule la technique Bubbles permet d'étudier directement l'utilisation des informations dimorphiques. En effet, l'oculométrie renvoie à des mesures *indirectes* — c.-à-d. basées sur des postulats — d'utilisation de l'information dimorphique, tandis

que la “reverse correlation” révèle plutôt la représentation visuelle des observateurs (voir Gosselin & Schyns, 2002). Troisièmement, seules cinq de ces études ont aligné leurs visages de sorte à contrôler pour la différence de taille absolue entre les visages d’hommes et de femmes. Comme nous l’expliquerons davantage à la section 1.2.2.1, il est peu probable que les mécanismes perceptifs utilisés au quotidien pour faire des jugements faciaux s’appuient sur la taille absolue des visages et de leurs attributs étant donné la difficulté de l’observateur humain à estimer la distance physique des objets distaux. Il est toutefois possible que certains participants aient décelé et utilisé de tels indices dans les études qui ne contrôlaient pas ces dimorphismes.

1.2.1.3.2. Études sur l’importance relative des indices achromatiques et chromatiques

Seules deux études ont évalué l’importance relative des indices achromatiques et chromatiques, soit l’étude de Nestor et Tarr (2008b) et de Hill, Bruce & Akamatsu (1995). Les résultats de l’étude de Nestor et Tarr (2008b), introduite dans un paragraphe antérieur, indiquent que l’information dimorphique contenue dans le canal rouge-vert (A) est plus importante que celle contenue dans le canal jaune-bleu (B) au niveau des représentations visuelles. Cette étude ne permet toutefois pas de dresser des conclusions claires quant à la différence perceptive relative entre le canal achromatique (L) et les canaux chromatiques (A et B), étant donné la variance anormalement élevée du bruit dans les canaux A et B. Hill, Bruce & Akamatsu (1995) ont pour leur part contrasté l’importance perceptive de deux indices : les ombrages provenant des formes tridimensionnelles du visage (« shape-from-shading ») et les indices chromatiques et achromatiques (c.-à-d. incluant certains ombrages). Afin d’isoler les indices d’ombrage, les auteurs ont d’abord extrait la forme tridimensionnelle de la surface de la peau de 16 visages (8 hommes et 8 femmes) à l’aide d’un scanner, et ont ensuite simulé l’effet sur cette forme 3D d’une lumière projetée à partir du haut. Le résultat est un visage gris dénué de tout indice de surface qui ressemble un peu à la sculpture d’une tête en ronde-bosse (voir Figure 7D à gauche). Quant aux indices achromatiques et chromatiques de surface, les auteurs les ont extraits en balayant latéralement les visages à l’aide d’une caméra vidéo et en projetant l’ensemble des images captées sur une surface 2D. Les stimuli ainsi créés sont représentés à la Figure 7D (droite). Durant l’expérience, les participants devaient catégoriser la race (c.-à-d. Caucasiens vs Japonais) et le sexe des visages ne comportant que les indices d’ombrage (Expérience 1) ou de

surface (Expérience 2). Les résultats sur la justesse des réponses suggèrent que les indices d'ombrage sont plus importants pour déterminer la race (88%) que le sexe (72%), tandis que les indices de surface sont plus importants pour le sexe (97%) que la race (90%). Selon ces données, nous pouvons également extrapoler que les indices de surface sont plus importants pour la catégorisation du sexe que les indices d'ombrage. Quoiqu'il en soit, la pertinence de ces résultats pour le domaine de la discrimination du genre et leur validité externe est discutable. D'abord, les visages utilisés sont tous invraisemblables. De plus, l'importance relative des indices de surface par rapport aux indices d'ombrage découle logiquement du fait que les premières englobent les dernières. En conclusion, aucune étude recensée n'a évalué adéquatement l'utilisation relative des indices achromatiques et chromatiques pour la catégorisation du sexe des visages.

1.2.2. Indices faciaux relationnels

Les observateurs humains présentent une certaine expertise en perception des visages (p. ex. Gauthier & Tarr, 2002; Gauthier et al., 2000). En effet, ils parviennent à identifier un visage, en catégoriser le sexe, en déceler l'expression faciale, en estimer l'âge et l'ethnie dans des conditions d'illuminations très variables, sous des angles de vue différents et à de multiples distances. Malgré cette flexibilité, il suffit d'orienter un visage vers le bas (c.-à-d. lui faire subir une rotation de 180° sur le plan de l'image) pour altérer le bon fonctionnement des processus faciaux. En effet, l'inversion des visages induit une réduction de l'exactitude de son identification (et de la catégorisation de son sexe) et une augmentation des temps de réponse. Qui plus est, cet effet d'inversion est significativement plus important pour les visages que pour d'autres objets mono-orientés (p. ex. maison, voiture, chien, table, etc.), ce qui suggère l'implication d'un processus spécifique aux visages (Hochberg et Galper, 1967; Bruce et al., 1993; Bruce et Langton, 1994; Leder et Carbon, 2006; Robbins et McKone, 2007; Sergent et Corballis, 1989; Yin, 1969; voir Rossion, 2008). Cet effet a été largement interprété comme la preuve de l'importance des processus faciaux dédiés à l'intégration des informations d'ordre spatial, c.-à-d. les *processus configuraux*. De tels processus se distingueraient des processus dits *analytiques* ou par composante ("componential") qui traitent des parties individuelles des visages, et ce, indépendamment les unes des autres.



Figure 9. Exemple de visages composites alignés (bas) et désalignés (haut). La portion supérieure de chaque image réfère à la même identité, disons A. La portion inférieure des images de gauche réfère à l'identité A (stimuli concordants) tandis que celle des images de droite réfère à l'identité B (stimuli discordants; tiré de Mondloch et al., 2007).

Selon Maurer, Le Grand et Mondloch (2002), les *processus configuraux* comprennent les processus *holistiques* ainsi que les processus traitant les *indices relationnels de premier et de deuxième ordre*. Les processus holistiques font référence à tous les processus de type *gestaltiste* traitant l'ensemble des éléments d'un visage en un tout indissociable. L'effet composite en serait un des signes diagnostiques (voir Young, Hellawell & Hay, 1987; Konar, Bennett & Sekuler, 2010; Richler, Cheung & Gauthier, 2011). Cet effet est généralement démontré dans des expériences présentant des visages dont la partie supérieure et inférieure peut représenter la même identité (stimuli concordants) ou des identités différentes (stimuli discordants, voir Figure 9). De plus, ces deux types de visages sont soit alignés (i.e. les deux parties sont verticalement alignées) ou soit désalignés (i.e. la partie inférieure est décalée horizontalement vers la droite ou la gauche, voir Figure 9). Les participants sont plus lents et moins habiles à identifier correctement le haut d'un visage discordant que concordant. De plus, l'élément clé de l'effet composite est que cette diminution de la performance s'estompe lorsque les visages sont désalignés. L'hypothèse proposée est que seuls les visages alignés engagent les processus holistiques et comme ces processus regroupent toutes les régions du visage en un tout indissociable, le bas des visages affecte donc la perception du haut (voir Richler, Cheung &

Gauthier, 2011). Notons que cet effet a également été observé dans des tâches de reconnaissance de l'expression faciale (p. ex. Calder et al., 2000) et de discrimination du sexe des visages (p. ex. Baudouin & Gallay, 2006).

Les indices relationnels de premier ordre caractérisent les régularités présentes dans l'organisation spatiale des attributs du visage: un nez en son centre, deux sourcils au-dessus de deux yeux placés latéralement au-dessus du nez, et une bouche directement sous le nez. Cette configuration stable se trouve chez tous les primates et chez plusieurs autres espèces. Ces indices permettraient de distinguer les visages parmi d'autres objets visuels.

Les indices relationnels de deuxième ordre font plutôt référence aux distances et aux angles entre les principaux attributs d'un visage. L'idée selon laquelle les relations spatiales entre les principaux attributs faciaux (c.-à-d. nez, bouche, yeux et sourcils) représenteraient un indice stable et invariant à la distance pouvant servir à la reconnaissance faciale a d'abord été proposée par Galton (1879). Haig (1984) fut l'un des premiers à tester directement cette hypothèse. Des participants devaient d'abord se familiariser à une série de visages caucasiens. Ensuite, on leur présentait plusieurs visages dont le positionnement relatif de leurs attributs avait été modifié. Les participants devaient départager les visages non modifiés des visages ainsi modifiés. Les résultats indiquent que la sensibilité —telle que mesurée par la « just-noticeable-difference »— à certains déplacements approchait la limite de l'acuité visuelle. Par exemple, les participants pouvaient déceler un déplacement de la bouche (vers le haut) d'une minute d'angle visuel. Ces résultats suggèrent l'existence d'un mécanisme très précis d'encodage des relations spatiales des visages. En fait, plusieurs autres études ont également démontré que le traitement des distances relatives entre des paires d'attributs faciaux (ci-après, les distances interattributs) était plus affecté par l'inversion des visages que le traitement de la forme des attributs et de leurs caractéristiques de surfaces (p. ex. pigmentation, texture, ombrages, etc.; Barton et al., 2001; Freire, Lee, et Symons, 2000; Le Grand et al., 2001; Rhodes et al., 2007; Sergent, 1984; voir Rossion, 2008, 2009, pour une revue et McKone et Yovel, 2009 pour une explication alternative).

Dans les deux sections suivantes, nous présenterons des études évaluant l'importance statistique (section 1.2.2.1) et perceptive (section 1.2.2.2) des indices relationnels de deuxième ordre pour la catégorisation du sexe des visages. Nous avons choisi d'investiguer ces indices au

détriment d'autres indices configuratifs principalement pour deux raisons. D'abord, tout un pan de la littérature sur la perception des visages porte sur l'importance des indices relationnels, et certaines études, que nous détaillerons à la section 1.2.2.2, suggèrent que ces indices jouent un rôle dans la catégorisation du sexe des visages. Deuxièmement, ces indices renvoient à des mesures physiques tangibles ce qui n'est pas le cas des processus holistiques dont la présence et l'influence doivent plutôt être inférées à partir d'effets expérimentaux comme l'effet composite. Notez que les indices relationnels de premier ordre sont aussi directement accessibles, néanmoins leur rôle dans les jugements faciaux est probablement de second ordre.

1.2.2.1. Dispersion et importance statistique des indices relationnels de deuxième ordre

Étant donné la différence de taille absolue entre les visages d'hommes et de femmes, il s'ensuit logiquement que certaines distances craniofaciales absolues seront plus grandes chez les hommes que chez les femmes (p. ex. la distance interpupillaire). De telles distances absolues contiendraient donc de l'information dimorphique. Au niveau perceptif cependant, ces indices ne sont accessibles qu'à de très courtes distances (<2 m). Ceci s'explique par le fait que la taille absolue des objets distaux doit être inférée en fonction de la distance des objets, ce qui n'est possible qu'à moins de 2 m via des indices oculomoteurs tels que la convergence oculaire et l'accommodation (Palmer, 1999). Pourtant, l'observateur humain peut aisément percevoir et catégoriser des visages situés à plus de 2 m. En effet, les résultats d'une étude montrent que la performance maximale atteinte dans une tâche d'identification des visages survient à une distance supérieure à 2 m (voir Taschereau-Dumouchel, et al., 2010; Rhodes, 1988). Pour ces raisons, seules les distances faciales *relatives* seront considérées dans cette thèse.

1.2.2.1.1. Travaux de Farkas

Farkas et Munro (1987) ont extrait des mesures craniofaciales *relatives* sur 1197 visages caucasiens mâles et 1367 visages caucasiens femelles âgés de 6 à 18 ans (Figure 10). Plus spécifiquement, ils rapportent 166 ratios entre paires de distances craniofaciales et analysent, entre autres, l'évolution des dimorphismes sexuels entre 6 et 18 ans. Le tableau 2 présente

quelques exemples de ratios significatifs dans les zones faciale, orbitale, nasale et orale de visages âgés de 18 ans.

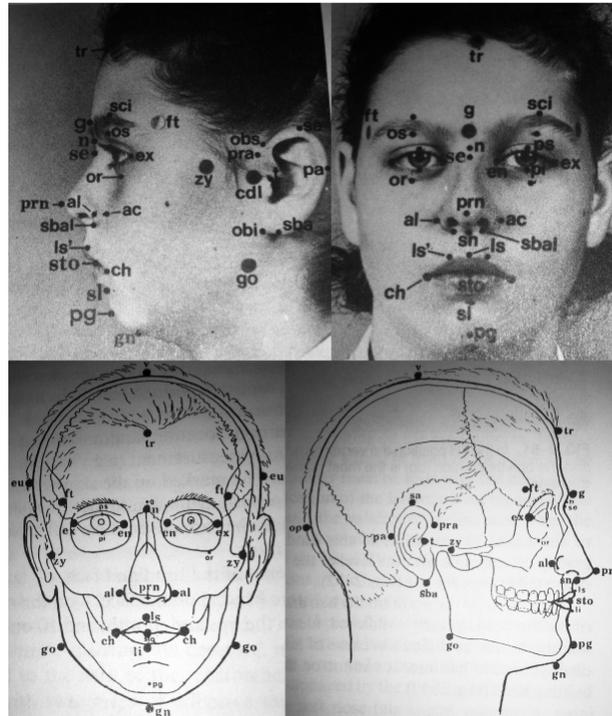


Figure 10. Repères craniofaciaux de surface (haut) et squelettiques (bas) en vue frontale et latérale utilisés dans Farkas (1981) et Farkas et Munro (1987).

Tableau 2. Exemples des ratios faciaux (en %) significatifs rapportés par Farkas et Monro (1987), et statistiques associées aux visages masculins et féminins.

Zone	Repères	N	Hommes		Femmes		t	p	d-cohen	
			M	SD	M	SD				
Faciale	(n-gn)/(zy-zy)	52	88.5	5.1	51	86.2	4.6	2.40	.018	0.48
Orbitale	(ps-pi)/(ex-en)	52	33.4	3.5	51	36.2	3.1	-4.3	.0000	-0.85
Nasale	(n-sn)/(al-al)	52	23.3	2.2	51	24.3	2.3	-2.3	.026	-0.45
	(ac-prn)/(n-sn)	52	65.3	4.4	51	63	4.2	2.71	.008	0.54
Orolabiale	(sn-sto)/[(sn-ls) + (ls-sto)]	13	92.2	6.2	8	85.7	5	2.50	.022	1.15

Une des conclusions générales tirée de cette enquête est que malgré le fait que plus de la moitié des ratios répertoriés varient significativement entre les mâles et les femelles, ces dimorphismes ne s'amplifient que très légèrement entre l'âge de 6 et 18 ans, à l'exception peut-être de la région orale. Ce résultat est quelque peu surprenant si l'on considère que la justesse de discrimination du genre des visages augmente significativement avec l'âge de ceux-ci (Wild et al., 2000). Quelques facteurs méthodologiques pourraient expliquer ce décalage. D'abord, il faut

savoir que les 166 ratios recensés dans l'ouvrage de Farkas et Monro (1987) représentent moins de 2% des ratios entre paires de distances qui auraient pu être calculés à partir des 132 distances que Farkas avait répertorié dans une étude précédente (voir Farkas, 1981). Cette sélection s'appuie principalement sur les mesures classiques issues d'études anthropologiques du 18^e et 19^e siècle, et s'inscrit dans la foulée d'ouvrages médicaux du 20^e siècle montrant un lien entre des disproportions corporelles et certains déséquilibres hormonaux, anomalies congénitales et syndrômes. Dans ce contexte, il est probable que les ratios sélectionnés et la méthodologie choisie ne se prêtent pas à l'étude des dimorphismes sur le plan perceptif. Plus particulièrement, l'utilisation de ratios basés sur un seul et même dénominateur aurait semblé plus appropriée si le but avait été de comparer leur apport respectif. C'est précisément la procédure qu'ont suivi Burton, Bruce et Dench (1993).

1.2.2.1.2. Étude de Burton

Plus spécifiquement, Burton, Bruce et Dench (1993) ont identifié 107 repères craniofaciaux sur le visage caucasiens de 91 hommes et 88 femmes (voir Figure 11). Le but de cette recherche était de sélectionner un ensemble de mesures *relatives* qui, lorsqu'introduit dans un modèle d'analyse discriminante de type "stepwise", pourrait rendre compte d'un taux de discrimination comparable à celui des humains, soit d'environ 0.96. Les mesures en question étaient soit de simples distances euclidiennes entre deux repères, soit des ratios de distances ou encore des angles. Il est à noter que les distances étaient toutes divisées par la distance interpupillaire. Burton et al. ont procédé à quatre tentatives plus ou moins fructueuses pour finalement aboutir à un pourcentage de discrimination correcte de 94%. Ce dernier modèle comprenait 16 mesures dont les cinq plus discriminantes étaient (1) la distance entre le lobe du nez et sa commissure (2) la hauteur du front, (3) la largeur du visage à la hauteur du menton, (4) la distance entre le lobe nasal et le sillon nasolabial ainsi que (5) la distance yeux-sourcils. Afin de valider leur modèle, les auteurs ont ensuite comparé ses erreurs de classification à celles faites par des sujets humains jugeant les mêmes visages. Les résultats indiquent que seuls 9% des visages incorrectement classés par le modèle étaient également incorrectement classés par au moins deux des 13 participants humains. De ce piètre résultat, les chercheurs dressent les conclusions suivantes. D'abord, il semble ardu de sélectionner, parmi les millions de mesures faciales relatives

calculables (c.-à-d. distances, ratios et angles), celles discriminant adéquatement les hommes des femmes. Ensuite, ce résultat suscite la question de l'importance de ces indices de distances relatives par rapport à d'autres dimorphismes comme les indices de surface (c.-à-d. la pigmentation de la peau et sa texture). Finalement, nous ajouterons que la haute performance atteinte par leur modèle ne reflète pas réellement celle obtenue par les observateurs humains: contrairement au modèle, les participants ne connaissaient pas les visages à discriminer et devaient donc se baser sur leurs connaissances *a priori*. En ce sens, une comparaison juste avec la performance humaine requerrait l'emploi de modèles prédictifs plutôt que descriptifs.

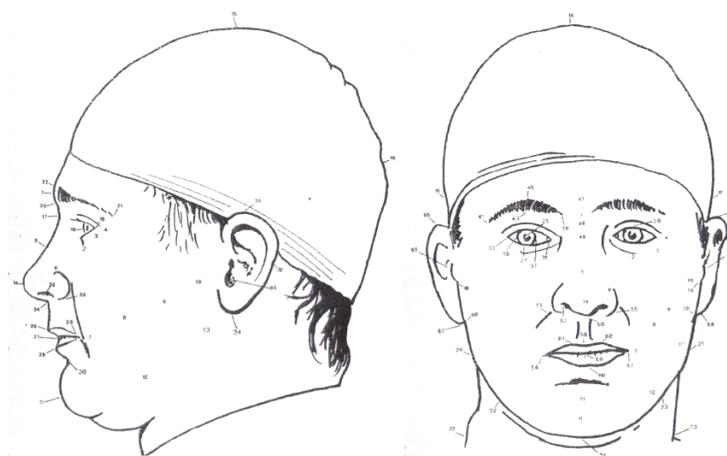


Figure 11. Repères craniofaciaux de surface en vue latérale et frontale utilisés par Burton, Bruce et Dench (1993).

1.2.2.1.3. Étude de Taschereau-Dumouchel

En 2010, Taschereau-Dumouchel, Rossion, Schyns et Gosselin (2010) ont testé l'hypothèse selon laquelle les distances relatives entre les principaux attributs faciaux (c.-à-d. nez, bouche, yeux et sourcils) sont cruciales à l'identification des visages. Pour ce faire, ils ont d'abord prélevé les distances interattributs sur 515 visages caucasiens adultes non modifiés (256 femmes et 259 hommes). Ensuite, ils ont demandé à des participants de sélectionner parmi deux visages, celui qui ressemblait le plus au visage cible qu'on venait juste de leur présenter (c.-à-d. une tâche de « matching-to-sample »). Les participants devaient ainsi appairer des visages dont seules les distances interattributs variaient (expérience 1) et des visages contenant tous les autres indices

achromatiques, mais dont les distances interattributs avaient été fixées à une valeur constante (expérience 2). La Figure 12 présente des exemples de visages utilisés dans ces deux expériences.

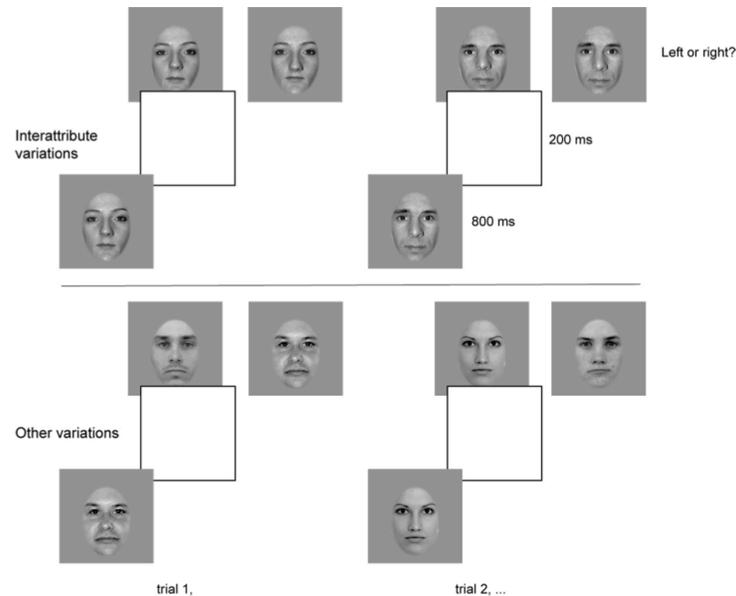


Figure 12. Exemples de stimuli utilisés dans l'expérience 1 (haut) et l'expérience 2 (bas) de Taschereau-Dumouchel, et al. (2010). Dans chaque essai des deux expériences, un visage cible était d'abord présenté pendant 800 ms ; ensuite une plage blanche était affichée (200 ms) ; et finalement deux visages apparaissaient côte à côte jusqu'à ce que le participant ait sélectionné celui qui ressemblait le plus à la cible. Les visages de l'expérience 1 (haut) varient en fonction de leurs distances interattributs tandis que les visages de l'expérience 2 (bas) contiennent toutes les sources de variation à l'exception des distances interattributs.

Les performances moyennes observées à l'expérience 1 (~65%) s'approchaient du niveau plancher tandis que celles enregistrées à l'expérience 2 (~99.5%) s'approchaient de la perfection. Ceci suggère que, contrairement aux résultats de plusieurs études antérieures, les distances interattributs ne sont pas cruciales à l'identification des visages. Afin de mieux comprendre ces résultats, les auteurs ont comparé les distances interattributs qu'ils avaient prélevées sur leurs 515 visages aux distances interattributs utilisées dans 14 autres études sur le sujet. Le résultat de cette comparaison est révélateur : ces 14 études avaient exagéré, par un facteur de 376%, les distances interattributs. Advenant que les participants se soient fiés à ce signal amplifié, ceci pourrait expliquer en partie les résultats démontrant l'importance de ces indices. Cette étude montre donc

l'importance d'utiliser des distances interattributs qui soient représentatives de celles observées dans l'environnement. Notez qu'une des causes sous-jacentes de cette amplification vient du fait que plusieurs de ces études manipulaient certaines distances individuellement sans tenir compte des covariations entre elles. Par exemple, en modifiant exclusivement le positionnement de la bouche, Haig (1984) se trouva en fait à modifier les distances bouche-nez, bouche-yeux et bouche-sourcils, affectant ainsi presque toute la matrice de covariance entre ces distances. Il s'agit donc non seulement de tenir compte des distances réelles, mais aussi des covariations entre celles-ci.

1.2.2.1.4. Dispersion et importance statistique des distances interattributs représentatives

Afin d'estimer l'importance statistique des distances interattributs *représentatives*, nous nous sommes basés sur les données de Taschereau-Dumouchel et al. (2010). Nous avons d'abord estimé l'information dimorphique contenue dans chacune des 15 distances interattributs possibles entre les six principaux attributs faciaux. Pour ce faire, nous avons d'abord appliqué, pour chacune de ces distances, une régression logistique entre le sexe des 515 visages (variable nominale dépendante) et leur distance correspondante (variable continue exploratoire). Afin de tenir compte des covariations entre ces distances, nous avons également appliqué une régression logistique multiple entre le sexe des 515 visages et l'ensemble des 15 distances. Dans le but de comparer nos résultats à d'autres dimorphismes sexuels rapportés dans la littérature, nous rapportons la taille d'effet des régressions en d de Cohen.

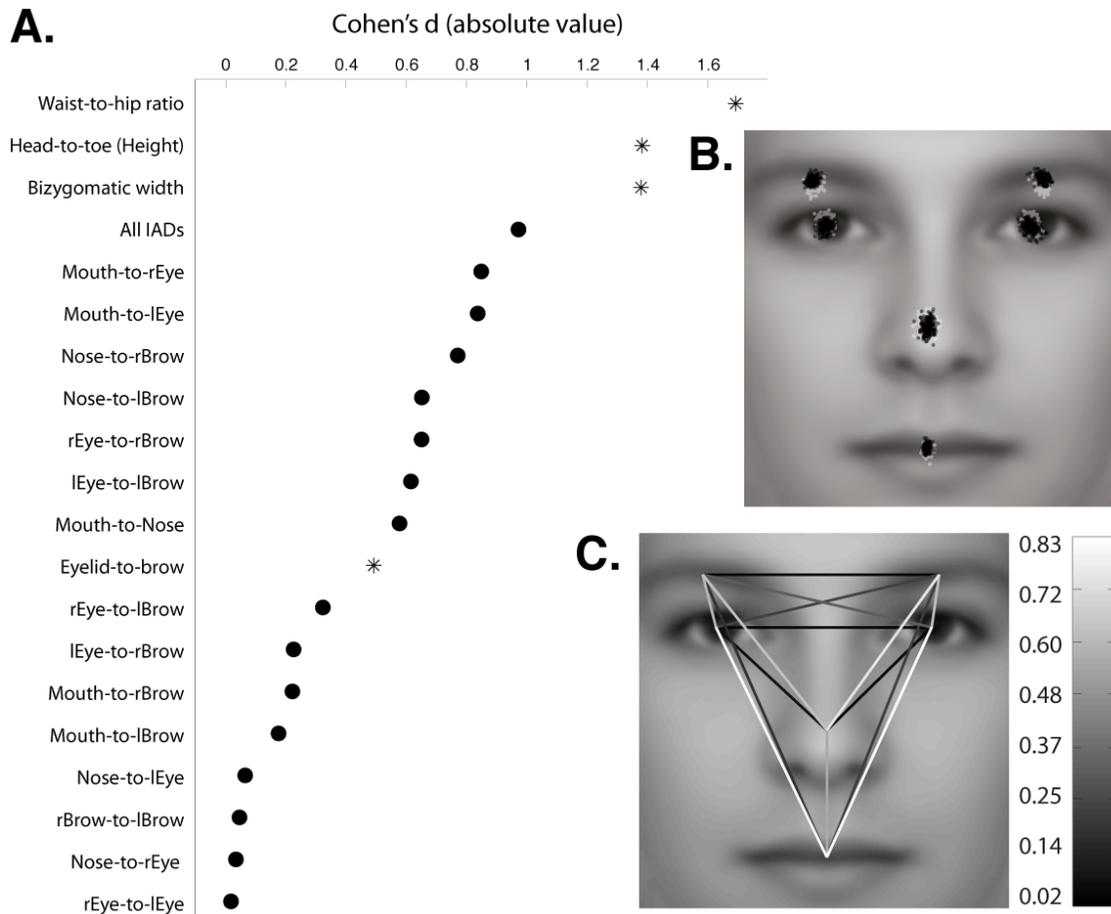


Figure 13. A) Taille d'effet (d de Cohen) associée à chacune des DIA individuelles et agrégées (points noirs), ainsi que la taille d'effet de quatre autres dimorphismes rapportés dans la littérature (astérisques). B) Centres de masse des attributs faciaux féminins (points noirs) et masculins (points blancs) juxtaposés sur un visage androgyne. C) Représentation imagée de la taille d'effet des DIA individuelles.

La Figure 13A présente la taille d'effet associée à chacune des régressions faites sur les distances interattributs individuelles et agrégées. À des fins de comparaison, nous affichons aussi la taille d'effet liée à quatre autres dimorphismes sexuels rapportés dans la littérature (astérisques), soit la distance cil-sourcil (Campbell et al., 1999), la largeur bizygomatique¹, la taille corporelle (Ogden et al., 2004) et le ratio taille-hanche (Dobbelsteyn et al., 2001). La Figure 13B présente les centres de masse des attributs faciaux féminins (points noirs) et masculins (points blancs) juxtaposés sur un visage androgyne. La Figure 13C offre une représentation plus

¹ Cet indice a été calculé à partir des données de Taschereau-Dumouchel et al., 2010.

intuitive des tailles d'effet individuelles présentées à la Figure 13A. Plus spécifiquement, les 15 distances interattributs sont représentées sous forme de segments dont la teinte de gris correspond à leur taille d'effet individuelle. Selon ces résultats, la distance interattribut relative la plus informative est la distance bouche-yeux tandis que la moins informative est la distance interpupillaire. Une fois intégrées dans une même régression, les distances interattributs relatives mènent à une taille d'effet de 0.94. Comparées à d'autres dimorphismes relatifs connus comme le ratio taille-hanche, les distances interattribut sont donc moyennement informatives. De plus, il semble que le gain de taille d'effet résultant de l'agrégation de toutes les distances interattributs est mince et suggère une covariance élevée.

1.2.2.2. Importance perceptive des indices relationnels de deuxième ordre

Des travaux suggèrent des interactions possibles entre certaines distances interattributs et la catégorisation du sexe des visages. Par exemple, Campbell, Wallace et Benson (1996) ont observé que les visages dont le regard était orienté vers le bas étaient jugés comme étant plus féminins que les visages dont le regard était orienté normalement. Ils ont interprété ce résultat comme provenant de l'effet de la direction du regard sur la distance oeil-sourcil, soit une distance assez informative sur le plan dimorphique selon Burton, Bruce, et Dench (1993).

Campbell et al., (1999) ont exploré l'impact de l'élévation et de l'abaissement des sourcils, soit deux positions naturellement impliquées dans certaines émotions (p. ex. joie et colère), sur la catégorisation rapide du sexe des visages. Leurs résultats démontrent que les visages mâles dont les sourcils étaient abaissés — notez que les visages n'étaient pas dynamiques — étaient catégorisés avec plus d'exactitude et de rapidité que tous les autres visages présentant des sourcils élevés ou abaissés.

Bien que ces études suggèrent une possible interaction entre certaines distances interattributs et la discrimination du sexe des visages, aucune d'entre elles n'a manipulé ces informations de façon distincte.

1.3. Présentation des articles de la thèse

1.3.1. *Premier article*

Comme nous l'avons vu à la section 1.2.1., la plupart des études examinant l'importance perceptive de certaines régions faciales dans la catégorisation du sexe des visages présentaient des limites quant à leur validité externe. Le premier article de cette thèse (voir Chapitre 2) vise donc à investiguer l'utilisation des indices achromatiques et chromatiques (sur l'axe xy) dans un contexte de plus grande validité externe. Pour ce faire, nous utilisons la technique Bubbles afin d'échantillonner l'espace xy de visages en couleurs n'ayant subi aucun contrôle ou transformation. Afin d'éviter les problèmes liés à la grande répétition des mêmes visages, nous utilisons un grand nombre de visages (c.-à-d. 300 visages caucasiens d'hommes et de femmes) et chaque visage n'est présenté qu'une seule fois au participant. La technique Bubbles a été préférée aux autres techniques principalement parce qu'elle permet d'étudier directement l'information P , soit l'utilisation de l'information dimorphique, et qu'elle ne requière pas la sélection *a priori* de zones faciales.

1.3.2. *Deuxième article*

Comme nous l'avons vu à la section 1.2.2, plusieurs études suggèrent que le traitement des distances interattributs est crucial à la perception des visages, incluant la catégorisation du sexe. L'étude de Taschereau et al. (2010; voir section 1.2.2.1.3) présente toutefois des données allant à l'encontre de cette hypothèse : les performances d'identification des visages étaient beaucoup plus faibles lorsque seules les distances interattributs réalistes étaient disponibles que lorsque toutes les autres informations faciales à l'exception des distances interattributs réalistes étaient disponibles. Quoi qu'il en soit, il est possible que la faible performance observée dans la condition où seules les distances interattributs étaient disponibles soit explicable non par une incapacité d'utiliser ces indices efficacement, mais plutôt par le peu d'information contenue dans ces indices. Un rejet définitif de l'hypothèse sur l'importance des distances interattributs nécessiterait donc de départager ces deux possibilités. Également, le rejet d'une telle hypothèse nécessiterait qu'elle soit testée sur d'autres jugements faciaux. Le deuxième article a donc comme objectif premier d'évaluer l'efficacité — une mesure de performance qui compense pour

la faiblesse de l'information disponible — des distances interattributs réalistes pour la catégorisation du sexe des visages. Afin de maximiser la validité externe, les distances interattributs manipulées respecteront la distribution et la matrice de covariance observées dans un large échantillon de visages (N=515).

Les études sur l'importance relative des indices achromatiques et chromatiques dans la discrimination du sexe des visages sont peu concluantes (voir section 1.2.1.2.3). Néanmoins, d'autres études sur l'identification des visages suggèrent un rôle de bas niveau pour la couleur. En effet, Yip et Sinha (2002) ont proposé l'hypothèse selon laquelle la couleur pourrait non seulement transmettre des informations supplémentaires sur l'identité des visages, mais pourrait également aider à la segmentation des surfaces et des parties du visage. Le second objectif de l'article est donc d'évaluer l'importance relative des indices chromatiques et achromatiques et de tester l'hypothèse de Yip et Sinha. Étant donné que des visages en couleurs contiennent nécessairement plus d'information sexuelle que des visages en tons de gris, une mesure d'efficacité est utilisée afin de comparer l'utilisation relative des informations fournies par ces deux types d'indices.

1.3.3. Troisième article

Dans le premier article, nous avons observé de fortes corrélations entre les réponses rapides/correctes et les pixels situés dans la région de la bouche. Compte tenu de la richesse des informations chromatiques contenues dans cette région (voir section 1.2.1.1) d'une part, et des preuves, tant en perception des objets qu'en perception des visages (p. ex. Tarr et al., 2001; Gegenfurtner & Rieger, 2000; Holcombe & Cavanagh, 2001) qui montrent, d'autre part, le traitement rapide de cette information, nous avons postulé que de telles informations pouvaient être extraites très rapidement par le système visuel. À l'origine, l'objectif du troisième article était donc de tester cette hypothèse dans le cadre de la catégorisation du sexe des visages en échantillonnant l'information achromatique et chromatique avec des *bulles* spatiotemporelles (voir Vinette, Gosselin & Schyns, 2004; Fiset et al., 2009). L'utilisation de ce type d'échantillonnage nous a également permis de faire des analyses temps-fréquences desquelles a découlé une découverte intéressante quant aux mécanismes d'encodage des informations

spatiales dans le temps. Pour cette raison, l'article trois porte davantage sur cette trouvaille que sur le test de notre hypothèse d'origine.

Chapitre 2: Article 1

Uncovering gender discrimination cues in a realistic setting

Nicolas Dupuis-Roy¹, Isabelle Fortin¹, Daniel Fiset² & Frédéric Gosselin¹

1. Département de psychologie, Université de Montréal.
2. Département de psychoéducation et de psychologie, Université du Québec en Outaouais

Address correspondence to

Nicolas Dupuis-Roy, Département de psychologie, Université de Montréal, PO Box 6128, Succ. Centre-Ville, Montréal, QC H3C 3J7, Canada.

Article published in *Journal of Vision* (2009), 9 (2), 1-8.

2.1. Abstract

Which face cues do we use for gender discrimination? Few studies have tried to answer this question and the few that have tried typically used only a small set of grayscale stimuli, often distorted and presented a large number of times. Here, we reassessed the importance of facial cues for gender discrimination in a more realistic setting. We applied Bubbles—a technique that minimizes bias toward specific facial features and does not necessitate the distortion of stimuli—to a set of 300 color photographs of Caucasian faces, each presented only once to 30 participants. Results show that the region of the eyes and the eyebrows—probably in the light-dark channel—is the most important facial cue for accurate gender discrimination; and that the mouth region is driving fast correct responses (but not fast incorrect responses)—the gender discrimination information in the mouth region is concentrated in the red-green color channel. Together, these results suggest that, when color is informative in the mouth region, humans use it and respond rapidly; and, when it’s not informative, they have to rely on the more robust but more sluggish luminance information in the eye-eyebrow region.

2.2. Introduction

Which face cues are we using for gender discrimination? Up until now, the small body of studies on this topic has highlighted the importance of the eyes, the eyebrows, the jaw and the face outline (e.g., Brown & Perrett, 1993; Nestor & Tarr, 2008a, 2008b; Russell, 2003, 2005; Yamaguchi, Hirukawa, & Kanazawa, 1995). Using Bubbles, Schyns, Bonnar, and Gosselin (2002; see also Gosselin & Schyns, 2001) found that relatively coarse eye and mouth information (5.62–22.5 cycles per face width for a face width subtending about 4 cycles per degree of visual angle) were significantly correlated with gender discrimination in humans. Relatedly, the distance between the brows and the upper eyelid was identified as the most reliable relational cue to gender in facial images (Burton, Bruce, & Dench, 1993; Campbell, Benson, Wallace, Doesbergh, & Coleman, 1999). Experiments investigating the role of pigmentation cues showed that human observers could rely on chromatic information—mostly on the red-green axis—to categorize gender especially when minimal discriminative shape information were revealed (Bruce & Langton, 1994; Hill, Bruce, & Akamatsu, 1995; Tarr, Kersten, Cheng, & Rossion, 2001; Tarr, Rossion, & Doerschner, 2002). The regions surrounding the eyes and the mouth were also found to be the most determinant chromatically (Nestor & Tarr, 2008b).

All the studies cited above suffer from at least one of the following three potentially serious limitations on external validity. First, all of them—except Gosselin and Schyns (2001), Nestor and Tarr (2008a), and Schyns et al. (2002)—manipulated specific features and regions of the face with techniques such as morphing and caricaturing. These manipulations could have distorted the natural characteristics of authentic faces. Moreover, selective manipulation of these features might have biased the results toward a limited sample of all the facial information available. Second, the face stimuli used in all of these studies—except the studies performed by Tarr and colleagues—were grayscale pictures or they were controlled for different aspects (e.g., hair and ears removed, no makeup). In fact, the skin and hair reflectance properties of males and females differ (makeup only exaggerates these spectral dimorphism—Russell, 2003) and, as we have mentioned above, human observers can use these differences reliably. Third, all of these studies—except Nestor and Tarr (2008a)—used a small set of faces that needed to be shown

many times to each participant. This context is likely to have promoted perceptual learning of the faces. Therefore, the results might reflect the peculiarities of the stimulus set rather than general characteristics of gender dimorphism. In fact, the repetition of the same face identity allows the subject to use a face identification strategy rather than a gender discrimination strategy. This may have artificially increased the role of eye region, a potent feature for face recognition (Gosselin & Schyns, 2001; Schyns et al, 2002; Sekuler, Gaspar, Gold, & Bennett, 2004).

Here, we reassess the importance of facial cues for gender discrimination in a more realistic setting: We apply Bubbles—a technique that minimizes bias toward specific facial features and does not distort stimuli—to a set of 300 color images of Caucasian faces that were presented only once to 30 participants.

2.3. Methods

2.3.1. Participants

Thirty students from the University of Montreal and McGill University were recruited to participate to the experiment. Participants were between 20 and 30 years of age. They all had normal, or corrected to normal vision. Informed consent was obtained before the beginning of the experiment and a monetary compensation was provided.

2.3.2. Stimuli

Stimuli were generated from 300 color images of Caucasian faces (150 females), chosen on Internet with the intent of ecological representativity. The only other characteristics required for selection were a clear gender membership, a neutral expression and a frontal view. Thus, no special attention was paid to lighting, file format, image size, age of depicted individual, etc. Subsequent transformations applied on the images were also kept to a minimal. Rotations, scalings, and translations in the image plane were applied to the face photographs in order to

minimize the distance between handpicked landmarks around the eyes (4 landmarks each), the eyebrows (2 landmarks each), the nose (4 landmarks) and the mouth (4 landmarks). The average interpupil distance was 40 pixels (1.03 deg of visual angle). Note that these affine transformations do not modify the relative distances between features. Six instances of the resulting face images are shown on Figure 1a.

Insert Fig. 1 near here

Stimuli were created by sampling the face images subtending 3.28 deg of visual angle by presenting them behind an opaque mask punctured by an adjustable number of randomly located Gaussian apertures having a standard deviation of 4 pixels or 0.1 deg of visual angle (henceforth called ‘bubble mask’). The result, shown in Figure 1b, is a sparsely sampled face on a mid-gray background.

2.3.3. Apparatus

The experimental programs were run on a Macintosh G4 in the Matlab environment, using functions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). All stimuli were presented on a Sony Trinitron monitor (1024 X 768 pixels at a refresh rate of 85 Hz). We determined the relationship between RGB values and luminance levels (measured with a Samsung SyncMaster 753 df photometer) for each color channel independently; the three best-fitted “gamma” functions were used in the computation of image statistics. Participants were seated in a dim ambient-lighted room at a distance of approximately 75 cm from the computer monitor.

2.3.3. Procedure

Each participant was submitted to 300 trials and, importantly, each trial involved a different face. The presentation order of the 300 faces was randomized. In a given trial, one stimulus—a sparsely sampled face—appeared at the center of computer monitor and remained there until the participant had indicated the gender of the stimulus by pressing a labeled keyboard key. No feedback was provided. The number of bubbles per image was adjusted on a trial-by-trial basis to maintain performance at 75% correct using QUEST (Watson & Pelli, 1983).

2.4. Results and discussion

Participants used an average of 27.06 bubbles and responded correctly on 74.74% of the trials. The average response time was 1.63 sec. The correlation between response time and accuracy was -0.1216 ($p < 0.001$). There was a slight bias toward responding “man” (52.18% of the trials, $p < 0.01$) rather than “woman”. No difference was observed between female and male participants (51.58% and 52.72%, ns).

2.4.1. Linear classification image analysis

To uncover which facial cues led more often to accurate or faster correct gender discrimination, we performed two least-square multiple linear regressions: one between discrimination accuracies (predictive variable) and bubble masks (explanatory variable) and another between quartiles of response time on correct trials and bubble masks¹. The outcome of these regressions are two 128 by 128 planes of regression coefficients which we call classification images (Eckstein & Ahumada, 2002; Gosselin & Schyns, 2004)². To compute group statistics, we summed classification images across participants and smoothed the resulting group classification images with a Gaussian kernel having a standard deviation of 6.93 pixels.

The statistical analysis was restricted to the area of the classification images that could contain face information; the complementary area, which was irrelevant to the task at hand, was used to estimate the mean and the standard deviation of the null distribution and to transform the group classification images into Z-scores. Any significant positive local divergence from uniformity in our group classification images would indicate that the corresponding part of the stimuli led to more accurate responses, or faster correct responses. We therefore conducted one-tailed Pixel tests (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005) on the group classification images transformed into Z-scored ($Sr = 3469$; for accuracy: $Z_{crit} = 3.7$ and $Z_{max} = 6.48$; for response time: $Z_{crit} = 3.5$ and $Z_{max} = 4.04$; $p < .05$). The statistical threshold provided by this test corrects for multiple comparisons while taking the spatial correlation inherent to our technique into account.

Insert Fig. 2 near here

Figure 2 displays the average women (column 1) and men (column 2) overlaid with a contour-plot representation of the accuracy and correct response time classification images. The colored pixels enclosed by the dotted black lines are statistically significant: the region of the eyes and eyebrows lead to more accurate and faster correct gender discrimination; this eye-eyebrow region is wider and more bilaterally distributed in the correct response time classification image (row 1) than in the accuracy classification image (row 2); and facial cues leading to fast correct responses also included the mouth region as well as the space between the mouth and the nose. To better understand the relation between the mouth region and our measurements, we ran an additional least-square multiple linear regression between quartiles of response time on incorrect trials and bubble masks. No pixel was significant in the resulting classification image (not shown).

2.4.1. Beyond linear classification Images

The linear classification image analyses confirmed that the eye-eyebrow region contains the most important cues for gender discrimination. However, they do not allow to identify more precisely the nature of these reliable cues, at least not directly. For example, we could wonder if these cues are mostly red-green pigmentation cues, as proposed by Tarr and colleagues? It's not so much a limit of the methodology than a limit of the search space we chose to explore—image location. In fact, Nestor and Tarr (2008b) have used classification images to probe the use of color directly during gender discrimination. On each one of 20,000 trials, color noise was added to the same androgynous morph and participants had to decide whether it looked more like a man or a woman. If we cannot address the color question directly, we can provide—based on the 300 faces of our face set—image statistics about the discriminative color information that was available within the eye-eyebrow region.

We converted these face images to Lab color space because its channels represent perceptually relevant color opponent processes: L corresponds to the light-dark process, a to the red-green process, and b to the yellow-blue process. Then, we computed d' on each pixel of the three Lab channels—we will call the resulting d' planes color maps. This metric could be interpreted as the information available in a given pixel of a given color map to discriminate the gender. More specifically, a pixel's d' is the distance (in standard deviation units) between the mean of the distribution of this pixel's value for male faces and the mean of the distribution of this pixel's value for female faces. The three color maps are represented as contour plots in Figure 3. Color lines delimit isovalued d' corresponding to percentiles of 95%, 85%, and 75%. Warm colors were used for regions where men are lighter, redder or yellower than women; and cold colors for regions where men are darker, greener or bluer than women. To help with interpretation, the contour plots were placed over an image of the average men (column 1) and women (column 2). Thick dotted lines were added to delineate the significant regression coefficients found in the accuracy (white) and the correct response time classification images (black).

The light-dark color map depicts the information that has been mainly investigated in the literature so far. It shows the availability of prominent gender cues in the temporal side of the brows and the eyes, over the upper lip, and under the commissure of the chin and the lower lip

(Russell, 2003, 2005). Note also the luminance information located on the face outline near the cheeks. On average, this channel has higher d 's than the other color channels (mean d 's: light-dark = 0.36, red-green = 0.27, yellow-blue = 0.21). This set of informative features overlap substantially with the features found in the accuracy and correct response time classification images. The most informative pixels in the red-green color map—the second most informative color channel—are localized on the lips but are also distributed on the maxilla region and near the chin-lower lip commissure. The upper lip is a feature also found in the correct response time classification image. In comparison, the yellow-blue channel contains less information allowing to distinguish males from females. The most informative yellow-blue cues are clustered on the temporal sclera, on the nasal side of the brows and on the outer portion of the hair. None of these features is found in the classification images.

Another cue that has already been targeted as the one of the most discriminative information for gender categorization is the eyelid-brow distance (Burton et al., 1993; Campbell et al., 1999), i.e. the distance between the center of the upper eyelid and the center of the bottom part of the eyebrow. If the participants used this cue they needed to see part of the eye, the eyelid and the brow together. Therefore, the performance observed in the trials in which these regions were presented together (see Table 1, first row) should be higher than the performance predicted by the linear combination of these regions presented individually with the appropriate weights from the accuracy classification image (see Table 1, second row).

Insert Table 1 near here

In fact, predictions made from the accuracy linear regression explains the performance observed when the eye, the eyelid and the brow are seen together. Moreover, image statistics computed on the 300 faces from our database indicate that this relational cue provide little discriminative information: the d ' of the eye-eyelids distance—measured from handpicked landmarks—is 0.91. In sum, these results do not support the use of the eyelid-brow distance in

our experiment. Further analyses would be required to assess the use of other distance cues. However, Nestor and Tarr (2008a) performed a similar analysis on all pairwise conjunctions between the forehead, the eyes, the ears, the upper and lower part of the nose, the cheeks, the mouth, and the chin, and failed to find evidence for nonlinear use of information during their gender discrimination task.

2.4. Conclusion

Which face cues do we use for gender discrimination? In this paper, we addressed this question in a more realistic setting than previous studies on the same topic. First, the face stimuli that have been used typically in gender discrimination experiments were grayscale photographs, normalized and controlled for different aspects. Our results can be considered as more representative of genuine gender discrimination because our face stimuli were real-life color photographs and, therefore, were not (artificially) controlled for luminance, chrominance, background, hair and makeup. Second, previous studies on facial gender discrimination cues used a small set of faces that needed to be shown many times to each participant; therefore, the results might reflect the peculiarities of small stimulus sets overlearned by participants rather than general characteristics of gender dimorphism. We used a set of 300 face photographs that were presented only once to each of our 30 participants. Third, gender discrimination studies typically manipulated specific features and regions of the face with techniques such as morphing and caricaturing. These manipulations probably altered the natural characteristics of faces, and biased the results. We sampled unaltered face photographs with minimum bias by presenting them behind mid-gray opaque masks punctured by a number of randomly located Gaussian apertures sufficient to maintain a 75% correct response rate. This sampling technique makes no assumption regarding feature processing - holistic or not. A comparison of the power spectrum of the 300 face photographs and the 9,000 face stimuli presented to participants (with bubbles) revealed a slight reduction of energy below 2.63 cycles per face width. This bias is unlikely to have interfered with normal face processing (e.g., Ruiz-Soler & Beltran, 2006).

To uncover which facial cues led more often to accurate or faster correct gender discrimination, we performed three classification image analyses: on accuracies, on correct response times, and on incorrect response times. The accuracy classification image confirmed that the eye-eyebrow region is the most important for gender discrimination. We do not know if participants used facial features, or makeup, or trimmed eyebrows within this region to perform the task. In any case, this main result is in agreement with previous findings obtained using different methods (e.g., Brown & Perrett, 1993; Russell, 2003, 2005; Yamaguchi et al., 1995).

Linear predictions made on a subset of our trials showed that participants did not use the eyelid-brow distance information, a distance cue that Bruce et al. (1993) and Campbell et al. (1999) proposed was one of the most reliable for gender discrimination. In fact, we discovered that the eyelid-brow distance has a small signal-to-noise ratio for gender discrimination.

We computed image statistics on the Lab channels of all 300 faces of our face set to capture the color information available to resolve the task. Our color maps do not inform us about the morphological or spectral gender dimorphisms of the real world. Nevertheless, they show, for example, that the highly discriminative information contained in the eye-eyebrow area is mostly concentrated in the light-dark channel. This suggests that humans discriminate face gender based on a linear combination of luminance cues within the eye-eyebrow region. There is no inconsistency between our results and Tarr and colleagues' results about the important role of color in face gender discrimination (Nestor & Tarr, 2008a, 2008b; Tarr et al., 2001, 2002). They showed that participants relied on pigmentation cues (especially from the red-green channel) when minimal or no luminance information is available. Similarly, Yip & Sinha (2002) showed that color cues play a role in face identification when shape attributes are degraded. Yip and Sinha proposed that the contribution of color may lie not so much in providing diagnostic cues to identity as in aiding low-level image analysis processes such as segmentation; and the same could be proposed about face gender discrimination. Parametric models aiming at automatic segmentation of facial features also focus on color information for the extraction of the lips (Evano, Clapier, & Coulon, 2004).

That being said, the correct response time classification image along with the additional analysis on incorrect response time suggest a more ubiquitous role for facial color during gender

discrimination. The mouth region is significantly correlated with correct fast responses (but not with incorrect fast responses) and the most discriminative information in the mouth region is concentrated in the red-green channel. This suggests that humans do use chromatic cues for discriminating face gender: When it's informative, they use it and respond rapidly (for evidence that color is perceived faster than shape, see Holcombe & Cavanagh, 2001; Moutoussis & Zeki, 1997a, 1997b); when it's not, they have to rely on the more robust and more sluggish luminance cues. The infero-temporal cortex, which is involved in both face perception and color perception (Clark et al., 1997; Edwards, Xiao, Keysers, Foldiak, & Perrett, 2003), provides the ideal locus for such a dual strategy.

2.5. Footnotes

1. For the least-square multiple linear regression on accuracy, the computations reduce to subtracting the mean of the bubble masks that led to an incorrect response from the mean of the bubbles masks that led to a correct response. And, for the regression on response time, the computations reduce to summing 1.5 times the mean of the bubble masks that led to a correct response and to a response time in the fastest quartile, 0.5 times the mean of the bubble masks that led to a correct response and to a response time in the second quartile, -0.5 times the mean of the bubble masks that led to a correct response and to a response time in the third quartile, and -1.5 times the mean of the bubble masks that led to a correct response and to a response time in the slowest quartile. Prior to these computations, every bubble mask was transformed into z-scores to give equal weight to all bubble masks. See Chauvin et al. (2005) for technical details.

2. Bubbles and reverse correlation experiments (e.g., Sekuler et al., 2004) result typically in linear classification images. However, the two techniques should not be confused. In a Bubbles experiment, the stimuli are sampled using multiplicative noise (or bubble masks) whereas, in a reverse correlation experiment, the stimuli are masked using additive noise. This apparently minor procedural difference has important functional consequences (e.g., Gosselin & Schyns, 2002; Murray & Gold, 2004).

2.7 Acknowledgments

This research was supported by FQRNT scholarship awarded to Nicolas Dupuis-Roy and Isabelle Fortin, and by NSERC and NATEQ grants awarded to Frédéric Gosselin.

2.8 References

- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.
- Brown, E., & Perrett, D. I. (1993). What gives a face its gender. *Perception, 22*, 829–840.
- Bruce, V., Burton, A. M., Hanna, E., Healey, P., Mason, O., Coombes, A., et al. (1993). Sex-discrimination: How well do we tell the difference between male and female faces. *Perception, 22*, 131–152.
- Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognizing the sex and identities of faces. *Perception, 23*, 803–822.
- Burton, A. M., Bruce, V., & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception, 22*, 153–176.
- Campbell, R., Benson, P. J., Wallace, S. B., Doesbergh, S., & Coleman, M. (1999). More about brows: How poses that change brow position affect perceptions of gender. *Perception, 28*, 489-504.
- Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision, 5(9):1*, 659–667, <http://journalofvision.org/5/9/1/>, doi:10.1167/5.9.1.
- Clark, V. P., Parasuraman, R., Keil, K., Kulansky, R., Fannon, S., Maisog, J. M., et al. (1997). Selective attention to face identity and color studied with fMRI. *Human Brain Mapping, 5*, 293–297.
- Eckstein, M. P., & Ahumada, A. J., Jr. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision, 2(1):i*, i, <http://journalofvision.org/2/1/i/>, doi:10.1167/2.1.i.
- Edwards, R., Xiao, D., Keysers, C., Foldiak, P., & Perrett, D. (2003). Color sensitivity of cells responsive to complex stimuli in the temporal cortex. *Journal of Neurophysiology, 90*, 1245–1256.
- Evano, N., Clapier, A., & Coulon, P.-Y. (2004). Accurate and quasi-automatic lip tracking. *IEEE Transactions on Circuits and Systems for Video Technology, 14*, 706–715.
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*, 2261–2271.
- Gosselin, F., & Schyns, P. G. (2002). RAP: A new framework for visual categorization. *Trends in Cognitive Science, 6*, 70–77.

- Gosselin, F., & Schyns, P. G. (Eds.) (2004). Rendering the use of visual information from spiking neurons to recognition. *Cognitive Science*, *28*, 141–301.
- Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: The role of shape and colour. *Proceedings of the Royal Society B: Biological Sciences*, *261*, 367–373.
- Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nature Neuroscience*, *4*, 127–128.
- Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society of London B: Biological Sciences*, *264*, 393–399.
- Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London B: Biological Sciences*, *264*, 1407–1414.
- Murray, R. F., & Gold, J. M. (2004). Troubles with bubbles. *Vision Research*, *44*, 461–470.
- Nestor, A., & Tarr, M. J. (2008a). The segmental structure of faces and its use in gender recognition. *Journal of Vision*, *8(7):7*, 1–12, <http://journalofvision.org/8/7/7/>, doi:10.1167/8.7.7.
- Nestor, A., & Tarr, M. J. (2008b). Gender recognition of human faces using color. *Psychological Science*, *19*, 1242–1246.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychological Research*, *70*, 273–292.
- Russell, R. (2003). Sex, beauty, and the relative luminance of facial features. *Perception*, *32*, 1093–1107.
- Russell, R. (2005). Face pigmentation and sex classification [Abstract]. *Journal of Vision*, *5(8):983*, 983a, <http://journalofvision.org/5/8/983/>, doi:10.1167/5.8.983.
- Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*, 402–409.
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, *14*, 391–396.
- Tarr, M. J., Kersten, D., Cheng, Y., & Rossion, B. (2001). It's Pat! Sexing faces using only red and green [Abstract]. *Journal of Vision*, *1(3):337*, 337a, <http://journalofvision.org/1/3/337/>,

doi:10.1167/1.3.337.

- Tarr, M. J., Rossion, B., & Doerschner, K. (2002). Men are from Mars, women are from Venus: Behavioral and neural correlates of face sexing using color [Abstract]. *Journal of Vision*, 2(7):598, 598a, <http://journalofvision.org/2/7/598/>, doi:10.1167/2.7.598.
- Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33, 113–120.
- Yamaguchi, M. K., Hirukawa, T., & Kanazawa, S.(1995). Judgment of gender through facial parts. *Perception*, 24, 563–575.
- Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, 31, 995–1003.

2.9. Figure and tables

Figure 1. (a) Three women and three men from our face database; and the average of all 150 women and 150 men from our face database. (b) A stimulus is generated by overlaying an opaque mid-gray mask punctured by a number of randomly located Gaussian apertures on a face.

Figure 2. Displays the average men (column 2) and women (column 1) superimposed on a contour plot of classification images derived from accuracy (row 1) and response time (row 2). The colored pixels enclosed by the dotted black lines are statistically significant ($p < .05$).

Figure 3. Contour plots of the color maps superimposed to the average man (column 1) and woman (column 2). Dotted lines define clusters significantly correlated with accurate (white) and correct fast responses (black). The contour plot summarizes the spatial modulation of available information (d' 's) in the dark-light (row 1), red-green channel (row 2) and yellow-blue (row 3) channels. The color-labeled lines of isovalued d' 's correspond to percentile 95%, 85%, 75%, 25%, 15%, and 5%.

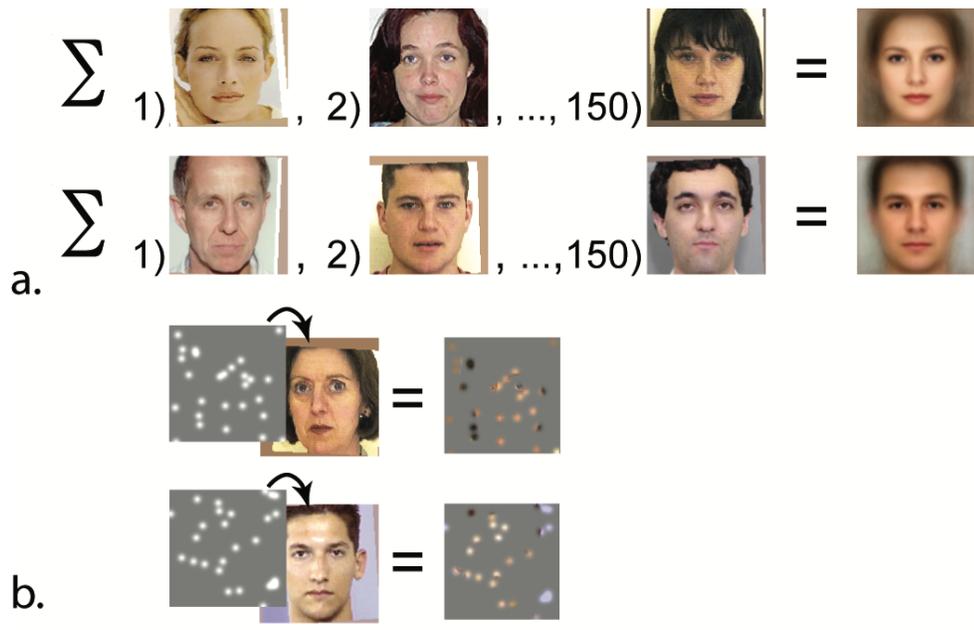


Figure 1

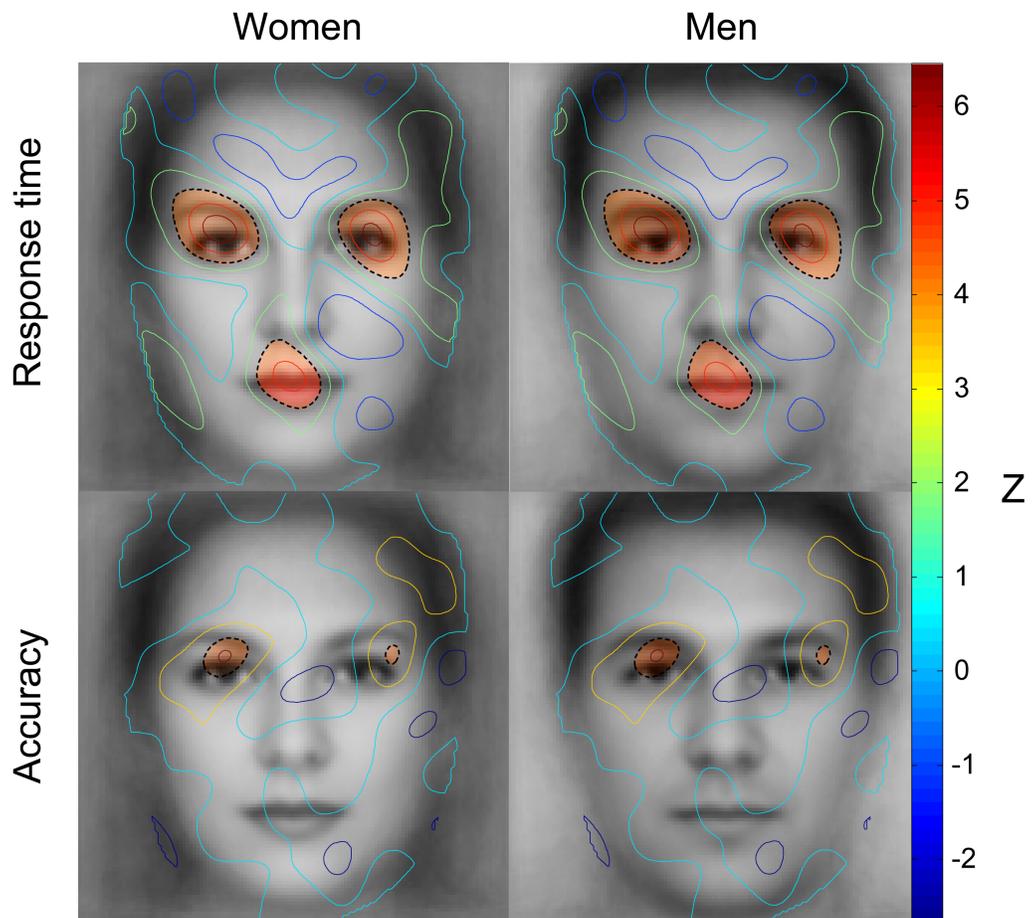


Figure 2

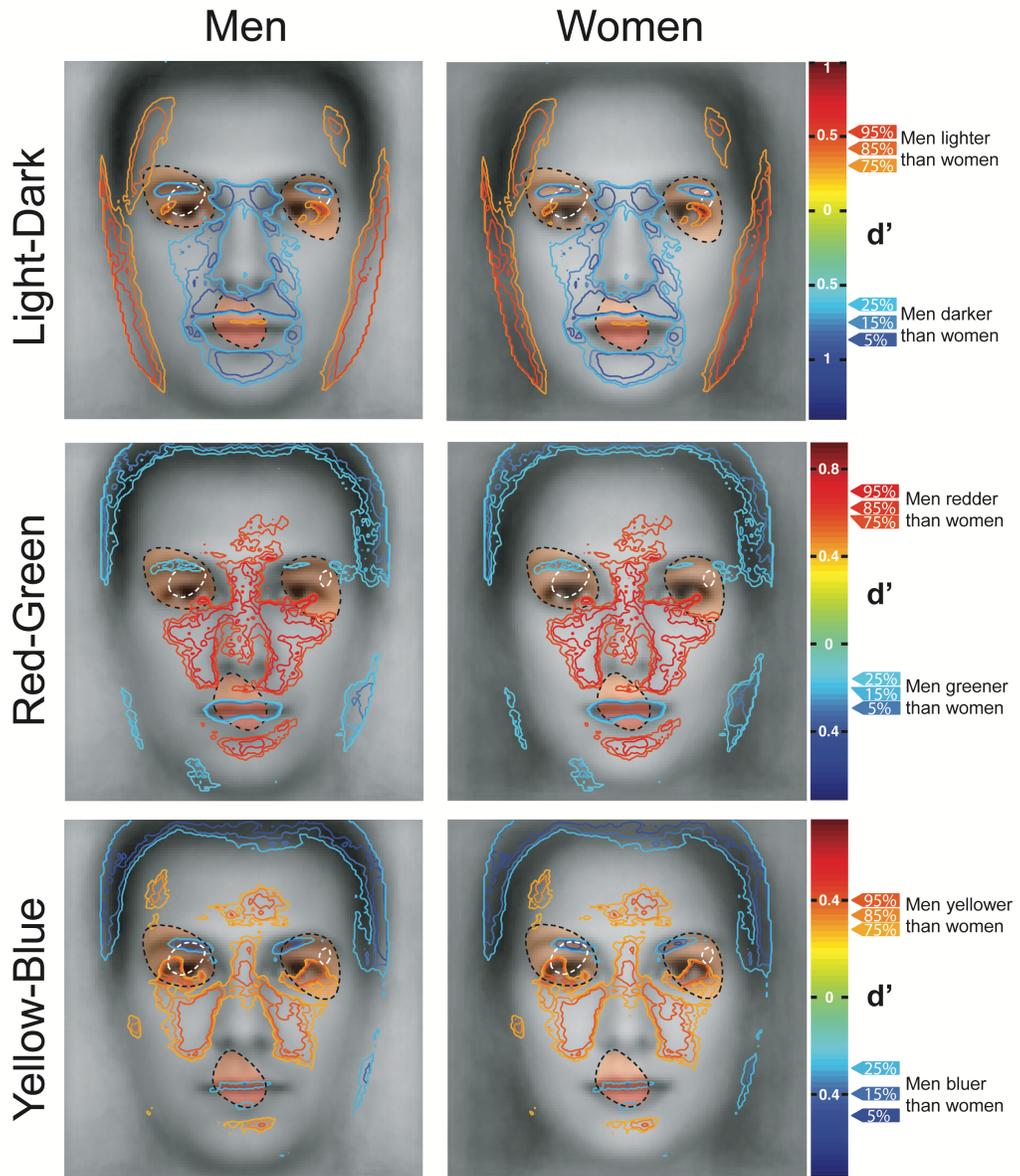


Figure 3

Table 1. The first row shows the mean accuracy observed when areas are revealed separately (columns 1–3) and together (column 4). The second row indicates the average accuracy predicted from linear regression. The last row displays the number of trials that were used to compute these statistics.

	Eyes	Brows	Eyelids	Eyes, brows and eyelids
Observed accuracy	0.7216	0.7329	0.6897	0.7626
Predicted Accuracy	0.6082	0.7471	0.6612	0.8216
N	194	87	307	269

Chapitre 3: Article 2

Real-World Interattribute Distances Lead to Inefficient Face Gender Categorization

Nicolas Dupuis-Roy¹, Daniel Fiset², Kim Dufresne¹, Laurent Caplette¹ & Frédéric Gosselin¹

1. Département de psychologie, Université de Montréal.
2. Département de psychoéducation et de psychologie, Université du Québec en Outaouais

Address correspondence to

Nicolas Dupuis-Roy, Département de psychologie, Université de Montréal, PO Box 6128, Succ. Centre-Ville, Montréal, QC H3C 3J7, Canada.

Article published in *Journal of Experimental Psychology: Human Perception and Performance* (2014), 40 (4), 1289-1294.

3.1. Abstract

The processing of interattribute distances is believed to be critical for upright face categorization. A recent study by Taschereau-Dumouchel et al. (2010) challenged this idea by showing that participants were nearly at chance when asked to identify faces on the sole basis of real-world interattribute distances, while they were nearly perfect when all other facial cues were shown. However, it remains possible that humans are highly tuned to interattribute distances but that the information conveyed by these cues is scarce. We tested this hypothesis by contrasting the efficiencies—a measure of performance that factors out task difficulty—of 60 observers in six face-gender categorization tasks. Most importantly, efficiencies for faces that varied only in terms of their interattribute distances were an order of magnitude lower than efficiencies for faces that varied in all respects except their interattribute distances or in all respects. These results provide a definitive blow to the idea that real-world interattribute distances are critical for upright face processing.

Keywords: Face-gender categorization, Efficiency, Ideal observer analysis, Chromatic and achromatic cues, Information integration

3.2. Introduction

Francis Galton was one of the first to propose that spatial relations between the main facial attributes (i.e. nose, mouth, eyes and eyebrows) was a reliable distant-invariant cue that could be used for face recognition (see Galton, 1879). Galton even advocated for the use of relational cues as a simple system to certify the identity of a person— just as he did for fingerprints (Galton, 1888). Nearly a century later, Haig (1984) investigated the impact of single attribute displacement on the recognition of unfamiliar faces. He found that the sensitivity to some displacement (e.g. mouth-up) approached the limit of visual acuity and concluded that humans are highly tuned to spatial relations in faces.

Studies of face inversion have also contributed to the idea that relative distances between attributes are fundamental for face recognition. Faces rotated by 180° in the picture-plane lead to lower recognition performances and higher response latencies (e.g., Hochberg and Galper, 1967). Because this impaired performance is disproportionately larger for faces than for other mono-oriented objects such as houses (Yin, 1969), researchers have employed face inversion as a tool to isolate what is special about upright face processing. It was discovered that processing of interattribute distances¹ (or IADs) is affected more by inversion than the processing of the local shape or surface-based properties of attributes (Barton et al., 2001; Freire et al., 2000; Le Grand et al., 2001; Rhodes et al., 2007; Sergent, 1984; for recent reviews, see Rossion, 2008, 2009, and see for opposite claim McKone & Yovel, 2009).

A recent study by Taschereau-Dumouchel et al. (2010) challenged the hypothesis that IADs are important for *real-world* face recognition. Taschereau-Dumouchel and colleagues used a sample of 515 face photographs to estimate the face recognition information available in real-world IADs. They found that the IADs from 86 face stimuli used in 14 previous studies on IADs studies had exaggerated this information up to a factor of 376% compared to real-world IADs. Furthermore, they showed that when human observers resolve a matching-to-sample (ABX) face identification task solely on the basis of real-world IADs, they perform poorly (<65%) across a broad range of viewing distances. In contrast, recognition is almost perfect when observers

recognize faces on the basis of real-world information other than interattribute distances such as attribute shapes and skin properties.

The Taschereau-Dumouchel et al. (2010) results raise serious doubts about the importance of IADs for face recognition but they fail to provide a definitive blow to the idea. A low performance with real-world IADs could be due to two possibly interacting causes: (1) scarce information could be available in real-world IADs to resolve face categorization or (2) observers might be inapt at using this real-world IADs information. Therefore, it remains possible that real-world IADs are important to face processing in the sense that although little real-world IADs information could be available, observers may exploit a high proportion of this information. In the present study, we tested this hypothesis for the first time in the context of face-gender categorization tasks, i.e. one of the most biologically relevant facial judgments.

To test our hypothesis, we used the efficiency measure, i.e. a performance index that factors out task difficulty (Tanner & Birdsall, 1958; Tjan et al., 1995). Efficiency can be expressed as the ratio of the quantity of noise required by a human observer to reach a given performance to the quantity of noise required by the ideal observer to reach the same performance. The ideal observer is a mathematical model that uses optimally all of the information available for the task at hand. Thus the ideal observer provides a benchmark for the highest possible performance, and the efficiency measure offers a grasp on the ability of humans to correctly exploit the available information. Within this framework, the hypothesized importance of IADs for upright face-gender processing should translate into high efficiency. In other words, human observers should exploit the most out of the real-world IADs information. The corollary of this hypothesis is that the efficiency for IADs should be relatively high compared to other face-gender cues. This was examined by contrasting the efficiency in three interattribute-distance conditions: the IADs condition, in which faces varied only in terms of their IADs; the noIAD condition, in which faces varied in all respects except their IADs; and the ALL condition, which featured the original and unmodified faces. Finally, we explored possible interactions between color and efficiencies by contrasting two *color* conditions: the LUM condition, which only exhibited the luminance of faces; and the COL condition, which showed the luminance and chrominance of faces.

3.3. Methods

3.3.1 Participants

Sixty healthy participants (30 women; $M=22.49$ years old, $SD=3.65$), with normal color vision and normal or corrected-to-normal visual acuity were recruited. An informed consent was obtained prior to the experiment, and a monetary compensation was provided upon its completion.

3.3.2 Stimuli

We used the Taschereau-Dumouchel et al. (2010) face database, which comprises 515 Caucasian frontal-view color portraits (256 females) presenting a neutral expression. The six main attributes of these faces (eyes, eyebrows, nose and mouth) were manually segmented; these attributes were aligned across faces using a linear conformal transformation, which preserves IADs.

The experiment comprised six stimulus sets: (LUM vs. COL) x (IADs vs. noIAD vs. ALL). Figure 1a depicts sample stimuli from each of these tasks.

3.3.2.1 IADs stimuli

Each of the 515 stimuli of this task was created as follows: First, the six facial attributes of all the faces of the database were displaced via cubic interpolation to the real-world attribute locations of one face of the database. Second, two gender prototypes were created by dot-averaging all faces belonging to the same gender. Third, the two gender prototypes were blended according to each subject point of subjective equality (see *Procedure*). Note that we measured one point of subjective equality for COL stimuli and another for LUM stimuli.

3.3.2.2 noIAD stimuli

The 515 face stimuli used in this task were created by displacing the six facial attributes of every face of the database via cubic interpolation to their average locations.

3.3.2.3 ALL stimuli

The 515 original faces of the database were used in this task.

3.3.3. Apparatus

The experimental programs were run on four computers in the Matlab (Mathworks Inc.) environment, using functions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The four high-resolution CRT monitors used for stimuli presentation were set to display 1024 by 768 pixels at a refresh rate of 85 Hz. The relationship between RGB values and luminance levels (measured with a Samsung SyncMaster 753 df photometer) was computed for each color channel independently; the three best-fitted gamma functions were then used to create luminance noise and analyze the luminance content of the stimuli. Participants were seated in a dim-lighted room. Their viewing distance was maintained by a chin-rest so that the stimuli interocular width was 1.05 deg of visual angle.

3.3.4. Signal and noise adjustments

The computation of efficiency measure requires the addition of visual noise to the stimuli to ensure that the ideal observer makes errors. Although using different external noise levels across conditions is not usually recommended, we were forced to do so in the present case due to the large discrepancy between the sensitivities in the IADs tasks and the other tasks. The noise and signal contrast levels were adjusted once and for all in a series of preliminary experiments (see Supplementary Material). We added luminance Gaussian noise to the LUM stimuli while we added luminance Gaussian noise independently to each of the RGB channels of the COL stimuli. The root mean square contrast of the face images and the power of the noise (expressed in contrast units) were, respectively, 27.58 and 7.32×10^{-3} in the IADs-COL condition; 69.73 and 1.08×10^{-3} in the IADs-LUM condition; 26.63 and 107.04×10^{-3} in the noIAD-COL condition; and

61.38 and 2.99×10^{-3} in the noIAD-LUM condition. For the ALL-COL and ALL-LUM conditions, we used the same signal contrast and noise power as in the noIAD-COL and noIAD-LUM conditions respectively.

3.3.5. Procedure

Subjects completed the experiment on two separate days. On day 1, they first did two preliminary tasks and, then, two of the six experimental conditions (15-20 min each). On day 2, they did the remaining experimental conditions. The order of the experimental conditions was randomized across participants.

In the preliminary tasks, each participant judged five blends of the gender prototypes (from 80% male-20% female to 20% male-80% female) 20 times on a male-to-female continuous scale. There were two preliminary tasks: one displaying LUM stimuli, and the other COL stimuli. The points of subjective equality were interpolated from fitted psychophysical curves, and were used to create subject-specific stimuli in the IADs tasks (see Stimuli—IADs stimuli).

Each experimental condition consisted in the unique presentation of the 512 stimuli created for that condition in a random order. Each trial was constructed as follows: A black fixation cross was shown for 750 ms on a gray background; a uniform gray screen was displayed for 250 ms; and finally, a noisy face stimulus revealed through an ellipse was presented at the center of a uniform gray screen and remained there until the participant had indicated its gender by pressing the appropriate keyboard key. No feedback was provided.

3.4. Results

Statistical analyses were conducted on 53 subjects (25 men): two subjects were excluded during the experiment because they were complaining of visual fatigue, and five others were rejected because their performance was at least three standard deviations below the group average in at least one experimental condition.

The average performance in the IADs-COL and IADs-LUM conditions—the most difficult ones—was above chance level (56.63%, $t(52)=16.51$, $p<0.001$; 56.08%, $t(52)=15.55$, $p<0.001$). Binomial tests performed on individual performance revealed that 34 participants reached significance in the IADs-COL condition and 31 in the IADs-LUM condition (see Supplementary Material, Table 1).

Insert Figure 1 about here

3.4.1. Efficiencies

The ideal observer maximizes the *a posteriori* probability $P(G_i|S)$ of selecting the proper gender, say G_i , when a face exemplar, say S , is embedded in Gaussian noise. Tjan et al. (1995) demonstrated that this ideal decision rule is equivalent to selecting the gender category i that maximizes the following term:

$$\text{Eq. (1)} \quad \sum_{jk} \exp\left(-\frac{1}{2\sigma^2} \|S - T_{ijk}\|^2\right),$$

the weighted sum of the square of the euclidean distance between the noisy face stimulus (S) and the face template (T_{ijk} —the j^{th} face exemplar of gender i with k independent chromatic channels). The ideal observer went through the same six tasks human observers did, except that each stimulus was presented 10 times at 10 different noise levels. The noise power needed by the ideal observer to reach a given subject's sensitivity index (d') was interpolated from power curves fitted to the average d' recorded at each noise level. The R-squares of the fitted curves were high, ranging from 96% to 100% ($M=98.5\%$).

Finally, an efficiency measure was computed following Tanner and Birdsall's (1958) formulation:

$$\text{Eq. (2)} \quad \eta = \frac{N_h}{N_i},$$

where N_h is the noise power required by a human observer to reach a given performance, and N_i is the noise power required by the ideal observer to reach the same performance. Figure 1b shows the individual and group efficiencies for the different conditions.

A MANOVA indicated two significant main effects (interattribute-distance conditions: Wilks' $\lambda=.09$, $F(2,50)=261.75$, $p<.001$; color conditions: Wilks' $\lambda=0.9$, $F(2,50)=494.32$, $p<.001$) and a significant interaction (Wilks' $\lambda=.09$, $F(2,50)=257.10$, $p<.001$). The interaction was decomposed into simple effects: the interattribute-distance conditions contained significant differences for both COL (Wilks' $\lambda=.09$, $F(2,50)=259.51$, $p<.001$) and LUM (Wilks' $\lambda=.15$, $F(2,50)=146.44$, $p<.001$). Pairwise comparisons indicated that the efficiencies for noIAD-COL ($M=13.4\%$) and ALL-COL ($M=13.2\%$) were significantly higher than the efficiency for IADs-COL ($M=0.276\%$; $p<.001$, Bonferroni-corrected). However, no significant difference was found between noIAD-COL and ALL-COL conditions. Similarly, the efficiencies for noIAD-LUM ($M=0.338\%$) and ALL-LUM ($M=0.255\%$) were significantly higher than the efficiency for IADs-LUM ($M=0.152\%$). The efficiency for noIAD-LUM was also significantly higher than for ALL-LUM ($p<.001$).

To estimate the available information in the different conditions, we computed the ratio of contrast energy and threshold noise power required for the ideal observer to reach a d' of 1.60, i.e. the average of all observed d' in the experiment: 1.26×10^5 (IADs-COL), 7.87×10^3 (IADs-LUM), 97.07 (noIAD-COL), 164.29 (noIAD-LUM), 91.51 (ALL-COL) and 151.10 (ALL-LUM).

3.5. Discussion

The idea that the processing of spatial relations between the main internal features of faces (e.g., nose, mouth, eyes and eyebrows) is a distinctive and critical aspect of upright face recognition is widely accepted in the face perception literature (e.g., Maurer et al., 2002). A recent study by Taschereau-Dumouchel et al. (2010) challenged this idea by showing that participants were nearly at chance level when asked to identify faces on the sole basis of real-world IADs, and that they were nearly perfect when all other facial cues were shown while IADs were kept constant across faces. However, a low performance with real-world IADs could be due

to two possibly interacting causes: (1) the information to resolve the task might be scarce, or (2) observers might be inapt at using the available information. Therefore, real-world IADs could be important for face processing inasmuch as observers exploit a high proportion of their meager information. We compared the efficiency—a performance index that factors out task difficulty—for face-gender categorization in six conditions: three *interattribute-distance* conditions (IADs: faces varied only in terms of their IADs; noIAD: faces varied in all respects except their IADs; and ALL: original faces) x two *color* conditions (LUM: only luminance of faces; COL: luminance and chrominance of faces).

The average accuracy recorded in the IADs-COL and IADs-LUM conditions was just above chance level, which confirms the main result of Taschereau-Dumouchel et al. (2010). More importantly, however, we found low efficiencies ($\sim 0.2\%$) in the IADs conditions, indicating that participants were poor at using real-world IADs. In other words, humans appear unable to exploit a high proportion of the real-world IAD information to categorize facial gender even when it is the only available information. Furthermore, the efficiencies found in the ALL conditions were much higher (i.e. $\sim 6.8\%$) and comparable to those reported for other face recognition tasks (e.g. Gold et al., 1999; Gold et al., 2013; Hammal, Gosselin & Fortin, 2009).

We also explored the impact of color on face-gender discrimination efficiencies and found significantly higher efficiencies for COL than for LUM faces, regardless of the interattribute-distance condition. This is consistent with studies on face-gender discrimination showing that human observers are sensitive to chromatic differences between men and women's facial pigmentation, and that they include chroma in their representation of the gender of a face (see Dupuis-Roy et al., 2009; Nestor & Tarr, 2008a, 2008b; Tarr et al. 2001, 2002; but see Bruce et al., 1999). In addition to the color face-gender information, two mid-level vision mechanisms might benefit from color and, in turn, improve gender discrimination. First, it has been proposed that color can enhance the parsing of a face into its constituent parts via surface segmentation and edge localization (see Yip & Sinha, 2002). Second, it is known that humans' contrast sensitivity to low spatial frequencies is higher for chromatic than achromatic variations (Mullen, 1985). Therefore, it is possible that color enhances the signal in the spatial frequency channels that are the most relevant to face-gender categorization. This hypothesis is also consistent with Yip and

Sinha (2002) who observed that the facilitating role of color in a face individuation task was further amplified as faces were increasingly low-passed (see also Tarr et al., 2001). In any case, our findings provide the first direct demonstration that the gains associated with chromatic variations in the face also lead to higher efficiencies.

We believe that the low efficiency for processing IADs could originate from multiple cue integration processes in the brain. Studies on the integration of multiple cues within a single sensory modality (e.g., Hillis et al., 2004; Young et al., 1993) and between sensory modalities (e.g., Alais & Burr, 2004; Battaglia et al., 2003; Ernst & Banks, 2002) have shown that humans tend to weigh multiple cues quasi-optimally as a function of their reliability. If humans also weigh the multiple facial cues quasi-optimally given the information they convey, then the scarce information conveyed by IADs should be largely underweighted compared to all the information contained in other local cues such as the 2D shape of the features, the shape-from-shading cues, the skin texture and pigmentation IADs cues. This multiple cue integration hypothesis also predicts that these weights could change as a function of their reliability. Interestingly, this might help explain the results of previous studies stating the importance of IADs for upright face recognition. Taschereau-Dumouchel et al. (2010) surveyed 14 IADs studies on face identification (Barton et al., 2001; Bhatt et al., 2005; Freire et al., 2000; Goffaux et al., 2005; Haig, 1984; Hayden et al., 2007; Hosie et al., 1988; Leder & Bruce, 1998, 2000; Leder et al., 2001; Le Grand et al., 2001; Rhodes et al., 1993; Sergent, 1984; Tanaka & Sengco, 1997) and found that most of these had greatly exaggerated the signal in their IADs, up to a factor of 376% compared to real-world IADs. By amplifying the signal in IADs, these studies might have artificially promoted their use. Altogether, these findings provide a definitive blow to the idea that real-world IADs are critical for face recognition mechanisms in the real-world.

3.6. References

- Alais, D., & Burr, D. (2004) The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257-262.
- Barton, J. J. S., Keenan, J. P., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. *British Journal of Psychology*, *92*, 527-

- Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, *20*, 1391-1397.
- Bhatt, R. S., Bertin, E., Hayden, A., & Reed, A. (2005). Face processing in infancy: Developmental changes in the use of different kinds of relational information. *Child Development*, *76*(1), 169-181.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433-436.
- Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, *5*(4), 339.
- Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision*, *9*(2), 1-8.
- Ernst, M. O. & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429-433.
- Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception*, *29*(2), 159-170.
- Galton, F. (1879). Composite portraits, made by combining those of many different persons into a single, resultant figure. *Journal of the Anthropological Institute*, *8*, 132-144.
- Galton, F. (1888). Personal identification and description. *Nature*, *38*, 173-202.
- Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & 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:10.1068/p5370

Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature*, 402, 176-178.

Gold, J.M., Barker, J.D., Barr, S., Bittner, J.L., Bromfield, W.D., Chu, N., Goode, R.A., Lee, D., Simmons, M. & Srinath, A. (2013). The efficiency of dynamic and static facial expression recognition. *Journal of Vision*, 13(5): 23; doi:10.1167/13.5.23

Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception*, 13, 104-109.

Hammal, Z., Gosselin, F. & Fortin, I. (2009). How efficient are the recognition of dynamic and static facial expressions? *Journal of Vision*, 9(8): 499; doi:10.1167/9.8.499

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., & Joseph, J. E. (2007). The development of expert face processing: are infants sensitive to normal differences in second-order relational information? *Journal of Experimental Child Psychology*, 97(2), 85-98.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, 4, 967–992.

Hochberg J., & Galper R. E. (1967). Recognition of faces: I. An exploratory study. *Psychonomic Science*, 9, 619–620.

Hosie, J. A., Ellis, H. D., Haig, N. D. (1988). The effect of feature displacement on the perception of well-known faces, *Perception*, 17, 461- 474

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *The Quarterly Journal of Experimental Psychology: Section A*, 51(3), 449-473.

- Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology: Section A*, 53(2), 513–536.
- Leder, H., Candrian, G., & Huber, O. (2001). Configural features in the context of upright and inverted faces. *Perception*, 30, 73-83.
- Le Grand, R., Mondloch, C. J., Maurer, D., Brent, H. P. (2001). Early visual experience and face processing. *Nature*, 410, 890.
- Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255-260.
- 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
- Mullen, K. T. (1985). The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *The Journal of Physiology*, 359(1), 381–400.
- Nestor, A., & Tarr, M. J. (2008a). Gender recognition of human faces using color. *Psychological Science*, 19(12), 1242–1246.
- Nestor, A., & Tarr, M. J. (2008b). The segmental structure of faces and its use in gender recognition. *Journal of Vision*, 8(7):7, 1-12.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437-442.

- Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, *47*(1), 25-57.
- Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, *103*(1), 34-79.
- Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, *128*(2), 274-289.
- Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta psychologica*, *132*(3), 300-312.
- Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, *75*, 221-242.
- Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*(5), 583-604.
- Tanner, W. P., & Birdsall, T. G. (1958). Definitions of d' and η as psychophysical measures. *The Journal of the Acoustical Society of America*, *30*(10), 922-928.
- Taschereau-Dumouchel, V., Rossion, B., Schyns, P. G., & Gosselin, F. (2010). Interattribute Distances do not Represent the Identity of Real World Faces. *Frontiers in psychology*, *1*:159, 1-10. doi:10.3389/fpsyg.2010.00159
- Tarr, M. J., Kersten, D., Cheng, Y., & Rossion, B. (2001). It's Pat! Sexing faces using only red and green [Abstract]. *Journal of Vision*, *1*(3):337, 337a, <http://journalofvision.org/1/3/337/>, doi:10.1167/1.3.337.
- Tarr, M. J., Rossion, B., & Doerschner, K. (2002). Men are from Mars, women are from Venus: Behavioral and neural correlates of face sexing using color. *Journal of Vision*, *2*(7):598,

598a, [http:// journalofvision.org/2/7/598/](http://journalofvision.org/2/7/598/), doi:10.1167/2.7.598.

Tjan, B.S., Braje, W.L., Legge, G.E. & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise, *Vision Research*, 35 (21), 3053-3069.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Attention, Perception, & Psychophysics*, 33(2), 113–120.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of experimental psychology*, 81(1), 141-145.

Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, 31, 995-1003.

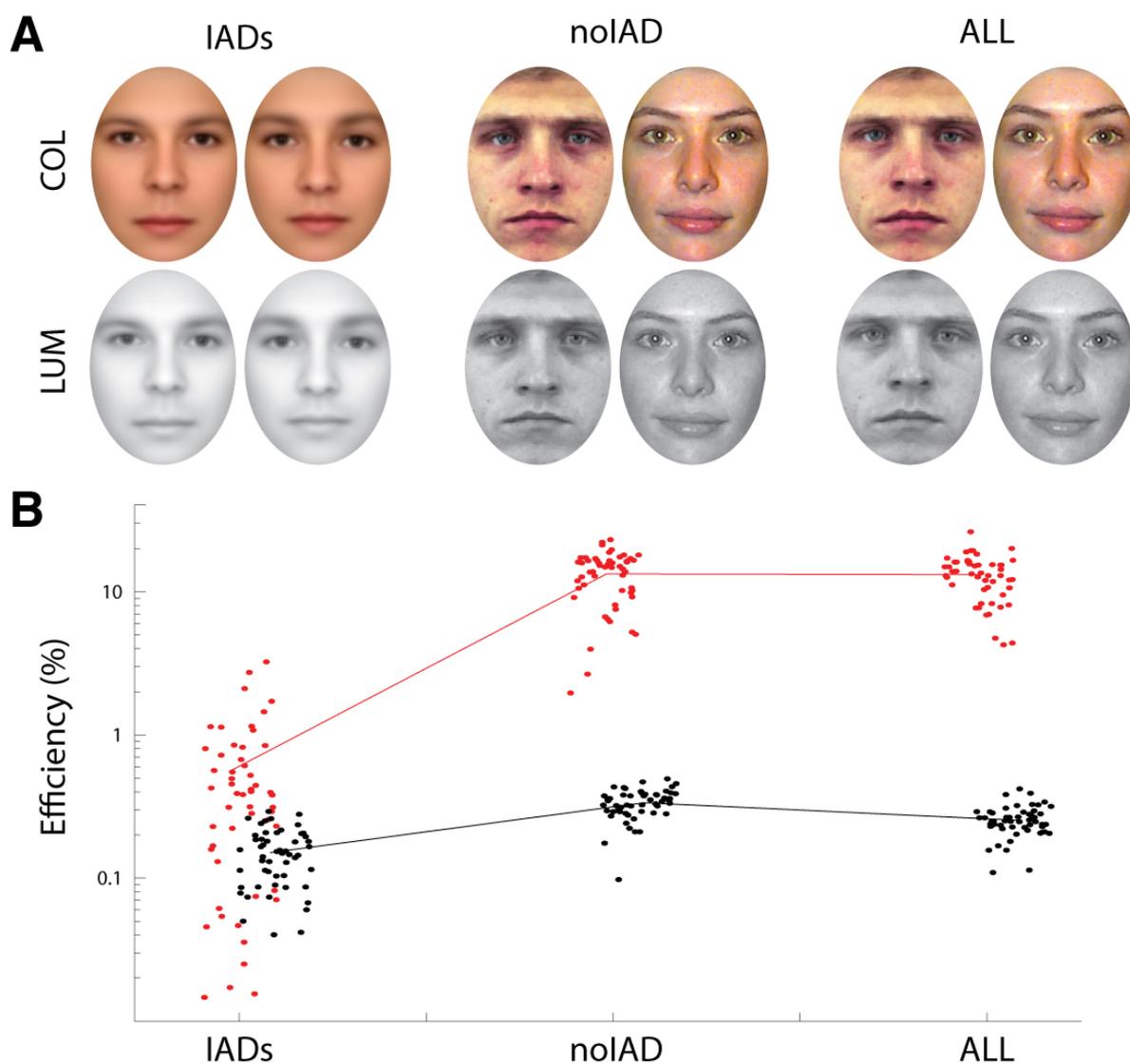
Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, 33, 2685–2696.

3.7. Footnotes

1. By “interattribute distances”, we mean relative distances between facial attributes that can be manipulated independently from the shapes of these attributes (e.g., the center of gravity to center of gravity interocular distance; e.g., Haig, 1984; Sergent, 1984; Hosie et al., 1988; Rhodes et al., 1993; Tanaka and Sengco, 1997; Leder and Bruce, 1998; Freire et al., 2000; Leder and Bruce, 2000; Barton et al., 2001; Leder et al., 2001; Le Grand et al., 2001; Bhatt et al., 2005; Goffaux et al., 2005; Hayden et al., 2007). This excludes, for example, the nasal-corner-to-nasal-corner interocular distance and the temporal-corner-to-temporal-corner interocular distance that cannot be manipulated jointly and independently from attribute size.

3.8. Figure

Figure 1. A) Example of masculine (left), feminine (right), COL (up) and LUM (bottom) face images used in the IADs, noIAD and ALL tasks. All images of a given gender correspond to a single face identity. B) Average efficiency recorded in the six tasks: the color condition is represented by the color of the line (red for COL and black for LUM). Red and black dots represent individual efficiencies.



3.9. Acknowledgements

This research was supported by a scholarship from the Social Sciences and Humanities Research Council awarded to Nicolas Dupuis-Roy, and by a grant from the Natural Sciences and Engineering Research Council of Canada awarded to Frédéric Gosselin.

3.10. Supplementary Material

3.10.1 Preliminary experiments for signal and noise adjustments

The noise and signal contrast levels were adjusted once and for all in a series of preliminary experiments. Two of the authors (L.C. and N.D.R) and another lab member first participated to a pilot experiment aiming at measuring the performance on the full-contrast noise-free stimuli in the six conditions. The results indicated a floor effect in the IADs-COL and IADs-LUM conditions (near-chance performances), and a ceiling effect in the ALL-COL and ALL-LUM conditions. The computation of efficiency measure requires the addition of visual noise to the stimuli to ensure that the ideal observer makes errors. Given the floor effect recorded in the noise-free IADs condition, we had no other choice but to use a minimum amount of noise and a maximum amount of signal. More specifically, we added luminance Gaussian noise to the LUM stimuli while we added luminance Gaussian noise independently to each of the RGB channels of the COL stimuli. The root mean square contrast of the face images and the power of the noise (expressed in contrast units) were, respectively, 27.58 and 7.32×10^{-3} in the IAD-COL condition, and 69.73 and 1.08×10^{-3} in the IAD-LUM condition. In the ALL-COL and ALL-LUM conditions, we wanted to prevent a possible ceiling effect. Although using different external noise levels across conditions is not usually recommended, we were forced to do it in the present case. We thus added a fixed amount of noise and adjusted the signal to reach 85% of correct responses. Assuming that face processing is unaffected by these external noise variations, this creates no theoretical difficulty for the computation of efficiencies and for their comparison. Another pilot experiment was run on three of the authors (F.G., L.C and N.D.R) and another lab member. The quantity of noise power was selected prior to the experiment and the signal contrast was adjusted online using QUEST (Watson & Pelli, 1983). The signal level chosen for each task corresponded to the mean of the last 30 trials averaged over the four participants. The root mean square contrast of the face images and the power of the noise were respectively, 26.63 and 107.04×10^{-3} in the ALL-COL condition, and 61.38 and 2.99×10^{-3} in the ALL-LUM condition. For the noIAD-

COL and noIAD-LUM conditions, we used the same signal contrast and noise power as in the ALL-COL and ALL-LUM conditions respectively.

Table 1. Individual sensitivities (d') observed in each condition.

Participant	IADs-COL	IADs-LUM	noIAD-COL	noIAD-LUM	ALL-COL	ALL-LUM
1	0.41	0.44	2.09	2.7	2.13	2.75
2	0.42	0.41	1.99	2.58	2.12	2.49
3	0.35	0.48	1.61	2.87	1.56	2.58
4	0.47	0.21	2.06	2.43	1.99	2.22
5	0.09	0.19	2.06	2.54	2.14	2.41
6	0.09	0.14	1.25	1.14	1.2	1.3
7	0.28	0.39	1.71	2.51	1.68	2.27
8	0.38	0.27	2.22	3.13	2.35	2.72
9	0.73	0.55	2.07	2.89	2.15	2.41
10	0.34	0.65	2.05	2.57	1.87	2.48
11	0.26	0.21	1.69	2.07	1.54	2.59
12	0.36	0.34	1.48	2.44	1.74	2.33
13	0.36	0.43	1.93	2.81	1.9	2.67
14	0.32	0.61	2.09	2.58	1.91	2.35
15	0.38	0.59	2.14	2.83	2.08	2.86
16	0.39	0.35	1.72	2.45	1.8	2.23
17	0.44	0.09	2.15	2.59	1.53	2.67
18	0.69	0.57	2.35	3.04	2.15	2.9
19	0.1	0.4	1.38	2.39	1.45	2
20	0.38	0.38	1.4	2.63	1.98	2.48
21	0.39	0.28	2.4	2.87	2.02	2.89
22	0.49	0.21	0.88	2.76	1.24	1.79
23	0.4	0.27	2.08	2.74	2.19	2.7
24	0.33	0.16	1.81	2.38	1.57	2.2
25	0.64	0.32	2.45	3.02	2.13	2.59
26	0.17	0.49	1.36	2.69	2.31	2.39
27	0.54	0.21	2.07	2.81	2.08	2.63
28	0.61	0.34	2.09	3.04	2.08	3.15
29	0.29	0.61	1.14	2	1.93	1.78

30	0.17	0.47	1.69	3.25	2.13	2.6
31	0.15	0.18	1.66	2.84	1.84	2.58
32	0.49	0.46	1.99	2.75	2.05	2.72
33	0.35	0.37	1.85	2.81	1.93	2.73
34	0.2	0.49	2.27	2.93	2.65	2.96
35	0.19	0.38	1.6	2.34	1.46	2.19
36	0.32	0.27	1.75	2.61	1.76	2.43
37	0.39	0.43	2.03	2.92	2.31	2.87
38	0.12	0.1	1.52	2	1.58	2.17
39	0.53	0.43	2.02	2.85	2.08	2.37
40	0.32	0.18	1.27	2.43	1.18	1.9
41	0.48	0.31	2.09	3.17	2.25	3.19
42	0.21	0.35	2.13	2.57	1.86	2.48
43	0.43	0.22	1.87	2.41	1.53	2.49
44	0.4	0.25	2.18	2.57	2.05	2.5
45	0.39	0.44	1.88	2.71	2.09	2.41
46	0.46	0.25	2.18	2.67	1.99	2.81
47	0.54	0.68	1.96	2.83	1.97	2.4
48	0.49	0.49	2.01	2.35	1.95	2.36
49	0.54	0.51	2.14	2.99	2.07	3.29
50	0.22	0.37	1.93	2.18	1.86	2.57
51	0.58	0.36	1.92	2.75	2.29	2.52
52	0.43	0.31	2.12	2.27	1.92	2.3
53	0.18	0.12	0.98	1.77	1.77	1.25

<i>M</i>	0.37	0.36	1.86	2.61	1.91	2.47
<i>SD</i>	0.15	0.15	0.36	0.37	0.30	0.39

Chapitre 4: Article 3

Frequency coding of facial parts

Nicolas Dupuis-Roy & Frédéric Gosselin

*Centre de Recherche en Neuropsychologie et Cognition, Département de Psychologie,
Université de Montréal, CP 6128, succ. Centre-ville, Montréal, H3C 3J7, CANADA*

Address correspondence to :

Nicolas Dupuis-Roy, Département de psychologie, Université de Montréal, PO Box 6128,
Succ. Centre-Ville, Montréal, QC H3C 3J7, Canada.

4.1. Abstract

How the brain samples spatiotemporal signals in order to form an accurate representation of its environment has been a long-standing issue in cognitive neuroscience. One hypothesis that has gained interest over the years is that the brain samples visual information through periodic and transient processes (see Tallon-Baudry & Bertrand, 1999; VanRullen & Koch, 2003; VanRullen & Dubois, 2011). Although traces of oscillatory processes have been repeatedly found in psychophysical experiments since the middle of the last century, efforts to map their frequency to specific aspects of visual processing remain elusive. One hundred and twelve participants performed 900 trials of a face gender categorization task in which the achromatic and isoluminant chromatic content of faces were partially revealed in space and time with random Gaussian apertures (see Gosselin & Schyns, 2001). Time-frequency wavelet transforms were performed on the time series recorded in the eyes and mouth regions to assess the frequency and latency at which they were sampled. Results show that achromatic and isoluminant chromatic information within the same facial part were sampled by the brain at the same frequency, whereas different facial parts were sampled at distinct frequencies (ranging from 6 to 10 Hz). This sampling code is consistent with recent electrophysiological evidence suggesting that facial features are multiplexed by the frequency of transient synchronized oscillations in the brain (see Schyns, Thut & Gross, 2011; Smith, Gosselin & Schyns, 2005, 2006, 2007; Thut et al., 2011; Romei, Driver, Schyns & Thut, 2011).

4.2. Introduction

How the brain samples spatiotemporal signals in order to form an accurate representation of its environment has been a long-standing issue in cognitive neuroscience. One hypothesis that has gained interest over the years is that the brain samples visual information through periodic and transient processes (see Uchida, Kepecs & Mainen, 2006; Laudau & Fries, 2012; Varela, et al., 1981; Kahala, Seelig & Malsen, 2001; Rodriguez, et al., 1999; Busch & VanRullen, 2010; Schyns, Thut & Gross, 2011; Smith, Gosselin & Schyns, 2005, 2006, 2007; Thut et al., 2011; Romei, Driver, Schyns & Thut, 2011; Tallon-Baudry & Bertrand, 1999; Tallon-Baudry et al., 1996; Tallon-Baudry et al., 1997; Engel & Singer, 2001; Fries, 2005). Traces of these oscillatory processes have been observed repeatedly in psychophysical experiments since the middle of the twentieth century.

One clear behavioral demonstration of an oscillating process in vision came from Purves, Paydarfar and Andrews (1996) who showed that the wagon wheel illusion (i.e. seeing a spoked wheel as rotating differently from its true rotation) does not only occur under stroboscopic presentation (e.g. in movies), but also arises under continuous illumination. The authors reasoned that if the cinematic version of the illusion results from temporal aliasing caused by the difference between the flickering rate of the camera and the rotation frequency of the spoked wheel, then the occurrence of the illusion in daylight implies that the brain samples visual information periodically, just as the snapshots of the camera. Later evidence further suggested that it was caused by a periodic visual sampling oscillating around 13 Hz (VanRullen, 2005), and that it depended on attentional resources (VanRullen, 2006; MacDonald, Cavanaugh, & VanRullen, 2013).

Other behavioral manifestations of these oscillations have been found in low-level (e.g. Latour, 1967; VanRullen, Carlson & Cavanagh, 2007; Venables, 1960; Andrews, 1996; Dehaene, 1993; Neri & Heeger, 2002; Neri & Levi, 2008) and high-level perceptual tasks (e.g. Blais, Arguin & Gosselin, 2013; Blais et al., 2009; Fiset et al., 2009; Vinette, Gosselin, & Schyns, 2004); using different measurements such as visual thresholds (Latour, 1967;

VanRullen, Carlson & Cavanagh, 2007), response times (Venables, 1960; Dehaene, 1993; Andrews, 1996) and accuracy rates (Landau & Fries, 2012); and with different techniques such as the reverse correlation (Blais, Arguin & Gosselin, 2013; Blais et al., 2009; Fiset et al., 2009; Gold & Shubel, 2006; Neri & Heeger, 2002; Neri & Levi, 2008; Vinette, Gosselin, & Schyns, 2004), and the wagon-wheel illusion under continuous illumination (Purves, Paydarfar, & Andrews, 1996; Reddy, Remy, Vayssiere, & VanRullen, 2011; Simpson, Shahani, & Manahilov, 2004; VanRullen, 2006, 2007; VanRullen, Reddy, & Koch, 2006; VanRullen et al., 2005).

The frequency at which the different targeted cognitive and perceptual processes were found to oscillate ranged from the delta (~0-4Hz) to the gamma band (~30-60Hz), although the vast majority fell within the theta (~4-7Hz), alpha (~7-14Hz) and beta band (~15-30Hz). Behavioral factors that were reported to influence this frequency include the difficulty of the task (Dehaene, 1993), the number of practice sessions (Dehaene, 1993), the attentional load (MacDonald, Cavanaugh, & VanRullen, 2013), the experimental method (Blais et al., 2013), and individual differences (Dehaene, 1993). For now, attempts at explaining the role played by the frequency of these oscillations have mainly focused on broad distinctions between perceptual, memory and attentional processes (for a review see VanRullen & Koch, 2003; VanRullen & Dubois, 2011). For instance, VanRullen and Koch (2003) argued that periodicities recorded in the alpha band could be associated with judgments involving explicit perception (including attentional processes), whereas those in the gamma band could be more directly correlated to behavior and reaction times. In the following experiment, we looked for a link between the frequency at which the brain sampled visual information and the nature of this information.

One hundred and twelve participants each completed 900 trials of a Bubbles face gender categorization task (Gosselin & Schyns, 2001). The chromatic and achromatic content of faces were partially revealed in space and time with random 3D gaussian apertures, allowing us to reverse correlate which chromatic and achromatic facial features led to accurate responses. Then, time-frequency wavelet transforms were applied on the time series recorded in these

specific facial features to assess the frequency and latency at which they were sampled. Results show that the frequency of the periodic visual sampling varied between the facial parts (i.e. right eye, left eye and mouth) but did not differ between the chromaticity type (i.e. chromatic or achromatic), suggesting that facial parts are coded by the frequency of transient and periodic visual sampling.

4.3. Methods

4.3.1. Participants

One hundred and twelve healthy participants (54 men) with normal color vision and normal or corrected-to-normal visual acuity were recruited on the campus of the Université de Montréal. Participants were aged between 20 and 30 years. An informed consent was obtained prior to the experiment and a monetary compensation was provided upon its completion. All procedures were carried out with the approval of the Université de Montréal ethics committee.

4.3.2. Face database

The face database contained 300 Caucasian frontal-view face images (150 males and 150 females) with a neutral expression. These faces were aligned on 12 handpicked landmarks. Interpupil distance was set to 1.03 degrees of visual angle during the experiment. For additional details on this database see Dupuis-Roy et al. (2009).

4.3.3. Stimuli

The chromatic and achromatic content of a face image were independently revealed partially in space (128 X 128 pixels) and time (17 frames) with random 3D Gaussian apertures. This required four steps. A face image was first converted from its original RGB luminance maps to the HSV colour maps. Second, two volumes of 128 X 128 X 17 voxels, henceforth M1 and M2, with approximately the same number of randomly located 3D

Gaussian apertures were generated. Each of these 3D Gaussian apertures ($\sigma_{xy} = 0.15^\circ$ of visual angle, $\sigma_{time} = 23.53$ ms) were scaled and truncated so that their probability density ranged from zero (at the tails $\pm 2.18 \sigma$) to one (at the center). Third, the following mathematical operations were performed 17 times on the H (hue), S (saturation) and V (value) channels of a given face image:

$$\begin{aligned} \text{Eq 1 :} \quad H' (:, :, i) &= H \\ S' (:, :, i) &= M1 (:, :, i) .* S \\ V' (:, :, i) &= M2 (:, :, i) .* (V - \mu) + \mu, \end{aligned}$$

where $i=[1, 17]$; H, S and V are the original channels of a given face image; H', S' and V' are the resulting sampled volumes (128 X 128 X 17 voxels); and μ is the average of all pixels in V. In conjunction with the H channel, each bubble of M1 reveals hue information with the maximum saturation at its centre. Each bubble of the M2 bubbles volume reveals the maximum value at its centre. When bubbles of the M1 and M2 volume do not overlap, the areas revealed through the M1 bubbles have approximately the same average luminance (± 5 cd/m² which represents 3.7% of the full range), whereas the areas revealed through the M2 bubbles have a null saturation and thus no hue. When bubbles of the M1 and M2 volumes overlap, they reveal all the information. Finally, the H', S' and V' volumes are combined and converted to RGB values. Figure 1a illustrates one iteration of the sampling process that was repeated 17 times to produce a stimulus.

4.3.4. Apparatus

The experimental programs were run on four Macintosh computers in the Matlab (Mathwork Inc.) environment, using functions from the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The four computer high-resolution CRT monitors used for stimuli presentation were set to display 1024 by 768 pixels at a refresh rate of 85 Hz. The relationship between RGB values and luminance levels (measured with a Samsung SyncMaster 753 df photometer) was computed for each color channel independently; the three best-fitted gamma

functions were then used to convert each image of the face database into HSV space. Participants were seated in a dim ambient-lighted room. Viewing distance was maintained constant by a chin-rest throughout the experiment.

4.3.5. Procedure

First, participants were screened for possible colour vision problems with the Ishihara Color Test (Ishihara, 1917). Then, they were brought in the experimental room where they completed three blocks of 300 trials in which all faces of the database were shown once, in a random order, and partially revealed. The sequence of events in a given trial unfold as follows (see Figure 1b): A fixation cross was shown for 750 ms; a uniform grey screen was displayed for 250 ms; a stimulus video containing 17 sampled-face images was projected at the center of the screen (at a rate of 85 Hz) for a total duration of 200 ms, and remained there until the participant had indicated the gender of the sampled face by pressing a labeled keyboard key. No feedback was provided. The number of bubbles was adjusted on a trial-by-trial basis with Quest (Watson, 1983) to maintain 75% of correct responses.

4.4. Results and Discussion

Participants required an average of 97.23 Gaussian apertures (SD = 35.95) to reach the target 75% of correct responses. There was a slight bias toward responding “woman” (50.23%) rather than “man” (49.77%). Female observers answered “man” more often (M=51.36%, SD=7.1%) than male observers did (M=48.12%, SD=6.80%). Four participants were excluded from the rest of the data analyses because their performance was at least three standard deviations away from the mean.

Insert Fig. 1 near here

4.4.1. Latency and frequency of transient synchronized information sampling

First, two least-square multiple linear regressions were performed: one between the volumes of chromatic Gaussian apertures (explanatory variables) and the participant's response accuracy (predictor variable); and another one between the volumes of achromatic Gaussian apertures and the participant's response accuracy. The outcome of these regressions were two classification volumes (128 X 128 X 17) of regression coefficients per participant revealing how each participant extracted spatial achromatic and chromatic face-gender information through time to respond.

Next, the frequency and the latency at which the chromatic and achromatic information were sampled were assessed with a time-frequency continuous wavelet transform. We focused on the time series recorded in three of the most prominent gender cues : the left eye, the right eye and the mouth regions (see Brown & Perrett, 1993; Dupuis-Roy, et al., 2009; Nestor & Tarr, 2008a; Roberts & Bruce, 1988; Yamaguchi, Hirukawa, & Kanazawa, 1995; Schyns, Bonnar, and Gosselin, 2002; Gosselin & Schyns, 2001; Burton, Bruce, & Dench, 1993). More specifically, all time series within a given facial part of a given participant's chromatic and achromatic classification volume were averaged. This led to six time series per subject, i.e. two for each facial part. A continuous wavelet transform was then applied on each averaged time series within the frequency range of 3 to 25 Hz. The complex Morlet wavelet used was of the form:

Eq 2 :
$$M(t, f) = e^{-\frac{t^2}{2\sigma^2}} * e^{i2\pi ft},$$

where f was the center frequency of the considered band with $\sigma_t = 6/f$. The power was then averaged across participants, leading to six grand average time-frequency spectra. The probability of the power magnitude was estimated via a bootstrapping procedure. First, this required the creation of a chromatic and achromatic random classification volumes, which were computed following a random permutation of each participant's responses. Then, 1,000 averaged time series were created for each facial part and for each chromaticity type by averaging m (the number of time series in a given facial region) randomly selected time series in the bootstrap classification volumes.

Insert Fig. 2 near here

Figure 2 shows the achromatic (left column) and the chromatic (right column) grand average time-frequency spectrum of the left (top row) and right eye (middle row) and the mouth region (bottom row). The colour scale is expressed in log units and reflects the probability of a given power magnitude: the smaller the probability, the higher the power magnitude. The thick black lines delimit the significant clusters ($p < 0.01$). The gravity centers of the significant clusters are identified by the color red for chromatic, by the color black for achromatic, by left-pointing triangles for the left eye, by right-pointing triangles for the right eye and by circles for the mouth. A single significant cluster was found in each grand average time-frequency spectrum except the achromatic right eye spectrum. For this spectrum, the gravity centre was computed on the top 1.5% power magnitude values.

Figure 3b shows an outline of all centers of gravity and significant clusters. The latency t_i and frequency f_i of each gravity center i are also indicated on the graph. Figure 3a presents the temporal profile of each gravity center computed over a frequency window of $f_i \pm 0.89$ Hz whereas Figure 3c shows the frequency profile of each gravity center computed over a time window of $t_i \pm 24$ ms. The power is expressed in normalized units so that each curve peaks at a 100%.

In the temporal domain, the right eye and the mouth were extracted almost simultaneously, first around 37 ms in the chromatic domain (i.e. at 48 ms and 26 ms respectively), and then around 136 ms in the achromatic domain (i.e. at 144 ms and 128 ms respectively). The chromatic and achromatic left eye did not follow the same temporal pattern: they were both extracted around 224.5 ms (i.e. at 233 ms and 216 ms respectively).

Insert Fig. 3 near here

In the frequency domain, each facial part was extracted at a distinctive frequency, regardless of whether it was chromatic or achromatic: the left eye was best processed around 8.54 Hz (chromatic = 8.49 Hz; achromatic = 8.58 Hz); the right eye around 6.58 Hz

(chromatic = 6.43 Hz; achromatic = 6.72 Hz); and the mouth around 10.68 Hz (chromatic = 10.54 Hz; achromatic = 10.82 Hz). The distances between the frequencies at the peak power of the spectrum are much smaller within ($M = 0.39 \pm 0.23$ Hz) than between facial parts ($M = 2.55 \pm 1.42$ Hz), which indicates that the facial parts are sampled by the visual system at different frequencies.

Next, we looked for evidence that the brain samples face parts at different frequencies in individual time-frequency spectra. To increase the signal-to-noise ratio, we selected participants ($N=84$) who had statistically significant power ($p<0.01$) over the range of 5 to 12 Hz, and computed the frequency at the peak power of their six time-frequency spectra. A 2 (chromatic, achromatic) by 3 (left eye, right eye, mouth) ANOVA for repeated measures was first performed on the frequency at the peak power. The results indicated a significant effect of the facial parts ($F(2,83)=3.06$, $p<0.05$), but no significant effect of the chromaticity type ($F(1,83)=0.04$, ns), and no significant interaction ($F(2,83)=0.02$, ns). More importantly, a t-test for paired samples was conducted to verify if distances between peak frequencies were smaller within than between facial parts. Within-part peak frequency distances were computed as the average of the chromatic vs achromatic distances, whereas between-part peak frequency distances were computed as the average of all distances between pairs of facial parts (left eye vs right eye, left eye vs mouth, and right eye vs mouth). The results revealed that peak frequency distances were smaller within ($M = 1.64 \pm 1.29$ Hz) than between facial parts ($M = 2.07 \pm 1.00$ Hz, $t(83)=2.20$, $p<0.05$). Once more, these results indicate that the brain samples facial parts at distinct frequencies.

This frequency sampling code is consistent with recent electrophysiological evidence suggesting that facial informations are multiplexed by the frequency of transient synchronized oscillations in the parietal brain. Indeed, Schyns, Thut and Gross (2011) found that EEG power in the theta band (~ 4 Hz) correlates with the use of the mouth information in a facial expression categorization task, whereas the power in the beta band (~ 12 Hz) correlates with the use of the eyes information. Furthermore, Smith et al., (2006) and Romei et al., (2011) showed that global, low-frequency, and local, high-frequency, visual information were

multiplexed in different frequency bands, i.e. the theta and beta bands respectively. Further investigations will be required to test this hypothesis and to unveil the link between behavioral and neural oscillations.

4.6. Acknowledgements

This research was supported by a scholarship from the Social Sciences and Humanities Research Council awarded to Nicolas Dupuis-Roy, and by a grant from the Natural Sciences and Engineering Research Council of Canada awarded to Frédéric Gosselin.

4.7. Figure legends

Figure 1. A) One iteration of the sampling process that is repeated 17 times to generate the full 128 X 128 X 17 face stimulus. A RGB face is inputted in the sampling function. This function (symbolized by a box) first creates two masks of randomly located 3D Bubbles. Only one frame of these two volumes is displayed in the box. Then, the RGB luminance values of the inputted image are converted into their HSV equivalent (see the second column inside the box). The S and V channels are sampled with the related Bubble mask, and the resulting H', S' and V' sampled maps are reconverted to the RGB colour map. B) The sequence of events in a trial: A fixation cross appears for 750ms; a 250ms-long blank screen follows; then a given sampled face is displayed for 200ms at a rate of 85Hz; finally, a blank screen remains until a response is provided.

Figure 2. The chromatic (left column) and the achromatic (right column) time-frequency spectrum of the left (top row) and right eye (middle row) and the mouth region (bottom row). The color scale is expressed in log units and reflects the probability of a given power magnitude. The thick black lines delimit the significant clusters ($p < 0.01$). The gravity centers of the significant clusters are identified by the color red for chromatic, by the color black for achromatic, by left-pointing triangles for the left eye, by right-pointing triangles for the right eye and by circles for the mouth.

Figure 3. A) Temporal profile of each gravity center computed over a frequency window of $f_i \pm 0.89$ Hz, where f_i is the frequency of gravity center i . B) Outline of all centers of gravity and significant clusters shown in Fig 2. The latency t_i and frequency f_i of each gravity center i are indicated on the graph. C) Frequency profile of each gravity center computed over a time window of $t_i \pm 24$ ms.

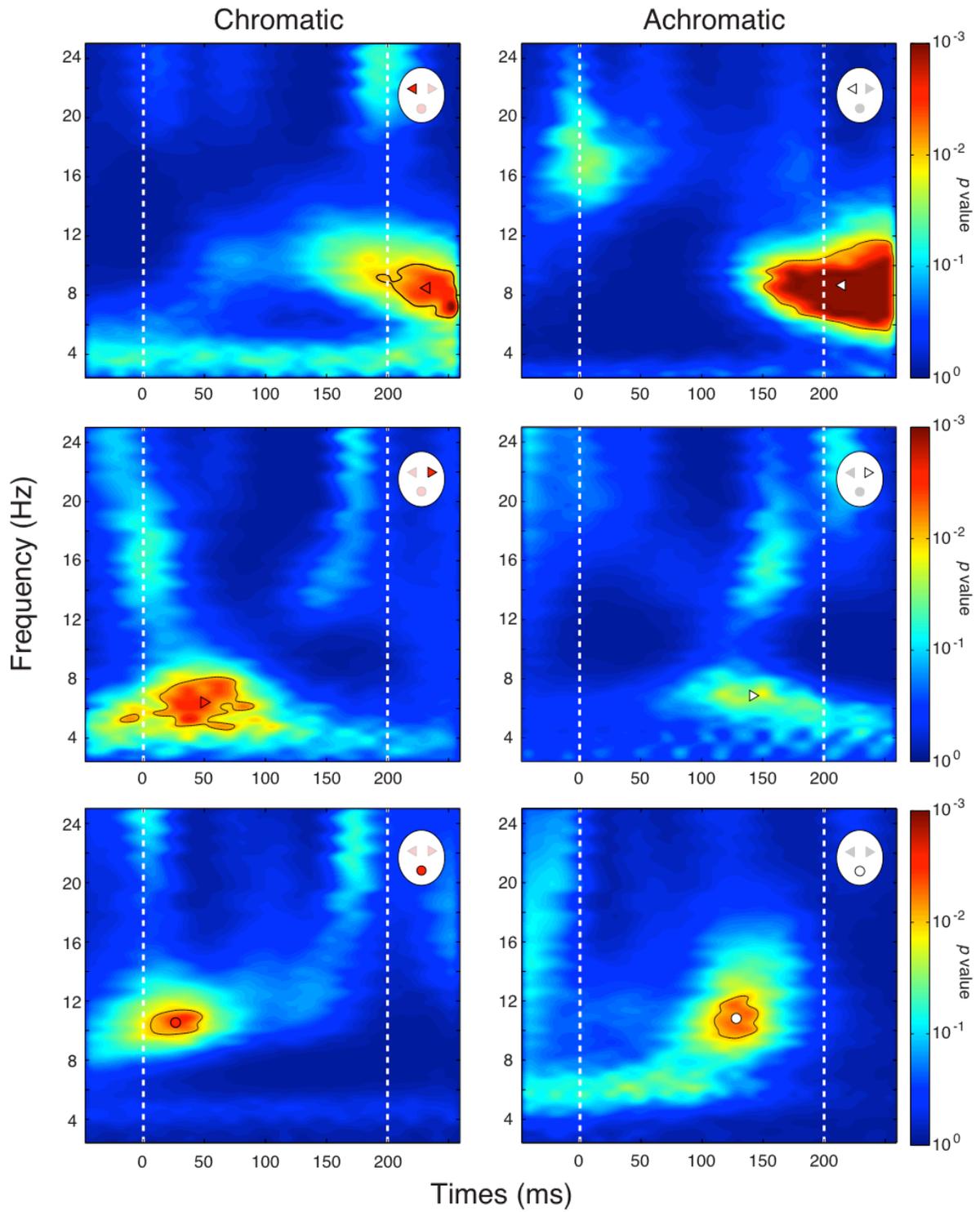


Figure 2. Dupuis-Roy & Gosselin

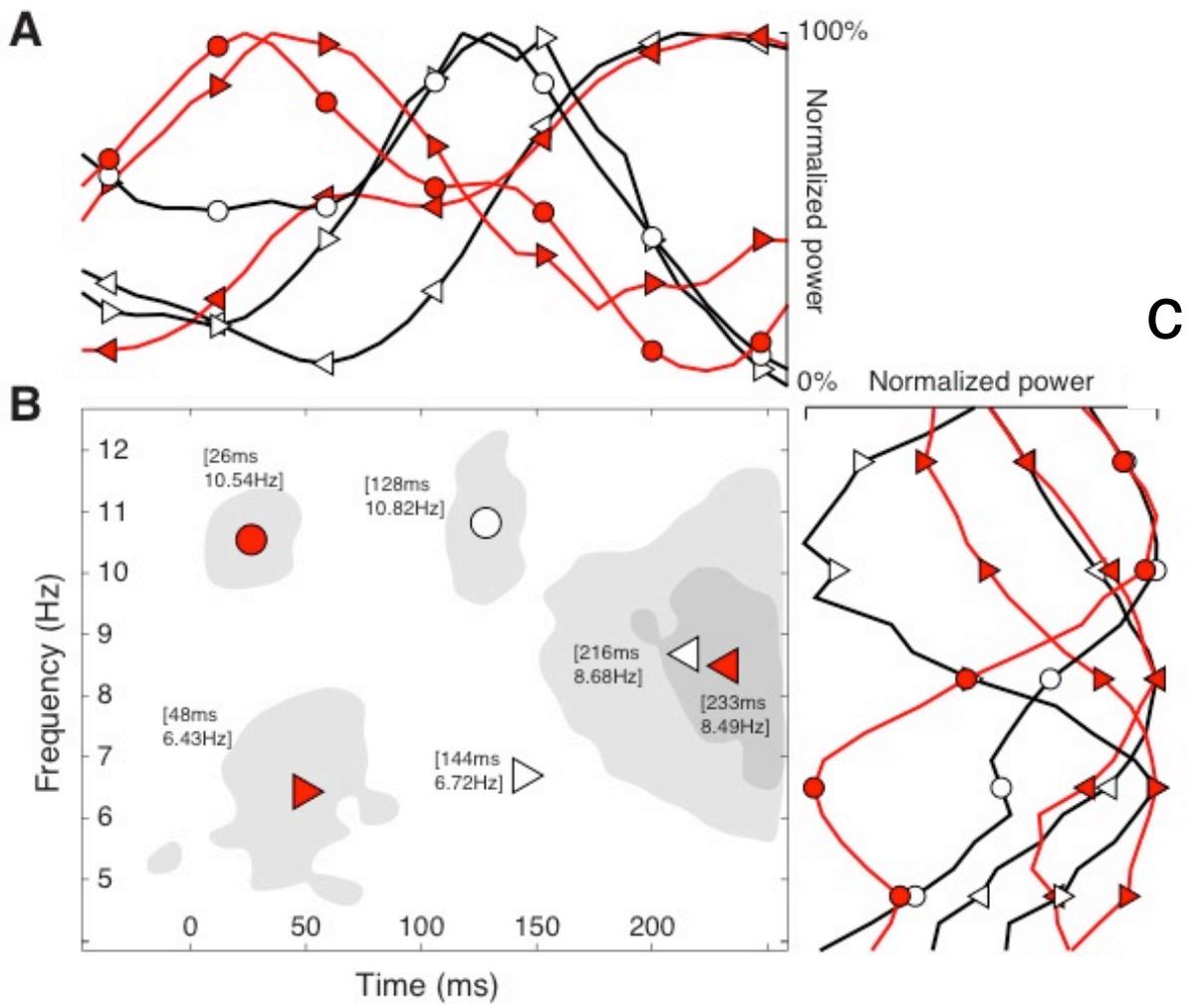


Figure 3. Dupuis-Roy & Gosselin

4.8. References

- Blais, C., Arguin, M., & Gosselin, F. (2013). Human visual processing oscillates: Evidence from a classification image technique. *Cognition*, *128*(3), 353–362. doi:10.1016/j.cognition.2013.04.009
- Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia*, *50*(12), 2830–2838. doi:10.1016/j.neuropsychologia.2012.08.010
- Burton, A. M., Bruce, V., & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception*, *22*(2), 153–176.
- Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(37), 16048–16053. doi:10.1073/pnas.1004801107/-/DCSupplemental
- Dehaene, S. (1993). Temporal oscillations in human perception. *Psychological Science*, *4*(4), 264–270.
- Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision*, *9*(2), 1-8.
- Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences*, *5*(1), 16–25.
- Fiset, D., Blais, C., Arguin, M., Tadros, K., Éthier-Majcher, C., Bub, D., & Gosselin, F. (2009). The spatio-temporal dynamics of visual letter recognition. *Cognitive Neuropsychology*, *26*(1), 23–35. doi:10.1080/02643290802421160

- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9(10)*, 474–480. doi:10.1016/j.tics.2005.08.011
- Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, *10(13)*, 805–808.
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research*, *41*, 2261–2271.
- Ishihara, S. (1917). *Tests for color-blindness*. Handaya : Tokyo, Hongo Harukicho.
- Kahana, M. J., Sekuler, R., Caplan, J. B., Kirschen, M., & Madsen, J. R. (1999). Human theta oscillations exhibit task dependence during virtual maze navigation. *Nature*, *399(6738)*, 781–784.
- Landau, A. N., & Fries, P. (2012). Attention Samples Stimuli Rhythmically. *Current Biology*, *22(11)*, 1000–1004. doi:10.1016/j.cub.2012.03.054
- Lee, K. J., & Perrett, D. I. (2000). Manipulation of colour and shape information and its consequence upon recognition and best-likeness judgments. *Perception*, *29(11)*, 1291–1312. doi:10.1068/p2792
- Lisman, J. E., & Idiart, M. A. (1995). Storage of 7±2 short-term memories in oscillatory subcycles. *Science*, *267(5203)*, 1512–1515.
- Neri, P., & Levi, D. (2008). Temporal dynamics of directional selectivity in human vision. *Journal of Vision*, *8(1)*, 22–22. doi:10.1167/8.1.22
- Nestor, A., & Tarr, M. J. (2008a). Gender recognition of human faces using color.

Psychological Science, 19(12), 1242–1246.

Nestor, A., & Tarr, M. J. (2008b). The segmental structure of faces and its use in gender recognition. *Journal of Vision*, 8(7):7, 1-12.

Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proceedings of the National Academy of Sciences of the United States of America*, 93(8), 3693–3697.

Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature*, 397(6718), 430–433.

Romei, V., Driver, J., Schyns, P. G., & Thut, G. (2011). Rhythmic TMS over parietal cortex links distinct brain frequencies to global versus local visual processing. *Current Biology*, 21(4), 334–337. doi:10.1016/j.cub.2011.01.035

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13, 402–409.

Schyns, P. G., Gosselin, F., & Smith, M. L. (2009). Information processing algorithms in the brain. *Trends in Cognitive Sciences*, 13(1), 20–26. doi:10.1016/j.tics.2008.09.008

Schyns, P. G., Thut, G., & Gross, J. (2011). Cracking the code of oscillatory activity. *PLoS Biology*, 9(5), e1001064. doi:10.1371/journal.pbio.1001064.s003

Smith, M. L., Gosselin, F., & Schyns, P. G. (2006). Perceptual moments of conscious visual experience inferred from oscillatory brain activity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(14), 5626–5631.

- Smith, M. L., Gosselin, F., & Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *NeuroImage*, *37*(3), 974–984. doi:10.1016/j.neuroimage.2007.05.030
- Stephen, I. D., & McKeegan, A. M. (2010). Lip colour affects perceived sex typicality and attractiveness of human faces. *Perception*, *39*(8), 1104–1110. doi:10.1068/p6730
- Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, *3*(4), 151–162.
- Tanaka, J. W., Weiskopf, D., & Williams, P. (2001). The role of color in high-level vision. *Trends in Cognitive Sciences*, *5*, 211–215.
- Tesche, C. D., & Karhu, J. (2000). Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(2), 919–924.
- Thut, G., Veniero, D., Romei, V., Miniussi, C., Schyns, P., & Gross, J. (2011). Rhythmic TMS causes local entrainment of natural oscillatory signatures. *Current Biology*, *21*(14), 1176–1185. doi:10.1016/j.cub.2011.05.049
- Uchida, N., Kepecs, A., & Mainen, Z. F. (2006). Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nature Reviews Neuroscience*, *7*(6), 485–491.
- VanRullen, R., & Dubois, J. (2011). The psychophysics of brain rhythms. *Frontiers in Psychology*, *2*. doi:10.3389/fpsyg.2011.00203/abstract
- VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, *7*(5), 207–213. doi:10.1016/S1364-6613(03)00095-0

- VanRullen, R., Carlson, T., & Cavanagh, P. (2007). The Blinking Spotlight of Attention. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(49), 19204–19209.
- Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*(4), 229–239.
- Vinette, C., & Gosselin, F. (2004). Spatio-Temporal Dynamics of Face Recognition in a Flash: It's in the eyes. *Cognitive Science*, *28*, 289–301, doi:10.1016/j.cogsci.2004.01.002
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Attention, Perception, & Psychophysics*, *33*(2), 113–120.
- Yamaguchi, M. K., Hirukawa, T., & Kanazawa, S. (1995). Judgment of gender through facial parts. *Perception*, *24*, 563–563.
- Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, *31*, 995–1003.

Chapitre 5: Discussion générale

5.1. Avant-propos

Les dimorphismes sexuels faciaux regroupent l'ensemble des différences phénotypiques entre les visages adultes féminins et masculins. Sachant que l'humain ne dispose que de ressources cognitives et perceptives limitées, la question principale de cette thèse a donc été de savoir lesquels de ces innombrables signaux se répercutant sur le spectre visible de la lumière sont utilisés par l'observateur humain afin de qualifier le sexe d'un visage. Cette question a fait l'objet du premier article duquel ont découlé d'autres questions, cette fois plus précises, quant à l'importance relative des différents indices faciaux achromatiques et chromatiques et à leur extraction spatiotemporelle. Plus spécifiquement, cette thèse comprenait trois articles examinant : l'utilisation des indices faciaux locaux (sur l'axe xy) dans un contexte de plus grande validité externe (article 1); l'importance relative des distances interattributs et des indices achromatiques et chromatiques (article 2); et l'extraction spatiotemporelle des indices faciaux achromatiques et chromatiques (article 3). Dans la présente discussion, nous résumerons et critiquerons d'abord les résultats de ces trois articles. Ensuite, nous en évaluerons la portée scientifique et les limites. Finalement, nous proposerons des projets scientifiques pouvant faire suite à ces travaux.

5.2. Utilisation des indices faciaux chromatiques et achromatiques

Les indices achromatiques et chromatiques locaux regroupent tous les indices qui produisent des variations locales de réflectance. Ceux-ci comprennent principalement les indices de pigmentation, de texture et de forme. Les études antérieures examinant l'importance perceptive de ces indices dans la catégorisation du sexe des visages souffraient d'au moins une des limites suivantes quant à leur validité externe. Premièrement, les stimuli utilisés n'étaient pas représentatifs des visages perçus dans la vie de tous les jours : certains avaient subi des transformations majeures, d'autres voilaient des signes sexuels faciaux et non-faciaux ostensibles tels que les cheveux, le maquillage ou les oreilles, la grande majorité étaient en tons de gris, et certains contenaient des indices accidentels de taille absolue. Deuxièmement,

l'échantillon des visages testés était parfois très restreint. Troisièmement, les méthodologies utilisées, comme l'isolement, le masquage et la substitution, requéraient la sélection de certaines zones faciales *a priori*.

Le premier article de cette thèse (voir Chapitre 2) visait à examiner l'utilisation des indices achromatiques et chromatiques locaux dans un contexte de plus grande validité externe. Afin de garantir un tel contexte, un plus grand nombre de visages (i.e. 300) fut utilisé, et chaque visage ne fut présenté qu'une seule fois à chaque participant. Un tel choix visait à minimiser l'utilisation de stratégies dites atypiques, basées sur l'apprentissage des caractéristiques idiosyncratiques des visages. Plus encore, nous avons employé une méthode d'échantillonnage aléatoire non biaisée, Bubbles (voir Gosselin & Schyns, 2001), ne supposant aucune hypothèse quant aux zones du visage à tester. La probabilité d'échantillonnage étant égale pour l'ensemble des pixels de l'image, aucun avantage statistique n'était donc associé à l'utilisation d'une zone faciale particulière, autre que le signal dimorphique sexuel qu'elle comportait. Finalement, les visages affichés lors de l'expérience étaient plus représentatifs des visages vus dans la vie courante : ils étaient en couleurs et comportaient certains signes sexuels externes comme la chevelure et le maquillage. C'est donc dans un contexte de plus grande validité externe que l'expérience se déroula.

Deux régressions linéaires multiples ont été réalisées sur les données de cette expérience : une première entre les masques de bulles et la justesse des réponses à la tâche de catégorisation du sexe des visages, et une seconde entre les masques de bulles et la latence des réponses correctes dans cette même tâche. Les résultats de ces deux analyses, présentés à la Figure 1, indiquent que la région orale et la région comprenant les yeux et les sourcils sont significativement corrélées à la catégorisation exacte et rapide du sexe des visages. L'échelle de couleur utilisée à la Figure 1 représente la probabilité cumulative d'observer un score Z étant donné que l'hypothèse nulle est fautive. Donc, plus la probabilité se rapproche de 1 (rouge), plus le pixel est statistiquement et positivement lié à la justesse ou la rapidité de réponse.

Le troisième article de cette thèse (voir Chapitre 4) visait à étudier l'extraction temporelle des indices achromatiques et chromatiques locaux dans un contexte similaire à celui de l'article 1. Dans ce troisième article, la luminance et la chrominance des visages fut échantillonnée indépendamment avec des bulles spatiotemporelles ($\sigma_{xy} = 0.15^\circ$ d'angle visuel, et $\sigma_{\text{temps}} = 23.53$ ms) dans l'espace des couleurs HSV. Afin de comparer les résultats de cette étude à ceux de la première étude, nous avons calculé des images de classification (ICs) bidimensionnelles en sommant les volumes de classification (voir section 4.4) sur la dimension temporelle. Plus spécifiquement, ces ICs sont le résultat de deux régressions linéaires multiples entre, d'une part, les coordonnées xy des centres des bulles chromatiques et achromatiques (variables indépendantes) incluses dans les volumes d'échantillonnage et, d'autre part, la justesse de réponse de chaque participant d'autre part (variable indépendante). Les coefficients de régressions résultant de cette analyse ont d'abord été convolués avec un filtre gaussien 2D ($\sigma_{xy} = 0.15^\circ$ d'angle visuel), puis transformés en scores Z à partir d'une simulation bootstrap (e.g. Efron et Tibshirani, 1986). Ces ICs représentent donc l'association linéaire entre la localisation xy d'une bulle chromatique (voir Figure 1B, image de gauche) ou d'une bulle achromatique (voir Figure 14B, image de droite) et la justesse de réponse à la tâche de catégorisation sexuelle des visages. L'échelle de couleur représente, comme précédemment, la probabilité cumulée d'observer un score Z étant donné que l'hypothèse nulle est fausse.

Les résultats des études 1 et 3 soulignent l'importance des informations chromatiques dans la région orale et des informations achromatiques dans les régions oculaires pour la catégorisation sexuelle des visages. Ces résultats confirment partiellement les découvertes des études précédentes sur le sujet.

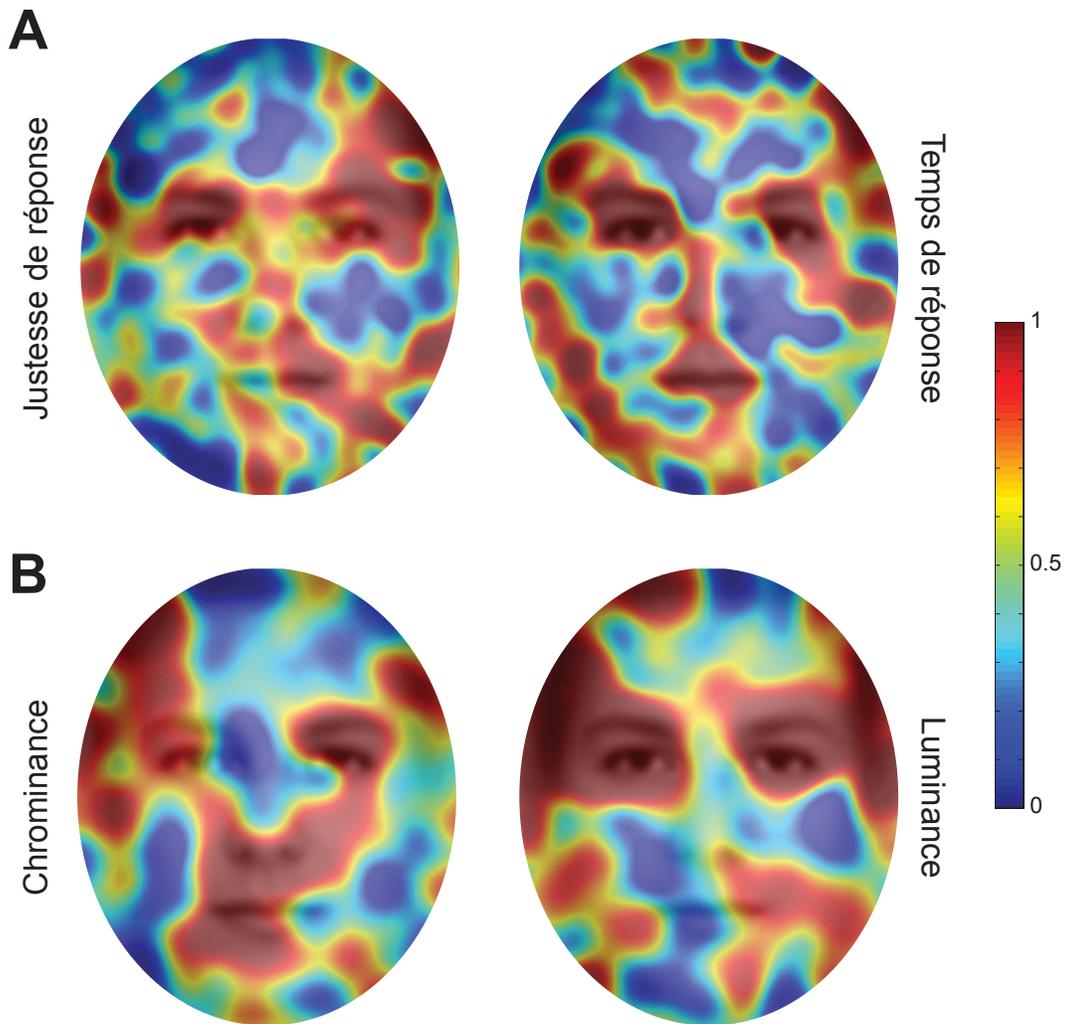


Figure 1. A) ICs moyennes pour les données de l'Article 1. L'IC de gauche exprime l'association entre les masques de bulles et la justesse de réponse, tandis que l'image de droite exprime l'association entre les masques de bulles et la latence des réponses correctes. B) ICs moyennes pour les données de l'Article 2. L'IC de gauche exprime l'association entre les masques de bulles chromatiques et la justesse de réponse, tandis que l'image de droite exprime l'association entre les masques de bulles achromatiques et la latence des réponses correctes. Les ICs sont exprimées en probabilités cumulatives affichées (voir échelle de couleur de gauche) et juxtaposées à un visage androgyne moyen.

Premièrement, les six études ayant testé des zones faciales sélectionnées *a priori*, c.-à-d. les études de Roberts et Bruce (1988), Bruce et al. (1993), Brown et Perret (1993) et Yamaguchi, Hirukawa et Kanazawa (1995) se sont soldées par des résultats significatifs et/ou une taille d'effet expérimental importante dans la grande région oculaire. Ces mêmes études soulignent aussi l'importance de la mâchoire et du menton, soit le contour du visage.

Deuxièmement, des résultats similaires ont aussi été observés dans les études ayant utilisé des méthodes sans *a priori* sur l'importance des zones faciales : l'étude de Saether et al. (2009) a montré l'importance de la zone infraorbitale à l'aide de l'oculométrie; l'étude de Nestor et Tarr (2008b) a montré l'importance des zones orbitales, nasales et orales grâce à la "reverse correlation"; l'étude de Gosselin et Schyns (2001) a montré l'importance des régions oculaire et orale en utilisant la technique Bubbles dans l'espace spatial (i.e. xy) ; tandis que l'étude de Schyns, Bonnar et Gosselin (2002) a montré l'importance de l'oeil gauche en hautes fréquences spatiales, des zones oculaires, nasale et orale en moyennes hautes et moyennes basses fréquences spatiales et de tout le contour du visage en moyennes basses fréquences spatiales au moyen de la technique Bubbles dans l'espace xy et des fréquences spatiales.

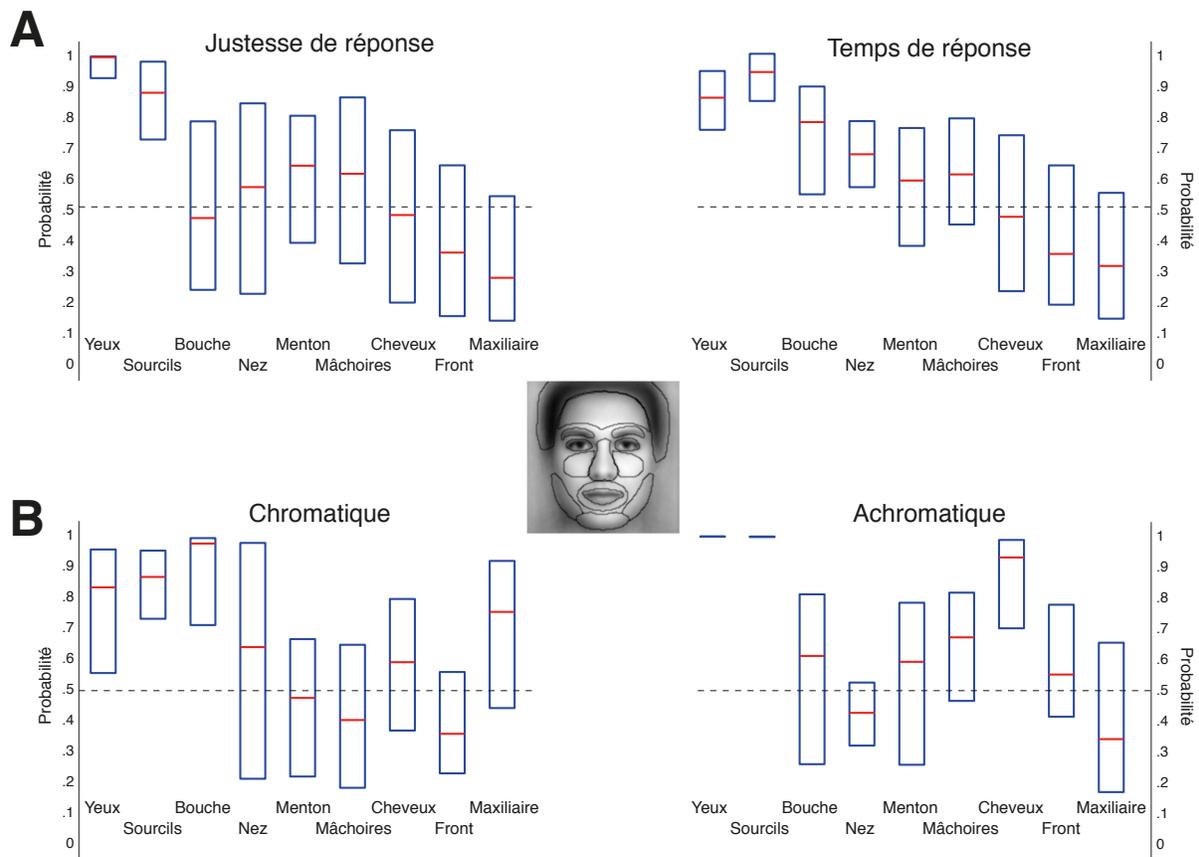


Figure 2. « Box plot » présentant les percentiles 75, 50 et 25 des probabilités cumulatives calculées dans neuf régions d'intérêt illustrées sur l'image au centre de la figure. Percentiles des probabilités observées dans les ICs de la première étude (panneau A) et de la troisième étude (panneau B).

Ainsi, les régions identifiées comme étant les plus importantes dans les études 1 et 3 s'accordent aux résultats de la majorité des études antérieures (voir section 1.2 pour une analyse plus poussée des résultats antérieurs). Cependant, d'autres régions faciales comme la zone nasale et le contour du visage ont été identifiées comme importantes dans plusieurs études mais n'ont pas atteint le seuil de signification dans notre expérience. Plusieurs facteurs peuvent expliquer ce décalage.

D'entrée de jeu, il faut préciser que le fait de ne pas avoir obtenu un résultat significatif n'implique pas nécessairement l'absence d'un effet. Les ICs de la Figure 1 présentent

également des probabilités cumulatives élevées dans d'autres zones faciales. Pour mieux comparer l'utilité de chacune des régions faciales, considérez la Figure 2 qui résume, pour chacune des ICs présentées à la Figure 1, les probabilités cumulatives enregistrées dans neuf régions faciales: les yeux, les sourcils, la bouche, le nez, le menton, les mâchoires, les cheveux, le front et les maxillaires. Les "box plot" représentent les percentiles 25, 50 et 75 des probabilités cumulatives observées dans chacune de ces régions faciales. Cette figure suggère l'apport potentiel de la zone maxillaire et nasale sur le plan chromatique, ainsi que des cheveux et des mâchoires sur le plan achromatique. Ceci dit, la présence d'un effet significatif dans une région faciale donnée est liée à deux facteurs statistiques: la taille de l'échantillon et la taille de l'effet expérimental. Étant donné que la technique Bubbles révèle l'information P , soit la portion des informations sexuelles disponibles A incluse dans les représentations visuelles R , la taille de l'effet est elle-même tributaire de A et de R . Ainsi, les chances de révéler des effets significatifs dans une région faciale donnée est fonction de la quantité d'information présente et de son importance au niveau de la représentation. Dans un tel contexte, il est donc difficile de révéler soit les régions qui contiennent très peu d'informations sexuelles mais qui sont tout de même importantes au niveau représentationnel, soit les régions riches en informations sexuelles mais qui ont une moindre importance au niveau des représentations. En 2002, Gosselin et Schyns ont montré qu'il était possible d'estimer R à partir des données d'une expérience avec Bubbles en comparant l'image de classification humaine, donc P , à l'image de classification de l'observateur idéal. Comme cet observateur est optimal, son image de classification P devrait épuiser toute l'information contenue dans A , et ce, compte tenu des limites statistiques de l'expérience (i.e. le nombre d'essais, le nombre de bulles, etc.). Ainsi, $A=P$ pour l'observateur idéal. La comparaison du P humain au P (ou A) idéal, bien qu'elle ne sera pas présentée dans le cadre de cette thèse, permettrait justement de révéler les régions qui renferment moins d'informations sexuelles mais qui sont tout de même importantes au niveau de R . En conclusion, nos résultats n'excluent pas la possibilité que d'autres régions faciales soient utilisées par les observateurs humains afin de catégoriser le sexe des visages, mais ils suggèrent que ces autres régions revêtent une importance secondaire sur le plan de l'information sexuelle et/ou de la représentation visuelle.

5.3. Importance relative et rôle des indices chromatiques

Le rôle que joue la couleur dans la perception des objets et des visages est encore peu connu. De récentes études suggèrent que la couleur pourrait intervenir dans le traitement de l'information de haut et de bas niveau (voir la revue de Tanaka, Weiskopf et Williams, 2001). Sur le plan de la perception des visages, des études ont suggéré que la couleur pourrait jouer un certain rôle dans plusieurs jugements perceptifs à caractère biologique comme la catégorisation du sexe et de l'ethnicité (Hill et al., 1995; Tarr et al., 2001), la perception de l'âge (Burt et Perrett, 1995; Fink, Grammer & Matts, 2006; Matts et al 2007), de la santé (Stephen et al., 2009a, 2009b), de l'attractivité (Fink, Grammer & Matts 2006; Matts et al 2007) et de l'identité (Lee et Perrett, 1996; Yip et Sinha, 2002). Par exemple, Lee et Perrett (1996) ont observé que les performances d'identification de visages étaient significativement plus élevées pour des visages présentés en couleurs que pour des visages présentés en tons de gris. Plus encore, ils ont enregistré des performances d'identification significativement plus élevées avec des visages dont le signal chromatique avait été amplifié qu'avec des visages en couleurs non-amplifiées. Similairement, Yip et Sinha (2002) ont enregistré des performances significativement plus élevées avec des visages en couleurs et en pseudo-couleurs³ qu'avec des visages en tons de gris, et ce, seulement lorsque ces visages étaient filtrés en basses fréquences spatiales. En fait, ces auteurs n'ont observé aucune différence de performance entre les visages en couleurs et les visages en pseudo-couleurs ce qui suggère que la valeur absolue de la tonalité chromatique n'est pas aussi importante que les différences relatives de tonalité chromatique. Sur la base de ces résultats, les auteurs ont postulé que la couleur améliorerait les performances d'identification non pas seulement en raison des informations sur l'identité qu'elle contient, mais aussi parce qu'elle pourrait faciliter la segmentation des différentes parties des visages. Une autre hypothèse pourrait expliquer partiellement le résultat de Yip et Sinha (2002). La sensibilité humaine au contraste dans les basses fréquences spatiales étant plus élevée pour les variations chromatiques que pour les variations achromatiques (Mullen,

³ Ces pseudo-couleurs ont été obtenues via une rotation du canal H dans l'espace H (« hue » ou tonalité), S (saturation) et V (valeur ou luminance).

1985), il est possible que la couleur ait amélioré l'accès à cette bande de fréquences spatiales qui est importante à l'identification des visages (e.g. Schyns, Bonnar & Gosselin, 2002).

L'article 2 avait pour premier but de comparer l'efficacité — un index de la proportion de l'information disponible *A* utilisée par le participant — de l'utilisation des distances interattributs (DIA) pour la catégorisation du sexe des visages à l'efficacité de l'utilisation de tous les autres indices faciaux. L'article 2 comparait aussi l'efficacité du traitement des visages en couleurs (condition COL) et des visages en tons de gris (condition LUM). Dans les conditions IADs (donc IADs-COL et IADs-LUM), la couleur ne comprenait aucune information sexuelle. En effet, ces conditions présentaient des visages androgynes dont seule la position des attributs changeait. Ainsi, la distribution des valeurs de luminance et/ou de chrominance était similaire pour tous les visages. Dans les conditions noIAD et ALL, la couleur contenait de l'information sexuelle et pouvait donc être utilisée pour catégoriser le sexe des visages. Une MANOVA fut réalisée afin de comparer les efficacités enregistrées dans les conditions COL et LUM. Le résultat de l'effet principal indique que les participants étaient plus efficaces dans les conditions COL que dans les conditions LUM. Des comparaisons multiples ont également révélées que les efficacités étaient significativement plus élevées dans la condition ALL-COL que dans la condition ALL-LUM, dans la condition noIAD-COL que dans la condition noIAD-LUM et dans la condition IADs-COL que dans la condition IADs-LUM.

Au moins deux possibilités doivent être envisagées pour expliquer ces résultats. La couleur pourrait améliorer les performances des participants soit (1) en facilitant le traitement de bas niveau des visages — via la segmentation des surfaces ou l'amplification du signal en basses fréquences spatiales — soit (2) en facilitant le traitement de plus haut niveau par l'ajout d'informations sexuelles supplémentaires. Compte tenu de la différence significative observée entre la condition IADs-COL et IADs-LUM, le premier facteur pourrait expliquer à lui seul les différences observées dans les autres conditions. Il nous semble toutefois plus probable que les deux facteurs interagissent ou contribuent chacun à divers degrés à cette augmentation de la performance.

5.3.1. Hypothèse sur l'utilisation rapide des informations chromatiques dans la région orale

Nous avons observé, dans l'article 1, de fortes corrélations entre les réponses rapides/correctes et les pixels situés dans la région orale. Compte tenu de la richesse des informations chromatiques contenues dans cette région d'une part, et des recherches, tant en perception des objets qu'en perception des visages (p. ex. Tarr et al., 2001; Gegenfurtner & Rieger, 2000; Holcombe et Cavanagh, 2009) montrant le traitement rapide de cette information d'autre part, nous avons proposé l'hypothèse selon laquelle de telles informations pouvaient être extraites très rapidement par le système visuel. Le but premier de l'article 3 était de tester cette hypothèse. Toutefois, nous n'avons pas rapporté ces résultats dans l'article 3 et avons plutôt misé sur les résultats d'analyses complémentaires concernant la fréquence d'échantillonnage des attributs faciaux. Dans cette section, nous présenterons l'étude 3 sous cette forme originelle et regarderons si l'information chromatique de la bouche est extraite en premier.

Afin de tester cette hypothèse, nous retournerons aux résultats des deux régressions linéaires multiples présentées à la section 4.4. De ces régressions ont découlé deux volumes (128 X 128 X 17) de coefficients de régression: un volume pour les bulles chromatiques et un autre pour les bulles achromatiques. Ces volumes de coefficients de régression, ou volume de classification (VCs), représentent l'association linéaire entre la localisation spatiotemporelle d'une bulle et la justesse de réponse à la tâche de catégorisation sexuelle des visages. En d'autres termes, ils révèlent l'extraction de l'information achromatique et chromatique spatiale à travers le temps.

La Figure 3 présente le VC chromatique (images du haut) et achromatique (images du bas) sous la forme d'une série temporelle de 17 images de classification 2D. Les scores Z sont exprimés selon une échelle de couleur, allant du bleu (scores les plus négatifs) au rouge (scores les plus positifs). Les lignes noires délimitent les scores Z significatifs. Le VC moyen chromatique révèle des voxels significatifs dans la région orale entre 24 et 35 ms et entre 165

et 200 ms. Le VC moyen achromatique montre des voxels significatifs dans les régions orbitales — incluant la région des tempes et des arcades zygomatiques — entre 12 et 200 ms.

Cette analyse indique aussi que les scores Z du VC achromatique sont supérieurs aux scores Z du VC chromatique, et ce, par un facteur de 233%. La Figure 4A présente le sommaire de ces résultats : tous les « clusters » significatifs des deux VCs ainsi que leurs latences respectives sont exprimées sur une seule image. Les pixels verts représentent le domaine achromatique tandis que les pixels rouges représentent le domaine chromatique. La latence et la position *xy* des voxels significatifs chez les observateurs hommes et femmes sont représentées sur les deux plus petites images de la Figure 4A.

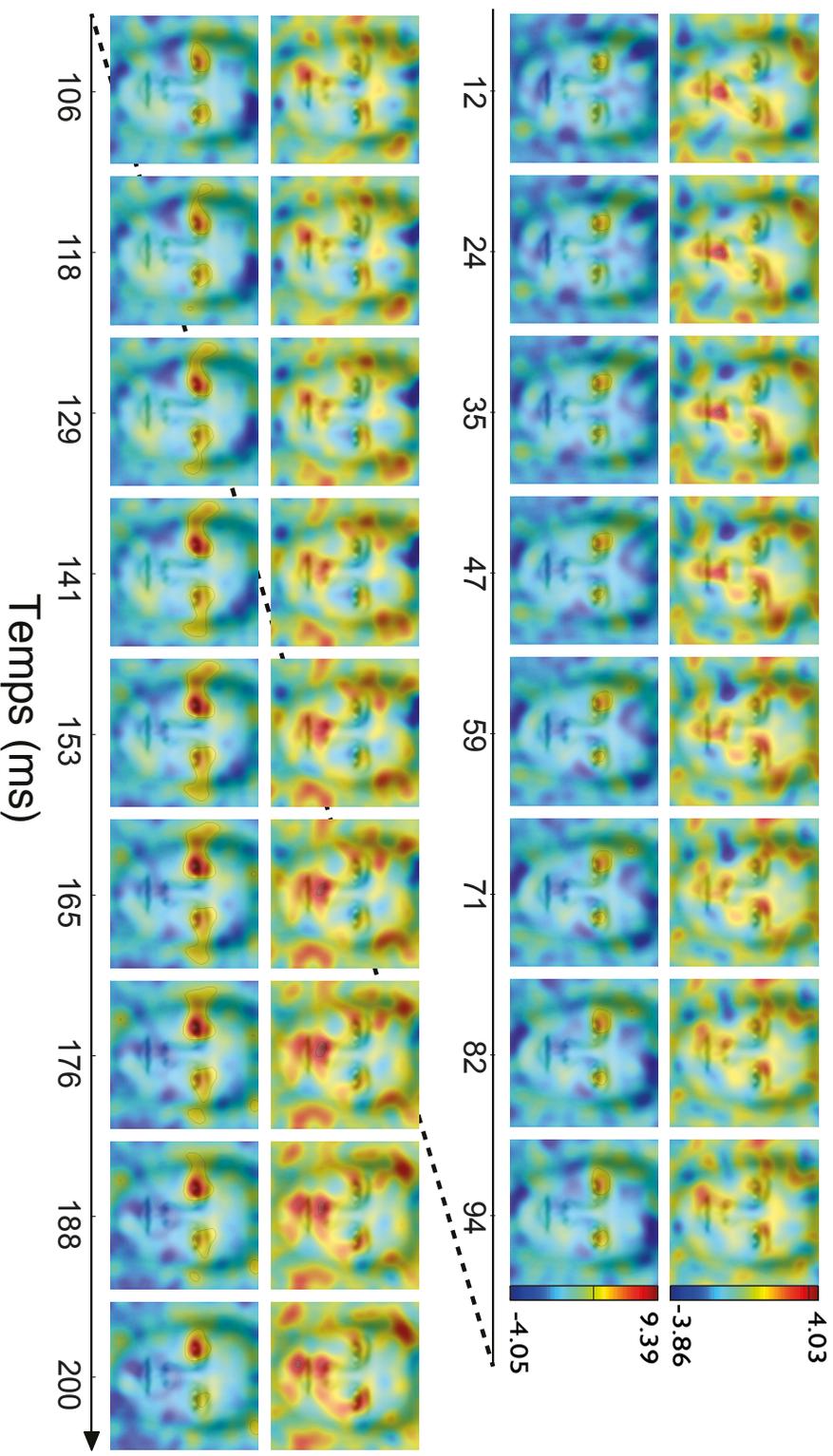


Figure 3. VC chromatique (images du haut) et achromatique (images du bas) juxtaposés à un visage androgyne moyen et exprimés sous la forme d'une série temporelle de 17 images de classification 2D. Les scores Z sont exprimés selon une échelle de couleur (voir barres de couleur en haut à droite), allant du bleu (scores les plus négatifs) au rouge (scores les plus positifs). Les lignes noires délimitent les scores Z significatifs.

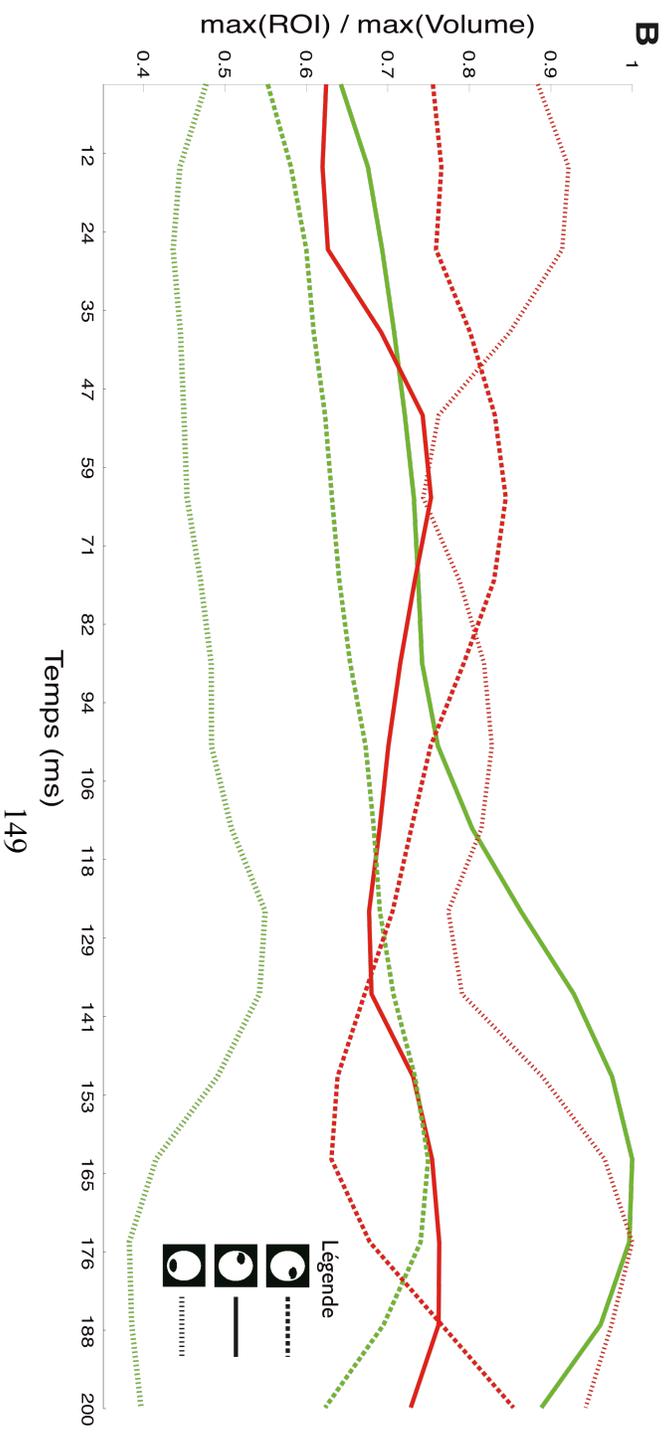
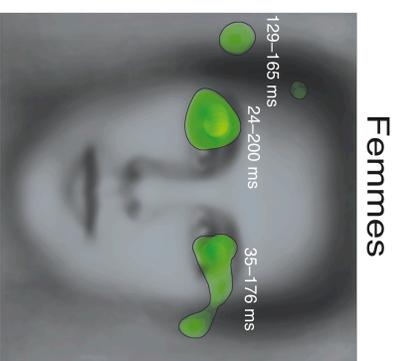
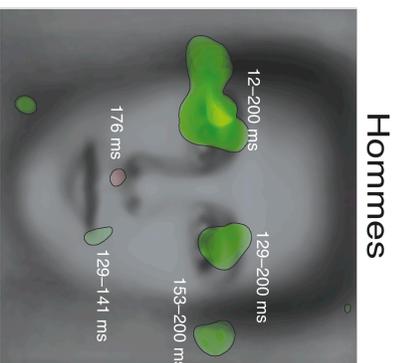
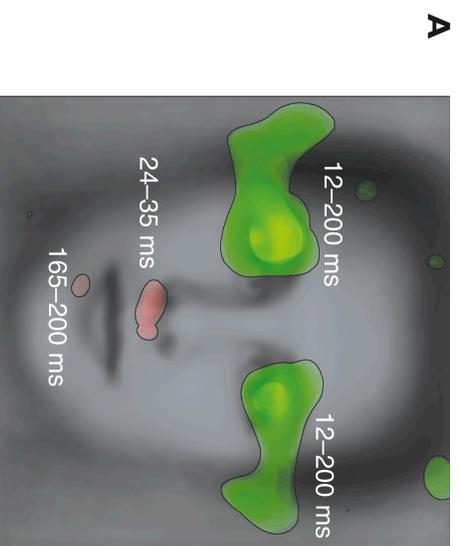


Figure 4. A) Sommaire des voxels significatifs dans les volumes de classification (VC) chromatique (pixels rouges) et achromatique (pixels verts). La position xy et la latence des voxels significatifs sont exprimés sur une seule image. L'image de gauche représente les VC de groupe, tandis que les deux autres images désignent le VC moyen des observateurs hommes (centre) et femme (droite). B) Courbes exprimant le ratio entre le Z_{max} calculés dans trois régions d'intérêts — c.-à-d. l'œil à la gauche de l'image (ligne pleine), l'œil droit (grand pointillé) et la bouche (petit pointillé) — et le Z_{max} du VC correspondant en fonction du temps.

Les résultats des VC révèlent une utilisation précoce des informations chromatiques dans la région orale et une utilisation précoce des informations achromatiques dans les régions oculaires. Il est toutefois difficile de tirer des conclusions claires de ces résultats compte tenu de la différence de magnitude entre les scores Z du VC chromatique et achromatique. Afin de compenser pour cette différence, nous avons divisé les scores Z de chaque VC en fonction de leurs scores Z maximaux. La Figure 4B présente les scores Z maximaux dans trois régions d'intérêts (c.-à-d. les deux régions orbitales et la région orale) en fonction du temps. Cette analyse montre que les scores Z maximaux dans la région orale sont atteints très tôt (12 ms) et très tard (176 ms), tandis que les scores Z maximaux dans la région de l'œil gauche sont atteints vers 165 ms.

Bien que des études aient répertorié des différences de couleur entre la peau des hommes et des femmes d'un même groupe ethnique il y a plusieurs années (voir Frost, 1988), ce n'est que tout récemment que des études ont suggéré que les observateurs humains pouvaient effectivement utiliser les différences de couleur dans le visage en tant que signal dimorphique sexuel (Nestor et Tarr, 2008a). Nos résultats fournissent des preuves additionnelles de cette capacité. En particulier, il s'agit d'une des premières indications à l'effet que des informations chromatiques mêmes *isoluminantes* dans la région orale peuvent être utilisées à des latences très brèves. Dans les sections subséquentes, nous aborderons certaines hypothèses quant au

rôle facilitateur de la couleur et quant à l'importance de la région orale pour la catégorisation du sexe des visages.

5.3.2. Rôle de la couleur et rapidité de traitement

Plusieurs études en perception des visages et en perception des objets suggèrent que la couleur serait un indice robuste et rapide à traiter. Par exemple, Moutoussis et Zeki (1997a) ont montré que les changements de couleur d'un stimulus périodique simple (e.g. un échiquier) étaient perçus 63 ms avant les changements d'orientation, et 115 ms avant les changements de direction du mouvement. Comme ces données se réfèrent aux percepts des participants (i.e. à leur expérience consciente) et non pas aux processus sous-jacents menant à ces percepts, on ne peut qu'en conclure que la couleur est priorisée à un certain niveau de traitement. Une preuve plus directe du traitement rapide de la couleur et de son importance dans les traitements de bas niveau vient de Gegenfurthner et Rieger (2000). Ces auteurs ont observé que la couleur améliorait significativement la reconnaissance de scènes naturelles dont le temps de présentation était très bref (16 ms) comparé aux mêmes scènes en tons de gris. Cet effet facilitateur de la couleur par rapport à la luminance n'était pas présent à des latences plus élevées. Les auteurs ont également confirmé que cet effet facilitateur de la couleur n'était pas simplement dû à la sensibilité accrue du système visuel aux basses fréquences spatiales des variations chromatiques. Les auteurs en ont conclu que l'effet facilitateur de la couleur se produirait à un très bas niveau de traitement, soit au niveau de la segmentation de l'image en surfaces homogènes.

L'effet rapide et facilitateur de la couleur a également été rapporté dans la littérature sur la reconnaissance des visages. Par exemple, les effets bénéfiques de l'amplification du signal chromatique enregistrés par Lee et Perrett (1997) ont été observés exclusivement à des temps de présentation très brefs (entre 67 et 100 ms). Des effets similaires ont été observés dans des tâches de catégorisation du sexe des visages. À titre d'exemple, Tarr et al. (2001) ont présenté des visages dont le canal *a* (ou rouge-vert) était soit représentatif d'un visage d'homme, soit représentatif d'un visage de femme. Les autres canaux (i.e. *L* et *b*) ne comportaient que des

informations neutres sur le plan sexuel. Les visages étaient donc androgynes à tous égards sauf au niveau du canal rouge-vert. Ces auteurs ont trouvé que ces variations de couleurs avaient un impact significatif sur la justesse de réponse seulement lorsque les visages étaient présentés à des temps de présentation très brefs (< 100 ms). Ils en ont conclu que les variations de tonalités dans le canal rouge-vert pouvaient être extraites rapidement pour catégoriser le sexe des visages.

Le fait que l'effet facilitateur de la couleur rapporté dans ces études n'était significatif qu'à des latences très rapides et pour des durées très brèves n'implique pas nécessairement que les indices chromatiques sont plus robustes et priorisés par le système visuel. En fait, les résultats de plusieurs études suggèrent que la couleur transmet beaucoup moins d'information que la luminance pour la reconnaissance des objets et des visages (voir Dupuis-Roy, et al., 2009; Nestor et Tarr, 2008a; Gegenfurthner et Rieger, 2000). Une possibilité est que les indices chromatiques sont utilisés très tôt alors que les indices achromatiques ne sont pas encore accessibles par le système visuel.

Ainsi, l'effet facilitateur de la couleur pourrait survenir très rapidement et très tôt dans le traitement de l'information et pourrait également intervenir à un plus haut niveau de traitement en fournissant des informations relatives à la tâche. Nos résultats indiquent plus spécifiquement que les variations chromatiques dans la région de la bouche seraient importantes pour la catégorisation du sexe des visages. La prochaine section proposera certaines hypothèses qui pourraient expliquer l'importance de ce dimorphisme sexuel.

5.3.3. Hypothèses sur l'importance de la région orale

D'abord, au niveau statistique, il semble que la région orale contienne effectivement des informations chromatiques et achromatiques pertinentes à la catégorisation sexuelle. En effet, des analyses statistiques (voir Dupuis-Roy et al., 2009; Nestor et Tarr, 2008a) effectuées sur des visages dans l'espace des couleurs CIE Lab indiquent la présence d'informations sexuelles dans la région orale, et ce, surtout dans les axes *L* et *a* (voir section 1.2.1.1). Les résultats de

Nestor et Tarr (2008b) suggèrent aussi que les observateurs humains incorporent ces différences chromatiques et achromatiques dans leurs représentations du sexe des visages. En 2009, Russell a présenté des résultats d'une analyse montrant que le contraste achromatique local entre les yeux et la peau environnante ainsi que le contraste entre les lèvres et la peau avoisinante étaient plus prononcés chez les femmes que chez les hommes d'un même groupe ethnique (voir section 1.2.1.2). L'intensité de ces contrastes locaux était également corrélée significativement avec les jugements de féminité/masculinité des visages. « L'Illusion du Sexe » que Russell a présentée en 2003 suggère également que ces contrastes interagiraient au niveau de la perception du sexe des visages (voir Figure 5A à la section 1.2.1.2).

Stephen et McKeegan (2010) ont testé l'effet de la manipulation du contraste local entre les lèvres et le reste du visage — dans l'espace des couleurs CIE Lab — sur l'attractivité et la masculinité/féminité des visages d'hommes et de femmes. Les résultats ont montré que l'augmentation du contraste local des lèvres dans l'axe rouge-vert maximisait l'attractivité des visages d'hommes et de femmes, ainsi que la féminité des visages de femmes. La masculinité des visages d'hommes était plutôt liée à une diminution du contraste local dans l'axe rouge-vert. Ensemble, ces études suggèrent (1) qu'il existe des informations chromatiques et achromatiques sexuelles dans la région orale et (2) que ces variations sont importantes au niveau des représentations visuelles du sexe des visages. Comment expliquer alors les résultats de nos analyses supplémentaires (voir section 5.3.1.1) montrant que seul le volume de classification chromatique contenait des pixels significatifs dans la région orale ?

D'abord, il faut noter une fois de plus que l'absence de résultat significatif n'implique pas nécessairement que les participants n'utilisaient pas les informations achromatiques de cette région faciale. Ces résultats indiquent plutôt que l'information achromatique dans les régions orbitales est beaucoup plus importante que celle contenue dans la région orale pour cette tâche. Qui plus est, Stephen et McKeegan (2010) ont formulé certaines hypothèses quant à l'utilisation des variations chromatiques dans la région orale. Ces auteurs évoquent la possibilité que la rougeur des lèvres provienne non seulement de représentations culturelles mais aussi d'un signal biologique. En effet, comme le rougissement — cosmétique — des

lèvres a été observé dans des cultures géographiquement et temporellement très éloignées (Corson, 1972; Schaffer, 2007), la possibilité d'une cause physiologique commune doit être envisagée. Selon ces auteurs, la rougeur des lèvres pourrait, par exemple, signaler une plus grande oxygénation sanguine elle-même associée (1) au taux d'œstrogène, (2) à l'excitation sexuelle, ou (3) à la santé cardio-respiratoire. Si les facteurs (2) et (3) ne sont pas exclusivement associés à l'un ou l'autre des deux sexes, il est tout de même possible que l'accentuation de la rougeur des lèvres soit plus prononcée chez la femme que chez l'homme étant donné que les lèvres des femmes ont tendance à être plus saillantes que celles des hommes en raison d'un taux d'œstrogène plus élevé (Farkas, 1981; Fink and Neave, 2005).

5.3.4. Chromatique vs achromatique

Quelle est l'importance relative des indices chromatiques et achromatiques pour la catégorisation du sexe des visages? À première vue, les résultats de la deuxième et de la troisième étude semblent contradictoires à cet effet. D'une part, les visages en couleurs ont mené à des efficacités significativement plus élevées que les efficacités enregistrées avec les visages en tons de gris. D'autre part, les scores Z observés dans le volume de classification achromatique étaient largement supérieurs (par un facteur de 233%) à ceux relevés dans le volume de classification chromatique. Une explication simple est que les indices chromatiques fournissent moins d'informations que les indices achromatiques, et ce, autant pour les visages que pour d'autres objets visuels (voir Gegenfurtner et al., 2000). Les statistiques d'image calculées dans l'espace CIE Lab et présentées à la section 1.2.1.1 confirment également ce déséquilibre. Les indices chromatiques pourraient donc être importants au niveau de la représentation visuelle sans toutefois qu'ils ne contiennent beaucoup d'information sexuelle. Ceci expliquerait pourquoi ils sont associés à des efficacités élevées et qu'ils sont plus difficiles à étudier dans un contexte où seulement une faible proportion de l'information chromatique est échantillonnée et où d'autres indices plus saillants sont présentés simultanément.

5.4. Mécanisme d'encodage des attributs faciaux

À l'origine, l'article 3 avait pour but de tester l'hypothèse de l'utilisation précoce des informations chromatiques dans la région orale. Après avoir calculé les volumes de classification chromatique et achromatique, nous avons examiné les séries temporelles moyennes dans les régions oculaires et orales (voir Figure 3 à la section 5.3.1). Les séries temporelles chromatiques et achromatiques de groupe obtenues dans chacune de ces régions présentaient des ondulations transitoires et des scores Z de magnitudes très différentes. Comme nous nous intéressions à la latence de ces ondulations sans égard à leur différences en matière de scores Z , nous avons employé une analyse temps-fréquence par ondelettes. Les résultats présentés aux Figures 2 et 3 de l'article 3 (section 4.4.1) indiquent que l'extraction des informations chromatiques de la bouche et de l'oeil situé à la droite de l'image oscille à des latences très faibles. Ce résultat est cohérent avec notre hypothèse de l'utilisation précoce de l'information chromatique dans la bouche. Ces analyses révélaient également des résultats clairs sur le plan fréquentiel: la fréquence d'extraction des informations chromatiques et achromatiques étaient la même pour une région faciale donnée, tandis que les fréquences d'extraction propres à chacune de trois régions faciales étaient distinctes, allant de 6 à 10 Hz. En d'autres mots, chaque région faciale renvoyait à une fréquence d'extraction distincte, tandis que les informations chromatiques et achromatiques qu'elles contenaient étaient extraites par une même fréquence. L'analyse des fréquences liées au "peak" d'activité dans les six temps-fréquences de chacun des participants confirme ce patron d'encodage fréquentiel. Ceci suggère la présence d'un multiplexage fréquentiel, soit le codage fréquentiel de différentes informations dans un même signal.

Plusieurs études sur les cortex sensoriels utilisant une combinaison des potentiels évoqués locaux et de "spike-train" ont démontré que différents attributs d'un même stimulus peuvent être multiplexés par différentes propriétés des décharges neuronales, telles que la fréquence, la phase et sa cohérence (voir Koepsell et al., 2010). De récentes études en électrophysiologie ont aussi découvert la présence d'un multiplexage fréquentiel à de plus hauts niveaux de traitement, soit au niveau cognitif et perceptif (Schyns, Thut et Gross, 2011; Smith, Gosselin et Schyns, 2005, 2006, 2007; Thut et al., 2011; Romei, Driver, Schyns et Thut, 2011). Par exemple, Smith et al., (2006) et Romei et al., (2011) ont montré que

l'information visuelle locale et globale était multiplexée à différentes bandes de fréquences, soit la bande thêta (~4 Hz) et bêta (~12 Hz) respectivement. Aussi, Schyns, Thut et Gross (2011) ont observé que la puissance des mesures EEG dans la bande thêta codait l'utilisation de l'information dans la zone orale pendant une tâche d'identification de l'expression faciale tandis que la puissance dans la bande bêta codait l'utilisation de l'information dans la zone oculaire.

À la différence des études électrophysiologiques sur les oscillations transitoires, celles observées dans notre étude ne peuvent être attribuées à un niveau de traitement particulier. Elles peuvent tout autant refléter une combinaison de traitements attentionnels, mnémoniques ou perceptifs qui dépendent de la tâche et qui sont en phase avec la présentation du stimulus. Ceci étant dit, les fréquences enregistrées dans notre étude sont cohérentes avec celles rapportées dans d'autres tâches perceptives (voir VanRullen, 2005; VanRullen, 2006; MacDonald, Cavanaugh, et VanRullen, 2013; Blais, Arguin et Gosselin, 2013). Aussi, la mesure de justesse de réponse utilisée dans notre étude pour les analyses temps-fréquences a été prélevée plusieurs secondes après la disparition du stimulus, ce qui a pu promouvoir la contribution de la mémoire à court terme. Sur ce point, les oscillations dans la bande thêta enregistrées dans l'hippocampe et dans d'autres régions corticales ont été répétitivement associées à la mémoire de travail, l'apprentissage et l'attention (Busch et VanRullen, 2010; Staudigl et Hanslmayr, 2013; Tesche et Karhu, 1999; Kahana, Seelig et Madsen, 2001; Lisman et Idiart, 1995). Bien que l'état des connaissances actuelles sur le lien entre les oscillations enregistrée avec l'EEG et celles observées au niveau comportemental ne permette pas de dresser des conclusions claires (voir VanRullen et Koch, 2003), il est possible que le multiplexage fréquentiel observé dans notre étude reflète l'encodage des régions faciales et le maintien de ces informations en mémoire à court terme. Des recherches plus approfondies seraient nécessaires afin d'éprouver cette hypothèse et clarifier le lien entre les oscillations comportementales et neurales.

5.5. Importance relative des indices relationnels de second ordre

Les indices relationnels de deuxième ordre font principalement référence aux distances et aux angles entre les principaux attributs d'un visage. Dans la présente thèse, nous nous sommes intéressés aux distances interattributs. Ce choix se justifie principalement par l'importance de ces indices dans la littérature en perception des visages, mais aussi par la facilité à mesurer et à manipuler ces indices.

5.5.1 Importance relative des distances interattributs

L'hypothèse selon laquelle les relations spatiales entre les principaux attributs faciaux (c.-à-d. les distances interattributs, ou DIA) représenteraient un indice stable et invariant à la distance pouvant servir à la reconnaissance faciale a d'abord été proposée par Galton (1879) et a depuis été corroborée par une série de travaux. Par exemple, plusieurs études ont montré que le traitement des DIA était davantage affecté par l'inversion des visages (ou rotation à 180°) que le traitement de leur forme ou de leurs propriétés de surface (Barton et al., 2001; Freire et al., 2000; Le Grand et al., 2001; Rhodes, Hayward & Winkler, 2007; Sergent, 1984; pour une revue de littérature récente voir Rossion, 2008, 2009, et pour l'hypothèse inverse voir McKone et Yovel, 2009). Certaines études suggèrent aussi l'importance de ces indices pour la catégorisation du sexe des visages (e.g. Campbell, Wallace & Benson, 1996; Campbell et al., 1999; Burton, Bruce, et Dench, 1993). Ainsi, plusieurs études appuient l'hypothèse selon laquelle le traitement des DIA serait crucial à la perception des visages, incluant la catégorisation du sexe. Cependant, l'étude de Taschereau et al. (2010) a récemment démontré que plusieurs de ces études avaient testé des visages dont les DIA avaient été amplifiées par 376%. De plus, ils ont observé que lorsque les participants devaient identifier des visages sur la seule base de DIA *réalistes*, les performances étaient proches du hasard. Cette étude montre donc l'importance d'utiliser des distances interattributs qui soient représentatives de celles observées dans l'environnement. Ceci implique également l'utilisation de covariances réalistes entre les distances. En effet, une des causes importante de l'amplification des DIA dans les études antérieures vient de la manipulation indépendante et individuelle des attributs. Par exemple, en modifiant exclusivement le positionnement de la bouche, Haig (1984) se trouva

en fait à modifier les distances bouche-nez, bouche-yeux et bouche-sourcils, affectant ainsi presque toute la matrice de covariance entre ces distances.

Dans ce contexte, nous nous sommes demandé si ces faibles performances pouvaient être expliquées non pas par une incapacité d'utiliser ces indices efficacement, mais plutôt par le peu d'information contenue dans ces indices. Un rejet définitif de l'hypothèse sur l'importance des DIA nécessitait donc de départager ces deux possibilités. C'est ce que nous avons testé dans l'article 2. Pour ce faire, nous avons évalué l'efficacité — une mesure de performance qui compense pour la faiblesse de l'information disponible — à l'utilisation des DIA dans une tâche de catégorisation du sexe des visages. Afin de maximiser la validité externe, les DIA manipulées étaient réalistes en ce sens qu'elles respectaient la distribution et la matrice de covariance des DIA observées dans un large échantillon de visages (N=515).

Les résultats de l'article 2 indiquent que les efficacités observées dans la condition où seules les DIA réalistes variaient (condition IAD) étaient significativement inférieures aux efficacités enregistrées dans les conditions où tous les autres indices faciaux variaient sauf les DIA (condition no-IAD) et où tous les indices faciaux variaient (condition ALL). De plus, l'analyse de l'observateur idéal indique que les DIA renferment moins d'information diagnostique du sexe que tous les autres indices faciaux combinés. Ainsi, ces résultats réfutent l'hypothèse selon laquelle les DIA seraient cruciales à la catégorisation du sexe des visages. Il est à noter que cette conclusion ne repose pas simplement sur la différence significative d'efficacité entre les conditions, mais aussi sur le fait que les efficacités enregistrées dans la condition IAD sont faibles si on les compare aux efficacités rapportées dans d'autres études sur la perception des visages.

En s'appuyant sur les travaux portant sur l'intégration de multiples informations dans une ou plusieurs modalités sensorielles (e.g., Hillis et al., 2004; Young et al., 1993; Alais et Burr, 2004; Battaglia et al., 2003; Ernst et Banks, 2002; Gold, Mundy et Tjan, 2012), nous avons formulé l'hypothèse selon laquelle l'observateur humain utiliserait moins les DIA puisqu'elles contiennent moins d'informations sexuelles. En fait, ces études suggèrent que les

humains pondèrent de façon quasi-optimale les différents indices en fonction de leur informativité. En plus de suggérer que l'observateur humain est adapté à son environnement en ce qui concerne l'information visuelle, cette hypothèse prédit également que la pondération des indices pourrait changer selon leur informativité. Dans une telle optique, une légère amplification d'un signal facial pourrait, au fil des essais, accroître son utilisation. Il serait intéressant de tester ces hypothèses dans le contexte des indices faciaux chromatiques et achromatiques, d'autant que certains chercheurs prêtent encore trop peu d'attention au réalisme des signaux faciaux qu'ils testent.

5.5.2 Importance des autres indices relationnels de deuxième ordre

Nous n'avons testé qu'un seul type d'indice facial relationnel de deuxième ordre, à savoir les distances interattributs. Les conclusions de l'article 2 ne peuvent donc pas être généralisées à l'ensemble des indices relationnels. À ce stade, il est toutefois légitime de se questionner sur l'importance des autres types d'indices relationnels de deuxième ordre, et sur la portée de nos conclusions. Compte tenu de la grande quantité d'articles scientifiques sur le sujet, nous nous pencherons exclusivement sur les études ayant évalué l'importance de ces indices à l'aide d'une mesure d'efficacité ou de méthodologies analogues à celles que nous avons utilisées dans cette thèse.

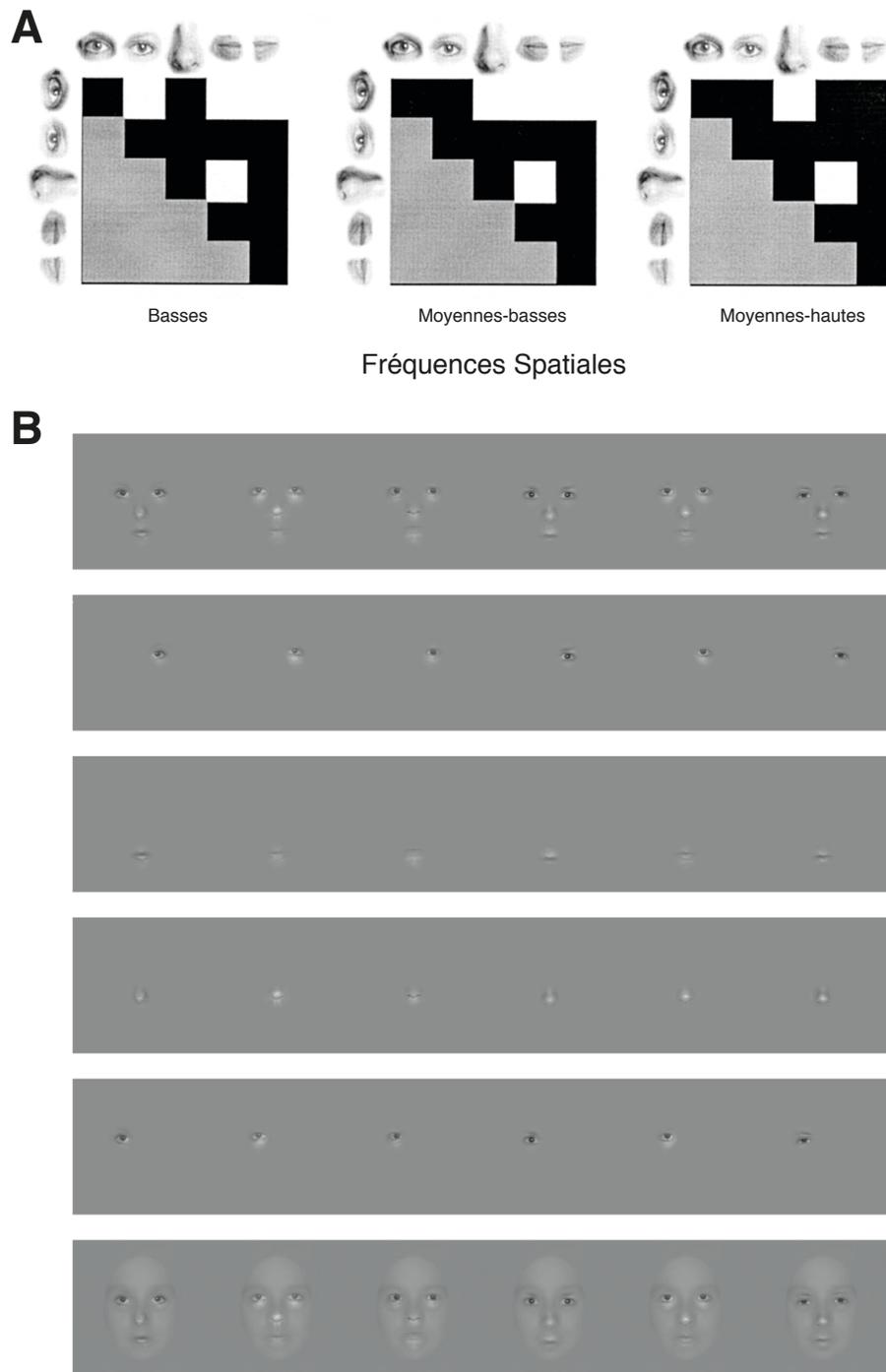


Figure 5. A) Résultat de l'analyse de conjonction du second degré sur les données de la tâche de catégorisation sexuelle des visages (tiré de Schyns, Bonnar et Gosselin, 2002). Les carrés blancs de la matrice de conjonction désignent des conjonctions significatives. B) Les stimuli utilisés dans les six conditions expérimentales de Gold, Mundy et Tjan (2012).

En 2002, Schyns, Bonnar et Gosselin ont testé l'utilisation conjointe de cinq attributs faciaux (œil droit, œil gauche, nez, coin droit de la bouche et coin gauche de la bouche) dans des tâches de catégorisation sexuelle, d'identification et de discrimination de l'expression faciale dans lesquels l'espace xy et les fréquences spatiales étaient échantillonnés avec la technique Bubbles. L'analyse statistique consistait en une régression linéaire multiple entre la conjonction de second degré des masques de bulles — représentant les dix combinaisons possibles de paires d'attributs — et l'exactitude des réponses (voir Figure 5A). Une telle analyse révèle non seulement l'utilisation des distances entre chaque paire d'attributs faciaux, mais aussi l'utilisation d'autres processus d'intégration par paire d'attributs. Les résultats dans la tâche de catégorisation du sexe indiquent la présence de conjonctions significatives ($p < 0.01$, non corrigé pour des tests multiples) de second degré entre l'œil gauche et la bouche en hautes et moyennes hautes fréquences spatiales, ainsi qu'entre le coin gauche de la bouche et le nez pour toutes les fréquences spatiales. Plusieurs conjonctions significatives ont également été trouvées dans les deux autres tâches. Bien que l'analyse utilisée était quelque peu libérale quant à l'erreur de première espèce, ces résultats suggèrent l'intégration spatiale conjointe des attributs ce qui est cohérent avec l'utilisation de certaines DIA. Quoiqu'il en soit, les données rapportées ne permettent pas d'estimer la taille de l'effet et donc d'évaluer l'importance de l'utilisation de ces conjonctions par rapport à l'utilisation locale des mêmes informations.

Dans la même veine, Gold, Mundy et Tjan (2012) ont testé l'hypothèse d'une intégration supra-optimale (ou non-linéaire) des principaux attributs faciaux dans une tâche d'identification des visages. Plus spécifiquement, ils ont mesuré l'efficacité à identifier un visage dans cinq conditions expérimentales : une condition où les quatre principaux attributs faciaux étaient présentés ensemble (i.e. œil droit, œil gauche, nez et bouche) et quatre conditions où chacun de ces attributs était présenté individuellement (voir Figure 5B). La logique de leur test était simple : si les observateurs humains combinaient les attributs faciaux de façon non-linéaire, alors l'efficacité dans la condition où tous les attributs étaient combinés devrait excéder la combinaison linéaire des efficacités dans les conditions où les attributs étaient présentés individuellement. Ce test devrait tenir compte de l'utilisation des DIA mais aussi de tous les autres traitements non-linéaires sur les principaux attributs. Les résultats

d'une analyse avec un intégrateur bayésien infirment cette hypothèse, suggérant plutôt que les différents attributs du visage étaient intégrés quasi-optimalement (donc linéairement). Gold et al. (2012) ont également testé l'effet de l'inversion des visages sur l'intégration des informations et ont montré une diminution significative de l'intégration des parties du visage lorsque présenté à l'envers. Ils en concluent que l'inversion du visage ne change pas la nature linéaire du traitement de l'information, mais qu'elle diminue quantitativement l'efficacité d'intégration de ces parties. Ce résultat dévoile toutefois une faiblesse de leur modèle : l'intégrateur bayésien ne tient pas compte de la perte d'efficacité liée à l'intégration des parties du visage. Supposons un observateur avec des ressources attentionnelles limitées et dont l'efficacité diminue en fonction de la quantité d'éléments à intégrer. Dans un tel cas, il y aura intégration sous-optimale dans la condition où toutes les parties du visage sont affichées. Supposons encore que cet observateur utilise tout de même partiellement le peu d'information contenue dans les DIA. Si tel est le cas, l'efficacité prédite par l'intégrateur linéaire pourrait encore être supérieure à l'efficacité observée, et ce, malgré l'utilisation des DIA. En conclusion, le modèle utilisé par Gold, Mundy et Tjan (2012) pour tester l'hypothèse des processus configuraux est très conservateur et ne peut exclure la présence d'une intégration non-linéaire tant pour les visages à l'envers qu'à l'endroit.

En 2004, Sekuler, Gaspar, Gold et Bennett ont utilisé la "reverse correlation" pour évaluer l'impact de l'inversion des visages sur le traitement des informations faciales locales. Une des hypothèses dominantes sur l'effet d'inversion suppose que l'inversion d'un visage affecte la nature même du traitement de l'information faciale : à l'endroit, les visages seraient traités de façon configurale et analytique, tandis qu'à l'envers, ils seraient traités de façon analytique seulement (voir Tanaka & Farag, 1993; Valentine, 1988). Une telle hypothèse prédit donc une plus grande difficulté à traiter les relations spatiales des visages lorsqu'ils sont orientés à 180 degrés. Une déclinaison formelle de cette hypothèse prédit que l'inversion des visages devrait altérer tous les traitements non-linéaires des informations faciales. Une telle définition ne présuppose pas l'importance de zones particulières du visage, et englobe à la fois les DIA et toutes autres formes de non-linéarité. En ce sens, elle se veut donc très générale et inclusive. Sekuler et al. (2004) ont comparé les images de classification (ICs) dans la

condition où les visages étaient à l'envers avec les ICs dans la condition où les visages étaient présentés à l'endroit. Les résultats de cette analyse révèlent la présence de pixels significatifs dans la région oculaire, et ce, dans une même proportion pour les deux conditions. Ceci suggère donc l'utilisation d'une même stratégie perceptive dans les deux conditions. Ces auteurs ont également comparé les efficacités mesurées chez leurs participants aux efficacités prédites par un modèle linéaire tenant compte du bruit interne. Si les participants intégraient différentes parties du visage de façon non-linéaire, alors l'efficacité observée devrait dépasser l'efficacité prédite. Plus encore, si la contribution de ces processus non-linéaires avait changé d'une condition à l'autre, l'adéquation entre les efficacités prédites et observées aurait aussi dû changer. Leurs résultats montrent que les efficacités observées dans les deux conditions étaient légèrement supérieures aux efficacités prédites, ce qui suggère une faible utilisation du traitement non-linéaire de l'information faciale, et ce, dans les deux conditions. Les auteurs en concluent que l'inversion des visages diminue l'efficacité du traitement des informations faciales sans toutefois affecter la nature de ce traitement. Ces conclusions sont cohérentes avec nos résultats montrant la faible importance des distances interattributs.

Une autre approche testant l'importance des traitements non-linéaires (ou "holistique") consiste à mesurer la proportion de variance des performances d'identification des visages qui est explicable par l'effet composite des visages (voir section 1.2.2). La logique sous-tendant ce test est simple: si les traitements holistiques sont importants pour l'identification des visages, alors les participants ayant de fortes performances dans cette tâche devraient également présenter un effet composite fort. Inversement, les participants ayant de faibles performances devraient présenter un effet composite faible. Konar, Bennett et Sekuler (2010) ont testé cette hypothèse dans deux expériences d'identification des visages incluant un total de 125 participants. Ils ont obtenu des corrélations non significatives entre deux mesures standard de l'effet composite et les performances dans une tâche d'identification des visages. Plus encore, des simulations tenant compte de l'erreur de mesure ont permis d'évaluer à 2.2% la part de la variance de la performance expliquée par l'effet composite. Ce résultat est aussi cohérent avec nos résultats indiquant une faible taille d'effet liée à l'utilisation exclusive des DIA.

En réponse à l'étude de Konar et al. (2010), Richler, Cheung, et Gauthier (2011) ont retesté cette hypothèse à l'aide d'un index supplémentaire de la performance d'identification des visages, le Cambridge Face Memory Test (Duchaine et Nakayama, 2006), ainsi que d'un index supplémentaire de l'effet composite des visages. Donc, en plus d'une tâche d'identification des visages par appariement ("matching-to-sample"), ils ont utilisé le CFMT principalement parce que ce test repose sur la mémoire des visages, ce qui est plus représentatif des conditions naturelles d'identification que ne l'est la tâche d'appariement de Konar et al. (2010). Contrairement à Konar et al. (2010) dont l'effet composite reposait sur deux types de combinaisons entre le haut et le bas des visages appris et testés, l'expérience de Richler et al. (2011) misait sur l'ensemble des combinaisons possibles (voir Figure 6) ce qui leur permettait de calculer l'interaction entre l'alignement du haut et du bas des visages d'une part et la congruence entre la réponse du participant dans la tâche de l'effet composite — soit une tâche où il faut indiquer si le haut du visage test est *pareil ou différent* au haut du visage appris — et le haut et le bas des visages testés d'autre part. La Figure 6 illustre le design complet (toutes les options) et partiel (options encadrées). Selon ces auteurs, un tel design tient compte des biais de réponse, à savoir la tendance à répondre 'pareil' lorsque le bas du visage testé est pareil au bas du visage appris, et à répondre 'différent' si le bas du visage testé est différent du bas du visage appris. Trente-huit participants ont complété les deux tâches d'identification ainsi que toutes les conditions associées au design complet dans la tâche pareil-différent utilisée afin d'estimer l'effet composite. Les résultats du design partiel répliquent les résultats non-significatifs de Konar et al. (2010) tandis que les résultats du design complet indiquent qu'en moyenne 12.55% de la performance d'identification était explicable par l'effet composite tel qu'indexée par l'interaction Alignement X Congruence. Ensemble, les résultats de Konar et al. (2010) et Richler et al. (2011) indiquent que l'effet composite ne rend pas compte d'une forte proportion de la variance des performances individuelles et montrent que la magnitude de cet effet dépend de la tâche. Ces résultats sont contraires à l'idée selon laquelle les processus configuraux et holistiques sont cruciaux pour la reconnaissance des visages.

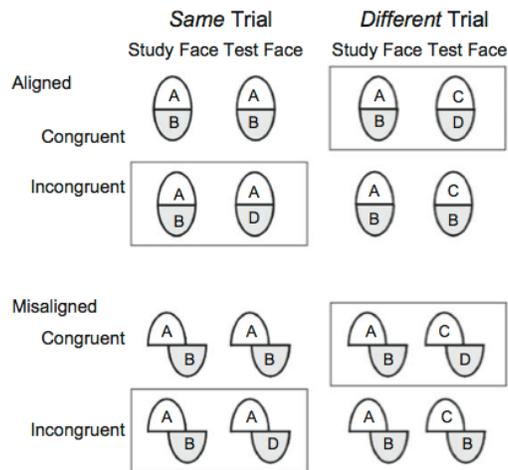


Figure 6. Le design complet (toutes les conditions illustrées) et partiel (seulement les conditions encadrées) de la tâche pareil-différent relative à l'effet composite (tiré de Richler, Cheung, et Gauthier, 2011)

À l'exception des études sur l'effet composite, les études rapportées ci-dessus conçoivent les indices relationnels comme des informations pouvant être extraites par des processus d'intégration non-linéaire. Une telle définition est tout à fait cohérente avec notre formulation des indices relationnels de deuxième ordre et a le mérite d'être formalisable et directement testable. L'ensemble des résultats discutés suggère que l'identification d'un visage et sa catégorisation sexuelle impliquent à la fois des traitements linéaires et non-linéaires de l'information, et que ces derniers sont d'une importance secondaire. Notez toutefois que ces conclusions se rapportent aux conditions expérimentales standard utilisées dans la littérature sur les visages et que la modification des conditions expérimentales pourrait avoir un impact sur l'importance relative des traitements non-linéaires. Par exemple, une étude récente de Yang, Shafai et Oruc (2014) suggère qu'il y a une augmentation du traitement non-linéaire pour les visages vus de très près (à moins de 1.25 m).

5.6. Limites et perspectives futures

Toute entreprise scientifique comporte des limites. Dans cette section, nous critiquerons les méthodologies employées, à commencer par la technique Bubbles, et tenterons de cerner les limites générales de la thèse.

5.6.1. Critiques et limites de la technique Bubbles

Une première critique pouvant être adressée à la technique Bubbles concerne l'emploi d'un modèle linéaire (c.-à-d. la régression linéaire multiple). Une telle analyse est inapte à dévoiler, par exemple, les non-linéarités possiblement impliquées dans l'utilisation simultanée de plusieurs régions faciales. En effet, une telle analyse linéaire serait incapable de révéler l'utilisation de certains signaux nécessitant la présence disjonctive ou conjonctive de deux informations, comme un clin d'œil. En effet, la présence d'un clin d'œil est évaluée selon l'opérateur logique *ou exclusif* : œil droit ouvert et œil gauche fermé, ou œil droit fermé et œil gauche ouvert. Notez que cette limite s'applique aussi à la « reverse correlation » et que des analyses non-linéaires ont déjà été réalisées dans des expériences avec la technique Bubbles (e.g. Schyns, Bonnar & Gosselin, 2002). D'ailleurs, nous avons également testé, dans l'article 1, une hypothèse locale impliquant l'utilisation non-linéaire (ou conjonctive) de l'œil et du sourcil. Cependant, l'essentiel de nos conclusions sont limitées par l'emploi de ce modèle linéaire.

Une autre critique concerne la vraisemblance des visages créés par la technique Bubbles (employée dans l'espace xy) et la possibilité que de tels visages induisent certaines stratégies atypiques (voir Murray & Gold, 2004 pour une critique détaillée). À ce sujet, plusieurs contre-arguments ont été proposés. D'abord, le bruit multiplicatif, c.-à-d. qui dépend de l'intensité du signal, est inhérent aux processus stochastiques impliqués dans la transmission de l'influx nerveux (e.g. Doshier et Lu, 1999). En ce sens, un tel bruit multiplicatif contamine même les expériences n'introduisant aucun bruit dans leurs stimuli. Ensuite, les objets visuels sont, dans la vie courante, rarement perçus dans leur entièreté. Que ce soit un bandeau dissimulant le front, des lunettes fumées voilant les yeux, un foulard masquant le menton et la mâchoire, ou une main couvrant la bouche, les exemples d'occlusion des attributs faciaux abondent. Le soi-

disant « bubble porn » qui fait rage actuellement sur les sites d'échange de photos montre de façon spectaculaire comment le cerveau peut compenser pour le manque d'information généré par un bruit multiplicatif (voir Figure 7).



Figure 7. Un exemple de « bubble porn ».

L'emploi d'un bruit multiplicatif n'est donc pas en rupture complète avec les processus perceptifs courants. L'autre volet de la critique concerne l'utilisation de stratégies atypiques. D'abord, il ne semble y avoir aucune preuve empirique soutenant l'utilisation de stratégies visuelles atypiques avec la technique Bubbles (voir Murray & Gold, 2004; Gosselin & Schyns, 2004 pour une discussion plus approfondie de ce point). Au contraire, les résultats obtenus à partir de cette technique ont été confirmés avec différentes méthodes (e.g. Dupuis-Roy, et al., 2009; Willenbockel et al., 2010; Gaspar, Sekuler et Bennett (2008); Solomon et Pelli, 1994; Fiset et al., 2008; Blais et al., 2009). Ensuite, comme nous l'avons précédemment expliqué, l'utilisation d'un échantillonnage aléatoire devrait, selon les statistiques, atténuer l'utilisation de stratégies locales due à la sélection d'un nombre restreint de zones faciales a priori (c.-à-d. tel que le prône la méthode de l'isolement).

En lien avec la critique précédente, certains ont évoqué la possibilité que la surface révélée par les masques de bulles puisse diminuer l'énergie des visages en basses fréquences spatiales, influençant ainsi les stratégies perceptives utilisées. Cette critique fut adressée dans

le contexte de l'article 1 où nous n'avons pas observé de scores Z significatifs dans certaines grandes zones du visage comme le contour et les cheveux. Comme Schyns, Bonnar et Gosselin (2002) avaient déjà observé des scores Z élevés sur le contour du visage en moyennes basses fréquences spatiales dans une tâche de catégorisation du sexe, cette critique était potentiellement pertinente. Afin d'évaluer les biais d'échantillonnage dans l'expérience de l'article 1, nous avons calculé le spectre fréquentiel des 9000 stimuli utilisés dans notre expérience, et les avons comparé aux spectres fréquentiels des 300 visages originaux non-filtrés. La Figure 8 montre les spectres fréquentiels moyens des visages filtrés (ligne pointillée) et non-filtrés (ligne pleine) par les masques de bulles.

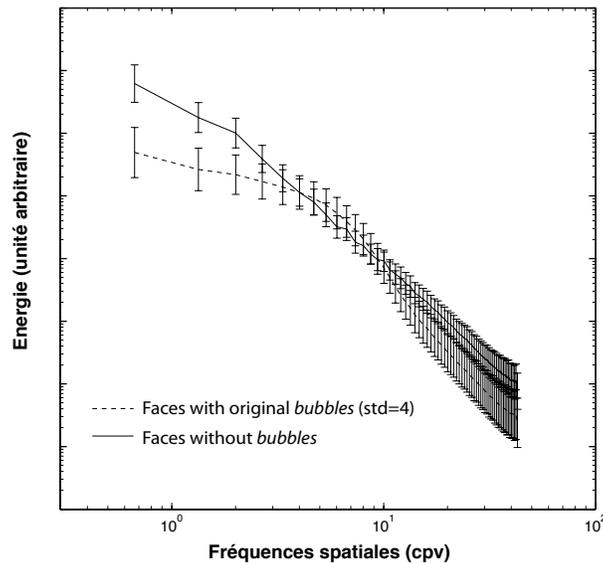


Figure 8. Profil spectral moyen des visages filtrés (ligne pointillée) et non-filtrés (ligne pleine) par les masques de bulles dans l'Article 1. Les barres d'erreur représentent deux écarts-types.

Ces deux courbes se chevauchent sur l'essentiel du spectre sauf sous les 2.63 cycles par visage (cpv). Les résultats de l'étude de Willenbockel et al. (2013) testant des visages d'hommes et de femmes filtrés par des masques de bulles appliqués au domaine des

fréquences spatiales, suggèrent que les fréquences spatiales sous les 2.63 cycles par visage sont très peu utiles à la catégorisation du sexe des visages.

Une critique similaire pourrait être faite pour les bulles spatiotemporelles utilisées dans l'expérience de l'article 3. Dans le cas du filtrage spatial, l'utilisation de plus grandes bulles ($\sigma = 7$ pixels au lieu de 4 dans l'article 1) et d'un plus grand nombre de bulles (97.23 bulles au lieu de 27.06 dans l'article 1) devrait augmenter l'accès aux très basses fréquences spatiales. Quant au filtrage temporel, il est possible que les paramètres d'échantillonnage utilisés dans l'expérience aient diminué l'accès à certaines fréquences temporelles. Pour en simuler l'effet, nous avons simplement convolué 10 000 séries temporelles de bruit blanc par le même filtre gaussien temporel que celui utilisé dans l'expérience. La Figure 9 montre la perte d'énergie en fonction des fréquences temporelles. Le profile spectral résultant est l'équivalent d'un filtre passe-bas avec un « cut-off » de -3 dB à 12.59 Hz et un « roll-off » de -9.45 dB par octave. Cette simulation suggère que le filtrage temporel a pu diminuer l'accès aux fréquences temporelles au-dessus d'environ 13 Hz, soit les fréquences situées dans l'intervalle bêta (15 Hz – 30 Hz). Bien que certaines études comportementales sur les oscillations aient effectivement enregistré des fréquences dans cet intervalle (p. ex. Dehaene, 1993), les fréquences observées dans la vaste majorité des études comportementales se situaient plutôt dans les bandes delta (~0-4Hz), thêta (~4-7Hz) et alpha (~7-14Hz; voir Blais et al., 2009 pour une revue). Quoi qu'il en soit, cet accès limité aux plus hautes fréquences temporelles ne peut pas expliquer le résultat principal des analyses temps-fréquence, à savoir le codage des attributs par une fréquence unique.

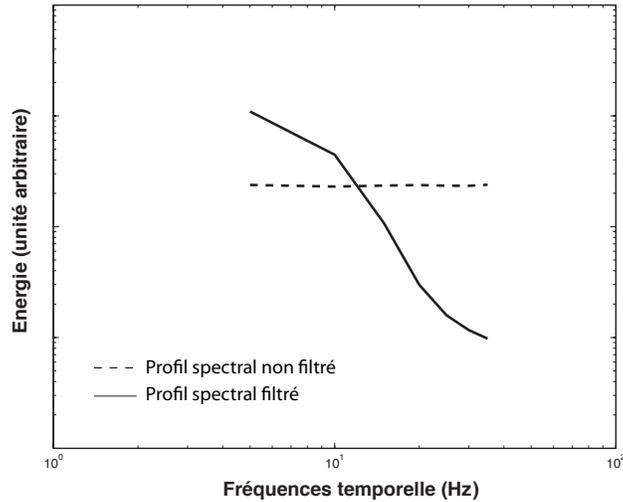


Figure 9. Profil spectral moyen des séries temporelles filtrées (ligne pleines) et non-filtrés (ligne pointillées) par des bulles temporelles ($\sigma_{\text{temps}} = 23.53$ ms).

En conclusion, la technique Bubbles comporte des limites mais, après un examen approfondi de celles-ci, il semble que leur impact sur les phénomènes perceptifs étudiés dans cette thèse soit négligeable.

5.6.2. Limite à la validité externe

Un des premiers objectifs que nous nous sommes fixés dans la thèse était d'étudier l'utilisation de l'information sexuelle faciale dans un contexte plus représentatif des conditions normales de vision. C'est pour cette raison que nous avons étudié l'effet de la couleur et que nous avons utilisé un échantillon important de visages qui furent présentés, au plus, trois fois chacun aux participants. C'est aussi dans cette optique que nous avons utilisé des distances interattributs correspondant à celles de visages réels. Malgré ces efforts, il serait difficile de généraliser nos résultats à des conditions de vision normales. En effet, la thèse se fonde sur le cadre théorique $R \otimes A \approx P$ qui suppose que l'information P provient d'une interaction entre l'information R et A . Dans cette perspective, toute condition visuelle affectant l'accès ou la nature de A peut possiblement affecter R et P . Ainsi, le changement de la distance entre

l'observateur et le visage ou de la luminosité ambiante aurait pu modifier l'accès à l'information A et affecter, en conséquence, le P . Également, certains facteurs comme l'origine ethnique ou même l'angle de présentation des visages (voir Saether et al. 2009) ont pu interagir avec la nature des dimorphismes sexuels faciaux, ce qui a pu, en partie, influencer l'utilisation et la priorisation de certains dimorphismes. Par exemple, l'étude de Saether et al. (2009) suggère que la zone nasale est davantage fixée du regard lorsque le visage est présenté de profil ou dans un certain angle. Advenant que le système visuel s'adapte à son environnement et donc qu'il modifie sa stratégie d'utilisation de l'information en fonction des informations présentes, il est très probable que de tels changements se répercutent sur P . Ainsi, la généralisation de P à des conditions normales de vision nécessiterait que l'information A soit elle-même représentative des conditions visuelles humaines normales.

Nous avons prôné l'utilisation d'une mesure plus réaliste des indices relationnels de deuxième ordre en argumentant que le déplacement des principaux attributs sans s'assurer de la conservation de distances et de covariances réalistes pouvait altérer le signal contenu dans ces indices. Cette conception des indices relationnels de second ordre repose toutefois sur certains *a priori* à savoir que les attributs sélectionnés sont les seuls qui sont importants pour la tâche et que la forme de ces attributs n'interagit pas avec leur positionnement. Nous avons vu, à la section 5.2 de cette discussion, que les attributs manipulés à l'article 2 (i.e. yeux, sourcils, bouche et nez) comptent parmi les plus importants pour la catégorisation faciale du sexe. Quoiqu'il en soit, il est possible que d'autres attributs comme la mâchoire et les cheveux comportaient des informations sexuelles de type relationnel et que le repositionnement des yeux, des sourcils, du nez et de la bouche sans égard à la position de ces autres attributs ait réduit le réalisme des visages ou même perturbé le signal de ces autres indices. Cet argument, s'il est poussé à l'extrême, suggère que tout repositionnement d'une région sans égard aux autres régions du visage peut engendrer des déformations d'ordre spatial. Bien que possible, il est toutefois peu probable que de telles déformations aient affectées davantage les visages de l'une des deux catégories sexuelles. En fait, il est plus probable que de telles déformations aient jouées un rôle à la fois positif et négatif dans l'utilisation des indices relationnels manipulés. Quoiqu'il en soit, cette conception très stricte des indices relationnels de deuxième

ordre est très peu courante dans la littérature sur le sujet et renvoie à l'impossibilité de manipuler de tels indices. De plus, nous avons vu à la section 5.4.2 que les études basées sur des définitions plus larges des indices relationnels avaient elles aussi conclu de la faible importance perceptive de ces indices.

5.6.3. Pistes de recherches futures

Les résultats présentés dans le cadre de cette thèse ouvrent de nouvelles perspectives de recherche. Nous en aborderons quelques-unes.

5.6.3.1. Sexe de l'observateur

Nous avons traité de la catégorisation du sexe des visages sans tenir compte du sexe de l'observateur. Pourtant, les résultats de toutes nos expériences indiquent des différences inter-gendre. D'abord, nous avons noté une tendance systématique des hommes à répondre 'femme' plus souvent que 'homme' et vice-versa. De plus, ce biais de réponse était toujours plus prononcé chez les hommes que chez les femmes. Sur le plan évolutif, la détection du sexe opposé pourrait sembler souhaitable dans un contexte où il n'y pas de danger imminent — comme c'est le cas dans nos expériences. De pair avec cette observation, les nouvelles analyses sur les données de l'étude 3 présentées à la section 5.3.1 révèlent également des différences entre les volumes de classification des observateurs hommes et femmes. En effet, seul le volume chromatique de classification moyen des observateurs hommes présente des voxels significatifs dans la région orale, ce qui suggère que cette région est plus importante pour ceux-ci. Afin d'accroître le réalisme de nos stimuli, certains des visages de femmes dans notre banque de visages arboraient du rouge à lèvres. Comme cet indice a pu interagir avec les résultats, il serait intéressant de valider ces résultats avec des visages sans maquillage et photographiés dans un environnement contrôlé. Dans un tel contexte où nous disposerions de toutes les informations nécessaires à la mesure de la réflectance des visages *dans le vrai monde*, nous pourrions tester l'hypothèse de Stephen et McKeegan (2010) selon laquelle la couleur des lèvres des femmes pourrait signaler certaines variations hormonales ou

physiologiques liées à l'excitation sexuelle et la santé cardio-respiratoire. Par exemple, nous pourrions d'abord estimer la quantité d'information chromatique véhiculée dans la région de la bouche par les variations hormonales liées aux différentes phases du cycle menstruel. Ensuite, nous pourrions tester la capacité des observateurs hommes et femmes à détecter ces variations chromatiques. Si les observateurs hommes déploient effectivement plus de ressources cognitives au traitement de ces informations, il est possible qu'ils soient aussi plus sensibles à ces indices.

5.6.3.2. Validation de l'hypothèse sur le multiplexage fréquentiel des régions faciales

L'article 3 présente des résultats inédits, à savoir le multiplexage fréquentiel des régions faciales dans une tâche de discrimination des visages. Bien qu'il soit très peu probable qu'un tel patron de résultats survienne par hasard, les analyses ayant menées à ces résultats présentent certaines limites. D'abord, l'analyse linéaire des images de classification sur laquelle repose l'analyse temps-fréquences suppose l'indépendance de chaque voxel échantillonné par les bulles. Donc, compte tenu de la nature aléatoire de la méthode Bubbles et de la faible proportion de voxels échantillonnés, la probabilité qu'une région du visage ait été échantillonnée à l'une des fréquences f observées comme significative, et ce, pendant un cycle complet, est certainement très faible. Suivant cette logique, il est encore moins probable que les participants aient été soumis aux patrons fréquentiels caractéristiques du multiplexage identifié dans les résultats.

Pour toutes ces raisons, il semble évident que les résultats de l'article 3 demandent d'être validés par une expérience supplémentaire et une méthodologie complémentaire. Une expérience de validation possible serait de tester plus directement l'hypothèse d'un multiplexage fréquentiel en réduisant considérablement l'espace de recherche. Pour ce faire, nous pourrions nous limiter à manipuler la fréquence et la phase des oscillations des informations chromatiques et achromatiques des trois régions faciales d'intérêt, soit la bouche et les deux yeux. Nous ajouterions du bruit additif aux stimuli pour maintenir l'exactitude moyenne des réponses à la tâche de discrimination du genre à 75%. Cette expérience nous

permettrait d'évaluer pour chaque sujet la paire fréquence-phase optimale pour chacune des régions d'intérêt.

5.6.3.3. Test d'une hypothèse sur l'origine de la faible efficacité des DIA.

En s'appuyant sur les travaux portant sur l'intégration de multiples informations dans une ou plusieurs modalités sensorielles (e.g., Hillis et al., 2004; Young et al., 1993; Alais et Burr, 2004; Battaglia et al., 2003; Ernst et Banks, 2002; Gold, Mundy & Tjan, 2012), nous avons formulé l'hypothèse selon laquelle l'observateur humain utiliserait moins les DIA puisqu'elles contiennent moins d'informations sexuelles. En fait, ces études suggèrent que les humains pondèrent de façon quasi-optimale les différents indices en fonction de leur informativité. En plus de suggérer que l'observateur humain est adapté à son environnement sur le plan de l'information visuelle, cette hypothèse prédit également que la pondération des indices pourrait changer selon leur informativité. Dans une telle optique, une légère amplification d'un signal facial pourrait, au fil des essais, accroître son utilisation. Il serait intéressant de tester ces hypothèses dans le contexte des indices faciaux chromatiques et achromatiques, d'autant que certains chercheurs prêtent encore trop peu d'attention au réalisme des signaux faciaux qu'ils testent.

Par exemple, nous pourrions comparer les efficacités associées aux DIA dans deux groupes de sujets ayant subi des traitements différents au préalable : au premier groupe de sujets, nous aurions présenté des stimuli faciaux avec des DIA ne comportant aucune information pour la discrimination du genre alors qu'au second groupe de sujets nous aurions présenté des stimuli faciaux avec des DIA hautement informatives pour la discrimination du sexe. Notre hypothèse prédit que les participants du second groupe auront, en moyenne, une efficacité à utiliser les DIA supérieure à celles des participants du premier groupe.

5.7. Conclusion

Dans cette thèse, nous avons étudié les informations signalant le sexe des visages ainsi que la part de ces informations utilisée par les observateurs humains. Si sur le plan expérimental nous nous étions donnés l'objectif d'étudier ces informations dans un contexte réaliste, nous avons bien vu à la section 5.6.2 que la généralisation de nos résultats à des conditions normales de vision était limitée. En effet, pour des raisons pratiques mais aussi afin de tracer des liens avec la littérature scientifique actuelle sur la perception des visages, nous nous sommes limités à l'étude d'un sous-ensemble des informations signalant le sexe des visages et n'avons testé que des visages d'ascendance caucasienne présentés en plan frontal. En plus de ces limites, toutes nos expériences ont présupposé 1) l'existence de dimorphismes sexuels faciaux stables et 2) l'existence de représentations mentales stables utilisées afin de catégoriser le sexe des visages. Le corollaire de ces présuppositions était que les informations *puissantes* pour la catégorisation du sexe des visages—l'intersection entre les informations disponibles et les informations représentées—étaient également stables. Dans un tel contexte, nos expériences visaient donc à découvrir les informations *puissantes* stables.

Bien que communément admise par une majorité de chercheurs en perception des visages, la stabilité des informations disponibles, représentées et *puissantes* est contestable. D'abord, les informations disponibles pour la catégorisation du sexe varient selon l'utilisation des cosmétiques faciaux (Russell, 2009) qui évolue au gré des saisons et des cultures, et ce, depuis l'Antiquité (Chaudhri et Jain, 2009). Également, le recours à la chirurgie cosmétique faciale est de plus en plus fréquent. Selon la International Society of Aesthetic Plastic Surgery (ISAPS), plus de 23 millions de chirurgies cosmétiques auraient été pratiquées dans le monde en 2013 alors que seulement 8 millions avaient été pratiquées en 2011. Ces interventions médicales changent avec les modes et les cultures et altèrent les informations faciales disponibles. Ainsi, il semble clair que les informations signalant le sexe des visages ne soient pas aussi stables que nous le présupposions.

D'autre part, plusieurs études indiquent que l'information puissante varie dans le temps. Par exemple, nous avons observé dans chacune de nos expériences une augmentation de la performance de nos participants au fur et à mesure des essais, indiquant la présence d'un

apprentissage perceptif/procédural. Si de tels apprentissages n'impliquent pas nécessairement la modification des représentations internes—la réduction du bruit interne, par exemple, aurait pour conséquence d'augmenter la performance sans changer les représentations sous-jacentes (voir Doshier & Lu, 1999)—plusieurs études indiquent que les représentations sont bel et bien dynamiques et changeantes (e.g. Dupuis-Roy & Gosselin, 2007; Gold et al, 1999). À titre d'exemple, nous avons récemment réussi à induire l'utilisation de l'œil gauche chez un groupe de participants et l'utilisation de l'œil droit chez un autre groupe de participants pendant une tâche de catégorisation du sexe au moyen de l'apprentissage opérant (Gosselin, Coüet-Garand et Dupuis-Roy, 2014). Ainsi, il semble clair, contrairement à nos présuppositions, que les représentations mentales et les informations puissantes liées à la catégorisation du sexe des visages ne soient pas nécessairement stables.

Bien que nos expériences s'appuyaient sur de fausses présuppositions, la validité et la portée de nos résultats demeurent. D'abord, les informations sexuelles faciales disponibles étaient maintenues constantes à travers nos expériences. Ensuite, d'après nos observations préliminaires, les changements ayant eu cours dans nos expériences indiquaient plutôt une augmentation de la part des informations disponibles contenue dans les représentations internes. Dans la mesure où les informations disponibles étaient représentatives des dimorphismes sexuels normaux et non-accessoires, on peut supposer que nos expériences ont effectivement permis d'étudier les informations représentées et puissantes invariantes impliquées dans la catégorisation du sexe des visages. Quoiqu'il en soit, la possibilité d'une variation des informations représentées et puissantes peut s'avérer problématique pour les études faisant fi de la représentativité des informations disponibles. Par exemple, la surreprésentation d'un type d'information—comme celles contenues dans les distances entre les attributs faciaux (voir section 5.6.3.3)—peut expliquer en elle-même les hautes performances liées à l'utilisation de ces informations si l'on admet l'idée selon laquelle l'observateur humain tend à pondérer de façon quasi-optimale les différents indices en fonction de leur informativité.

En somme, la variation des informations disponibles, représentées et puissantes pose plusieurs défis majeurs. Le premier consiste, pour le chercheur, à mieux cerner les limites liées à la représentativité des informations disponibles à travers le temps, et les limites quant à l'impact des transformations des informations représentées et puissantes durant l'expérience.

Un second défi concerne les aspects théoriques. Si intuitivement les informations représentées devraient tendre asymptotiquement vers les informations disponibles (e.g. Dupuis-Roy et Gosselin, 2007), aucune loi ni modèle ne prédit précisément l'évolution des informations représentées. Ainsi plusieurs questions demeurent béantes: Les représentations intègrent-elles sélectivement de plus en plus d'informations disponibles, allant des régions faciales les plus informatives aux moins informatives; ou évoluent-elles simplement en donnant de plus en plus d'importance aux zones faciales avoisinant les zones déjà représentées? Comment cette évolution tient-elle compte des limites perceptives et des coûts associés au traitement de tels ou tels indices faciaux? Les différents types d'indices faciaux (e.g. les indices chromatiques, achromatiques, relationnels) suivent-ils des régimes d'apprentissage différents?

Un troisième et dernier défi concerne la difficulté à cerner la *dynamique* de ces variations. Si l'apprentissage perceptif tel que nous l'avons rapporté exprime clairement un changement des informations représentées en réponse au bouleversement des informations disponibles, certains phénomènes comme le maquillage, l'épilation des sourcils ou la chirurgie cosmétique suggèrent plutôt la dynamique inverse. "A man's face is his autobiography. A woman's face is her work of fiction.", comme le disait Oscar Wilde. En effet, qu'est-ce qui pourrait bien pousser un individu à se faire raboter le nez ou à se faire injecter du collagène dans les lèvres? Une possibilité fascinante est que ce serait précisément les représentations du sexe des visages de l'individu qui serviraient de guide à ces altérations, ce qui n'est pas sans rappeler l'idée des *stimuli supernormaux* proposés par Tinbergen (1951).

Références

- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P. G. & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*, 68-72.
- Alais, D., & Burr, D. (2004) The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257-262.
- Ahumada A. J. Jr., Lovell J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, *49*, 1751–1756.
- Barragan-Jason G, Lachat F and Barbeau EJ (2012) How fast is famous face recognition? *Frontiers in Psychology*, *3*, 454. doi: 10.3389/fpsyg.2012.00454
- Barton, J. J. S., Keenan, J. P., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. *British Journal of Psychology*, *92*, 527-549.
- Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, *20*, 1391-1397.
- Baudouin, J. Y., & Gallay, M. (2006). Is face distinctiveness gender based?. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 789.
- Bhatt, R. S., Bertin, E., Hayden, A., & Reed, A. (2005). Face processing in infancy: Developmental changes in the use of different kinds of relational information. *Child Development*, *76*(1), 169-181.

- Blais, C., Arguin, M., & Gosselin, F. (2013). Human visual processing oscillates: Evidence from a classification image technique. *Cognition, 128*(3), 353–362. doi:10.1016/j.cognition.2013.04.009
- Blais, C., Fiset, D., Arguin, M., Jolicoeur, P., Bub, D., & Gosselin, F. (2009). Reading between eye saccades. *PloS one, 4*(7), e6448.
- Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia, 50*(12), 2830–2838. doi:10.1016/j.neuropsychologia.2012.08.010
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.
- Brown, E., & Perrett, D. I. (1993). What gives a face its gender. *Perception, 22*, 829–840.
- Bruce, V., Burton, A. M., Hanna, E., Healey, P., Mason, O., Coombes, A., et al. (1993). Sex-discrimination: How well do we tell the difference between male and female faces. *Perception, 22*, 131–152.
- Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied, 5*(4), 339.
- Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognizing the sex and identities of faces. *Perception, 23*, 803–822.
- Bruce, V. & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327.
- Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces:

Computer graphic manipulation of shape and colour information. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 259(1355), 137-143.

Burton, A. M., Bruce, V., & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception*, 22, 153–176.

Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences of the United States of America*, 107(37), 16048–16053. doi:10.1073/pnas.1004801107/-/DCSupplemental

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature neuroscience*, 3(11), 1077-1078.

Campbell, R., Benson, P. J., Wallace, S. B., Doesbergh, S., & Coleman, M. (1999). More about brows: How poses that change brow position affect perceptions of gender. *Perception*, 28, 489-50.

Campbell, R., Wallace, S., & Benson, P. (1996). Real men don't look down: Direction of gaze affects sex decisions on faces. *Visual Cognition*, 3(4), 393–412.

Cate, ten, C., Bruins, W. S., Ouden, den, J., Egberts, T., Neevel, H., Spierings, M., et al. (2009). Animal Behaviour. *Animal Behaviour*, 77(4), 795–802. doi:10.1016/j.anbehav.2008.12.020

Chaudhri, S. K., & Jain, N. K. (2009). History of cosmetics. *Asian Journal of Pharmaceutics*, 3(3), 164.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, 5(9):1, 659–667,

<http://journalofvision.org/5/9/1/>, doi:10.1167/5.9.1.

- Clark, V. P., Parasuraman, R., Keil, K., Kulansky, R., Fannon, S., Maisog, J. M., et al. (1997). Selective attention to face identity and color studied with fMRI. *Human Brain Mapping*, 5, 293–297.
- Corballis, M. C., & Sergent, J. (1989). Hemispheric specialization for mental rotation. *Cortex*, 25(1), 15-25.
- Cornell, E. H. (1974). Infants' discrimination of photographs of faces following redundant presentations. *18(1)*, 98-106.
- Corson, R. (1972). *Fashions in makeup*. London: Peter Owen.
- DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *Journal of Neurophysiology*, 69, 1118–1118.
- Dehaene, S. (1993). Temporal oscillations in human perception. *Psychological Science*, 4(4), 264–270.
- Dosher, B. A., & Lu, Z. L. (1999). Mechanisms of perceptual learning. *Vision Research*, 39, 3197–3221.
- Dostoevski, F. (1957). *L'Idiot*. Gallimard, Paris, 1364 pages.
- Dobbelsteyn, C. J., Joffres, M. R., MacLean, D. R., & Flowerdew, G. (2001). A comparative evaluation of waist circumference, waist-to-hip ratio and body mass index as indicators of cardiovascular risk factors. The Canadian Heart Health Surveys. *International journal of obesity and related metabolic disorders: journal of the International Association for*

the Study of Obesity, 25(5), 652-661.

Dosher, B., and Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision Research*, 39, 3197-3221.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic faces. *Neuropsychologica*, 44, 576–585.

Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision*, 9(2), 1-8.

Dupuis-Roy, N., Fiset, D., Dufresne, K., Caplette, L., & Gosselin, F. (2014). Real-world interattribute distances lead to inefficient face gender categorization. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1289–1294. doi:10.1037/a0037066

Eckstein, M. P., & Ahumada, A. J., Jr. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision*, 2(1):i, i, <http://journalofvision.org/2/1/i/>, doi:10.1167/2.1.i.

Edwards, R., Xiao, D., Keysers, C., Foldiak, P., & Perrett, D. (2003). Color sensitivity of cells responsive to complex stimuli in the temporal cortex. *Journal of Neurophysiology*, 90, 1245–1256.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 1, 54–75.

Ellis, H. D. (1986). Processes underlying face recognition. *The neuropsychology of face perception and facial expression*, 1-27.

- Enlow, D. H. (1990). *Facial growth*, 3rd edn. Philadelphia : W.B. Saunders.
- Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences*, *5(1)*, 16–25.
- Ernst, M. O. & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429-433.
- Evano, N., Clapier, A., & Coulon, P.-Y. (2004). Accurate and quasi-automatic lip tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, *14*, 706–715.
- Farkas, L. G. (1981). *Anthropometry of the head and face in medicine*.
- Farkas, L. G., & Munro, I. R. (1987). *Anthropometric facial proportions in medicine*. Charles C. Thomas Publisher.
- Fink, B., & Neave, N. (2005). The biology of facial beauty. *International journal of cosmetic science*, *27(6)*, 317-325.
- Fink, B., Grammer, K., & Matts, P. J. (2006). Visible skin color distribution plays a role in the perception of age, attractiveness, and health in female faces. *Evolution and Human Behavior*, *27(6)*, 433-442.
- Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science*, *19(11)*, 1161-1168.
- Fiset, D., Blais, C., Arguin, M., Tadros, K., Éthier-Majcher, C., Bub, D., & Gosselin, F. (2009). The spatio-temporal dynamics of visual letter recognition. *Cognitive Neuropsychology*, *26(1)*, 23–35. doi:10.1080/02643290802421160

- Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29*(2), 159–170.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences, 9*(10), 474–480. doi:10.1016/j.tics.2005.08.011
- Frost, P. (1988). Human skin color: A possible relationship between its sexual dimorphism and its social perception. *Perspectives in Biology and Medicine*.
- Galton, F. (1879). Composite portraits, made by combining those of many different persons into a single, resultant figure. *Journal of the Anthropological Institute, 8*, 132-144.
- Galton, F. (1888). Personal identification and description. *Nature, 38*, 173–202.
- Gaspar, C. M., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research, 48*, 2817–2826.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature neuroscience, 3*(2), 191-197.
- Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 431.
- Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology, 10*(13), 805–808.

- Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & 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:10.1068/p5370
- Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature, 402*, 176-178.
- Gold, J. M., Mundy, P. J., & Tjan, B. S. (2012). The perception of a face is no more than the sum of its parts. *Psychological science, 23*(4), 427-434.
- Gold J. M., Murray R. F., Bennett P. J., Sekuler A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology, 10*, 663–666.
- Gold, J.M., Barker, J.D., Barr, S., Bittner, J.L., Bromfield, W.D., Chu, N., Goode, R.A., Lee, D., Simmons, M. & Srinath, A. (2013). The efficiency of dynamic and static facial expression recognition. *Journal of Vision, 13*(5): 23; doi:10.1167/13.5.23
- Gosselin, F., Couet-Garand, A., & Dupuis-Roy, N. (2014). Greater usage of the left eye causes better facial gender discrimination. *Journal of Vision, 14*(10), 1440–1440.
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*, 2261–2271.
- Gosselin, F., & Schyns, P. G. (2002). RAP: A new framework for visual categorization. *Trends in Cognitive Science, 6*, 70–77.
- Gosselin, F. & Schyns, P. G. (2003). Superstitious perceptions reveal properties of memory representations. *Psychological Science, 14*, 505-509.
- Gosselin, F., & Schyns, P. G. (Eds.) (2004). Rendering the use of visual information from

spiking neurons to recognition. *Cognitive Science*, 28, 141–301.

Gosselin, F., Spezio, M.L., Tranel, D. & Adolphs, R. (2011). Asymmetrical use of eye information from faces following unilateral amygdala damage. *Social Cognitive and Affective Neuroscience*, 6, 330-337.

Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception*, 13, 104-109.

Hammal, Z., Gosselin, F. & Fortin, I. (2009). How efficient are the recognition of dynamic and static facial expressions? *Journal of Vision*, 9(8): 499; doi:10.1167/9.8.499

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., & Joseph, J. E. (2007). The development of expert face processing: are infants sensitive to normal differences in second-order relational information? *Journal of Experimental Child Psychology*, 97(2), 85-98.

Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: The role of shape and colour. *Proceedings of the Royal Society B: Biological Sciences*, 261, 367–373.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, 4, 967–992.

Hochberg J., & Galper R. E. (1967). Recognition of faces: I. An exploratory study. *Psychonomic Science*, 9, 619–620.

Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nature Neuroscience*, 4, 127–128.

Hosie, J. A., Ellis, H. D., Haig, N. D. (1988). The effect of feature displacement on the

perception of well-known faces, *Perception*, 17, 461- 474.

Ishihara, S. (1917). *Tests for color-blindness*. Handaya : Tokyo, Hongo Harukicho.

Jablonski, D. 2000. Micro- and macroevolution scale and hierarchy in evolutionary biology and paleobiology. In D. H. Erwin and S. L. Wing (Eds.) *Deep time: Paleobiology's perspective*. *Paleobiology*, 26, (Suppl. to No. 4), 15–52.

Kahana, M. J., Seelig, D., & Madsen, J. R. (2001). Theta returns. *Current opinion in neurobiology*, 11(6), 739-744.

Kahana, M. J., Sekuler, R., Caplan, J. B., Kirschen, M., & Madsen, J. R. (1999). Human theta oscillations exhibit task dependence during virtual maze navigation. *Nature*, 399(6738), 781–784.

Koepsell, K., Wang, X., Hirsch, J. A., & Sommer, F. T. (2010). Exploring the function of neural oscillations in early sensory systems. *Frontiers in neuroscience*, 4.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2010). Holistic processing is not correlated with face-identification accuracy. *Psychological Science*, 21(1), 38-43.

Landau, A. N., & Fries, P. (2012). Attention Samples Stimuli Rhythmically. *Current Biology*, 22(11), 1000–1004. doi:10.1016/j.cub.2012.03.054

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *The Quarterly Journal of Experimental Psychology: Section A*, 51(3), 449-473.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology: Section A*, 53(2), 513–536.

- Leder, H., Candrian, G., & Huber, O. (2001). Configural features in the context of upright and inverted faces. *Perception, 30*, 73-83.
- Leder, H., & Carbon, C. C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology, 97*(1), 19-29.
- Lee, K. J., & Perrett, D. (1996). Presentation-time measures of the effects of manipulations in colour space on discrimination of famous faces. *Perception, 26*(6), 733-752.
- Lee, K. J., & Perrett, D. I. (2000). Manipulation of colour and shape information and its consequence upon recognition and best-likeness judgments. *Perception, 29*(11), 1291–1312. doi:10.1068/p2792
- Le Grand, R., Mondloch, C. J., Maurer, D., Brent, H. P. (2001). Early visual experience and face processing. *Nature, 410*, 890.
- Leinbach, M. D. & Fagot, B. I. (1993). Categorical habituation to male and female faces: Gender schematic processing in infancy. *16*(3), 317-332.
- Lisman, J. E., & Idiart, M. A. (1995). Storage of 7+/-2 short-term memories in oscillatory subcycles. *Science, 267*(5203), 1512–1515.
- Macdonald, J., Cavanagh, P., & VanRullen, R. (2014). Attentional sampling of multiple wagon wheels. *Attention, Perception & Psychophysics, 76*, 64-72.
- Mangini M. C. & Biederman I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science, 28*, 209–226.
- Matts, P. J., Fink, B., Grammer, K., & Burquest, M. (2007). Color homogeneity and visual

- perception of age, health, and attractiveness of female facial skin. *Journal of the American Academy of Dermatology*, 57(6), 977-984.
- Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255-260.
- 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. J., Pathman, T., Maurer, D., Le Grand, R., & de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence of adult-like holistic face processing. *Visual Cognition*, 15(5), 564-577.
- Mouchetant-Rostaing, Y. & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *Journal of Cognitive Neuroscience*, 15 (6), 900-910.
- Mouchetant-Rostaing, Y., Giard, M. H., Delpuech, C., Echallier, J. F., & Pernier, J. (2000). Early signs of visual categorization for biological and non-biological stimuli in humans. *NeuroReport*, 11, 2521–2525.
- Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society of London B: Biological Sciences*, 264, 393–399.
- Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London B: Biological Sciences*, 264, 1407–1414.

- Mullen, K. T. (1985). The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *The Journal of Physiology*, 359(1), 381–400.
- Murray, R. F., & Gold, J. M. (2004). Troubles with bubbles. *Vision Research*, 44, 461–470.
- Neri, P., & Levi, D. (2008). Temporal dynamics of directional selectivity in human vision. *Journal of Vision*, 8(1), 22–22. doi:10.1167/8.1.22
- Nestor, A., & Tarr, M. J. (2008a). The segmental structure of faces and its use in gender recognition. *Journal of Vision*, 8(7):7, 1–12, <http://journalofvision.org/8/7/7/>, doi:10.1167/8.7.7.
- Nestor, A., & Tarr, M. J. (2008b). Gender recognition of human faces using color. *Psychological Science*, 19, 1242–1246.
- Ogden C, Fryar C, Carrol M, Flegal K, 2004 Mean Body Weight, Height, and Body Mass Index, United States 1960-2002 Advance data from vital and health statistics, no 347 (Hyattsville, MD: National Center for Health Statistics)
- O'Toole, A. J., Peterson, J. & Deffenbacher, K. A. (1996). An 'other-race effect' for categorizing faces by sex. *Perception* 25(6), 669-676.
- Palmer, S. E. (1999). *Vision science: Photons to phenomenology* (Vol. 1). Cambridge, MA: MIT press.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proceedings of the National Academy of Sciences of the United States of*

America, 93(8), 3693–3697.

- Reddy, L., Wilken, P., & Koch, C. (2004). Face-gender discrimination is possible in the near-absence of attention. *Journal of Vision*, 4(2), 1–12. doi:10.1167/4.2.4
- Rhodes, G. (1988). Looking at faces – 1st order and 2nd order features as determinants of facial appearance. *Perception* 17, 43–63.
- Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, 47(1), 25-57.
- Rhodes, G., Hayward, W. G., and Winkler, C. (2007). Expert face coding: configural and component coding of own-race and other-race faces. *Psychon. Bull. Rev.* 13, 499–505.
- Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science*.
- Ringach, D., & Shapley, R. (2004). Reverse correlation in neurophysiology. *Cognitive Science*, 28, 247-166.
- Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, 103(1), 34-79.
- Roberts, T., & Bruce, V. (1988). Feature saliency in judging the sex and familiarity of faces. *Perception*, 17(4), 475-481.
- Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature*, 397(6718), 430–433.

- Romei, V., Driver, J., Schyns, P. G., & Thut, G. (2011). Rhythmic TMS over parietal cortex links distinct brain frequencies to global versus local visual processing. *Current Biology*, *21*(4), 334–337. doi:10.1016/j.cub.2011.01.035
- Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, *128*(2), 274-289.
- Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta psychologica*, *132*(3), 300-312.
- Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychological Research*, *70*, 273–292.
- Russell, R. (2003). Sex, beauty, and the relative luminance of facial features. *Perception*, *32*, 1093–1107.
- Russell, R. (2005). Face pigmentation and sex classification [Abstract]. *Journal of Vision*, *5*(8):983, 983a, <http://journalofvision.org/5/8/983/>, doi:10.1167/5.8.983.
- Russell, R. (2009). A sex difference in facial contrast and its exaggeration by cosmetics. *Perception*, *38*(8).
- Sæther, L., Van Belle, W., Laeng, B., Brennen, T., & Øvervoll, M. (2009). Anchoring gaze when categorizing faces' sex: Evidence from eye-tracking data. *Vision Research*, *49*(23), 2870–2880. doi:10.1016/j.visres.2009.09.001
- Sartre, J-P. (1938). *La Nausée*. Paris, Gallimard, 233 pages.
- Schaffer, S. E. (2007). Reading our lips: The history of lipstick regulation in Western seats of power. *Food & Drug LJ*, *62*, 165.

- Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science, 13*, 402–409.
- Schyns, P. G., Gosselin, F., & Smith, M. L. (2009). Information processing algorithms in the brain. *Trends in Cognitive Sciences, 13(1)*, 20–26. doi:10.1016/j.tics.2008.09.008
- Schyns, P. G., Thut, G., & Gross, J. (2011). Cracking the code of oscillatory activity. *PLoS Biology, 9(5)*, e1001064. doi:10.1371/journal.pbio.1001064.s003
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14*, 391–396.
- Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75*, 221-242.
- Sergent, J., Ohta, S., & McDonald, B. (1992). Functional neuroanatomy of face and object processing. *Brain, 115*, 15–36.
- Sheperd, J. (1981). Social factors in face recognition. In G. Davies, H. Ellis & J. Shepherd (Ed.s), *Perceiving and remembering faces* (pp. 55-79). San Diego, CA : Academic Press.
- Smith, M. L., Gosselin, F., & Schyns, P. G. (2006). Perceptual moments of conscious visual experience inferred from oscillatory brain activity. *Proceedings of the National Academy of Sciences of the United States of America, 103(14)*, 5626–5631.
- Smith, M. L., Gosselin, F., & Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *NeuroImage, 37(3)*, 974–984. doi:10.1016/j.neuroimage.2007.05.030

- Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, *369*(6479), 395-397.
- Staudigl, T., & Hanslmayr, S. (2013). Theta Oscillations at Encoding Mediate the Context-Dependent Nature of Human Episodic Memory. *Current Biology*, *23*(12), 1101–1106. doi:10.1016/j.cub.2013.04.074
- Stephen, I. D., Coetzee, V., Law Smith, M., & Perrett, D. I. (2009a). Skin Blood Perfusion and Oxygenation Colour Affect Perceived Human Health. *PloS One*, *4*(4), e5083. doi:10.1371/journal.pone.0005083.s006
- Stephen, I. D., Law Smith, M. J., Stirrat, M. R., & Perrett, D. I. (2009b). Facial Skin Coloration Affects Perceived Health of Human Faces. *International Journal of Primatology*, *30*(6), 845–857. doi:10.1007/s10764-009-9380-z
- Stephen, I. D., & McKeegan, A. M. (2010). Lip colour affects perceived sex typicality and attractiveness of human faces. *Perception*, *39*(8), 1104–1110. doi:10.1068/p6730
- Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, *3*(4), 151–162.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46*(2), 225-245.
- Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*(5), 583-604.
- Tanaka, J. W., Weiskopf, D., & Williams, P. (2001). The role of color in high-level vision. *Trends in Cognitive Sciences*, *5*, 211–215.

- Tanner, W. P., & Birdsall, T. G. (1958). Definitions of d' and η as psychophysical measures. *The Journal of the Acoustical Society of America*, 30(10), 922-928.
- Tarr, M. J., Kersten, D., Cheng, Y., & Rossion, B. (2001). It's Pat! Sexing faces using only red and green [Abstract]. *Journal of Vision*, 1(3):337, 337a, <http://journalofvision.org/1/3/337/>, doi:10.1167/1.3.337.
- Tarr, M. J., Rossion, B., & Doerschner, K. (2002). Men are from Mars, women are from Venus: Behavioral and neural correlates of face sexing using color [Abstract]. *Journal of Vision*, 2(7):598, 598a, <http://journalofvision.org/2/7/598/>, doi:10.1167/2.7.598.
- Taschereau-Dumouchel, V., Rossion, B., Schyns, P. G., & Gosselin, F. (2010). Interattribute Distances do not Represent the Identity of Real World Faces. *Frontiers in psychology*, 1:159, 1-10. doi:10.3389/fpsyg.2010.00159
- Tesche, C., & Karhu, J. (1999). Interactive processing of sensory input and motor output in the human hippocampus. *Cognitive Neuroscience, Journal of*, 11(4), 424-436.
- Tesche, C. D., & Karhu, J. (2000). Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2), 919-924.
- Thut, G., Veniero, D., Romei, V., Miniussi, C., Schyns, P., & Gross, J. (2011). Rhythmic TMS causes local entrainment of natural oscillatory signatures. *Current Biology*, 21(14), 1176-1185. doi:10.1016/j.cub.2011.05.049
- Tinbergen, N. (1951). *The Study of Instinct*. Oxford, Clarendon Press.
- Tjan, B.S., Braje, W.L., Legge, G.E. & Kersten, D. (1995). Human efficiency for recognizing

- 3-D objects in luminance noise, *Vision Research*, 35 (21), 3053-3069.
- Uchida, N., Kepecs, A., & Mainen, Z. F. (2006). Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nature Reviews Neuroscience*, 7(6), 485–491.
- Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British journal of psychology*, 79(4), 471-491.
- VanRullen, R., Reddy, L. & Koch, C. (2005). Attention-driven discrete sampling of motion perception. *Proceeding of the National Academy of Science USA*, 102(14), 5291-5296.
- VanRullen, R. (2006). The continuous Wagon Wheel Illusion is object-based. *Vision Research*, 46(24), 4091-4095.
- VanRullen, R., & Dubois, J. (2011). The psychophysics of brain rhythms. *Frontiers in Psychology*, 2. doi:10.3389/fpsyg.2011.00203/abstract
- VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, 7(5), 207–213. doi:10.1016/S1364-6613(03)00095-0
- VanRullen, R., Carlson, T., & Cavanagh, P. (2007). The Blinking Spotlight of Attention. *Proceedings of the National Academy of Sciences of the United States of America*, 104(49), 19204–19209.
- Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2(4), 229–239.
- Vinette, C., & Gosselin, F. (2004). Spatio-Temporal Dynamics of Face Recognition in a Flash: It's in the eyes. *Cognitive Science*, 28, 289-301, doi:10.1016/j.cogsci.2004.01.002

- Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113–120.
- Wild, H. A., Barrett, S. E., Spence, M. J., O'Toole, A. J., Cheng, Y. D., & Brooke, J. (2000). Recognition and Sex Categorization of Adults' and Children's Faces: Examining Performance in the Absence of Sex-Stereotyped Cues. *Journal of Experimental Child Psychology*, *77*(4), 269–291. doi:10.1006/jecp.1999.2554
- Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., ... & Gosselin, F. (2010). Does face inversion change spatial frequency tuning?. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(1), 122.
- Willenbockel, V., Lepore, F., Bacon, B. A., & Gosselin, F. (2013). The informational correlates of conscious and nonconscious face-gender perception. *Journal of vision*, *13*(2), 10.
- Yamaguchi, M. K., Hirukawa, T., & Kanazawa, S. (1995). Judgment of gender through facial parts. *Perception*, *24*, 563–575.
- Yang, N., Shafai, F., & Oruc, I. (2014). Size determines whether specialized expert processes are engaged for recognition of faces. *Journal of vision*, *14*(8), 17.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141-145.
- Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, *31*, 995–1003.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face

perception. *Perception*, 16(6), 747-759.

Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, 33, 2685–2696.