

Université de Montréal

Les distances entre les attributs internes du visage humain

par
Vincent Taschereau-Dumouchel

Département de psychologie
Faculté des arts et des sciences

Mémoire présenté à la faculté des études supérieures
en vue de l'obtention du grade de maître (M.Sc.)
en psychologie

Août, 2009

© Vincent Taschereau-Dumouchel, 2009

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

Les distances entre les attributs internes du visage humain

Présenté par :

Vincent Taschereau-Dumouchel

A été évalué par un jury composé des personnes suivantes :

Martin Arguin
Président-rapporteur

Frédéric Gosselin
Directeur de recherche

Sven Joubert
Membre du jury

Résumé

La zeitgesit contemporaine sur la reconnaissance des visages suggère que le processus de reconnaissance reposerait essentiellement sur le traitement des distances entre les attributs internes du visage. Il est toutefois surprenant de noter que cette hypothèse n'a jamais été évaluée directement dans la littérature. Pour ce faire, 515 photographies de visages ont été annotées afin d'évaluer l'information véhiculée par de telles distances. Les résultats obtenus suggèrent que les études précédentes ayant utilisé des modifications de ces distances ont présenté 4 fois plus d'informations que les distances inter-attributs du monde réel. De plus, il semblerait que les observateurs humains utilisent difficilement les distances inter-attributs issues de visages réels pour reconnaître leurs semblables à plusieurs distances de visionnement (pourcentage correct maximal de 65%). Qui plus est, la performance des observateurs est presque parfaitement restaurée lorsque l'information des distances inter-attributs n'est pas utilisable mais que les observateurs peuvent utiliser les autres sources d'information de visages réels. Nous concluons que des indices faciaux autre que les distances inter-attributs tel que la forme des attributs et les propriétés de la peau véhiculent l'information utilisée par le système visuel pour opérer la reconnaissance des visages.

Mots clés : Reconnaissance des visages, distances inter-attributs, information configurale, relations de second-ordre, processus holistique.

Abstract

According to an influential view, based on studies of development and of the face inversion effect, human face recognition relies mainly on the treatment of the distances among internal facial features. However, there is surprisingly little evidence supporting this claim. Here, we first use a sample of 515 face photographs to estimate the face recognition information available in interattribute distances. We demonstrate that previous studies of interattribute distances generated faces that exaggerated 4 times this information compared to real-world faces. When interattribute distances are sampled from a real-world distribution, we show that human observers recognize faces poorly across a broad range of viewing distances (with a maximum accuracy of 65%). In contrast, recognition performance is restored when observers only use facial cues of real-world faces other than interattribute distances. We conclude that facial cues other than interattribute distances such as attribute shapes and skin properties are the dominant information of face recognition mechanisms.

Keywords : Face recognition, interattribute distances, relational information, second-order relations, holistic processing.

TABLE DES MATIÈRES

RÉSUMÉ	iv
MOTS-CLÉS	iv
ABSTRACT	v
KEYWORDS	v
TABLE DES MATIÈRES	vi
Liste des figures.....	viii
Liste des abréviations.....	x
Remerciements.....	xi
INTRODUCTION GÉNÉRALE	1
1. La localisation anatomique d'un processus spécialisé pour la reconnaissance des visages.....	3
1.1 FFA: Processus qualitativement différent pour le traitement des visages? ..4	
2. Les évidences comportementales d'un processus cognitif spécialisé pour la reconnaissance des visages	5
2.1 Les attributs internes du visage humain	6
2.2 L'hypothèse holistique/configurale.....	7
2.2.1 L'effet d'inversion	8
2.2.2 L'effet composite	10
2.2.3 L'effet partie-tout	12
3. Critique de la conception et de l'utilisation des distances inter-attributs.....	14
3.1 La notion de distance inter-attributs.....	17
4. Problématique	19
5. Contribution des auteurs	21
ARTICLE	23
Interattributs distances do not represent the identity of real-world face.....	24
ABSTRACT	25
INTRODUCTION	26
Experiment 1	
METHODS	29
Participants	29
Stimuli	29
Apparatus	30
Procedure.....	30
RESULTS	32

Experiment 2	
METHODS	36
Participants	36
Stimuli	37
Apparatus	39
Procedure	39
RESULTS	41
GENERAL DISCUSSION	43
REFERENCES	46
APPENDIX : HOW TO CREATE REALISTIC INTERATTRIBUTE DISTANCES?	52
SUPPLEMENTARY MATERIAL	55
DISCUSSION GÉNÉRALE	58
1. La notion de distances inter-attributs et l’hypothèse holistique.....	59
1.1 Discussion de l’effet d’inversion.....	60
1.1.1 La technique de classification	60
1.1.2 La technique Bubbles	62
1.1.3 Les distances verticales vs les distances horizontales	63
1.2 Discussion de l’effet composite	65
1.3 Critique de l’étude psychophysique des processus holistiques.....	65
2. Les limites de l’étude	67
3. Les pistes d’études futures	69
3.1 L’efficacité des distances inter-attributs et des attributs eux-mêmes	69
3.2 L’effet d’inversion et le réalisme des distances inter-attributs.....	71
RÉFÉRENCES	73
APPENDICE	82

LISTE DES FIGURES

- Figure 1. (A) Visages créés par la manipulation des distances entre les attributs internes
(B) Visages créés par le remplacement des attributs internes par ceux d'autres visages tout en préservant les distances inter-attributs (Mondloch et al., 2002).....10
- Figure 2. (A) Visages composites générés par l'utilisation du même haut de visage (au dessus de la ligne blanche) en combinaison avec plusieurs bas de visage différents (sous la ligne blanche). (B) Lorsque ces visages sont inversés l'effet composite disparaît (Rossion, 2008).....11
- Figure 3. Illustration d'une tâche visant à étudier l'effet partie-tout (McKone, Kanwisher & Duchaine, 2006).....13
- Figure 4. Leftmost column: Sample faces annotated in Experiment 1. The 20 blue crosses show, for these faces, the average annotations across participants. These 20 annotations were reduced to 6 attribute positions—green dots—by averaging the xy-coordinates of the annotations belonging to every attribute. Rightmost column: We translated, rotated, and scaled the attribute positions of each face to minimize the mean square of the difference between them and the average attribute positions across faces. The residual differences between aligned attribute positions—green dots—is the interattribute variance in the real-world.....31
- Figure 5. Distribution of post-alignment attribute positions (green clouds) of the 515 annotated faces, with standard-deviation-length eigenvectors (red segments) centered on the distributions, and overlaid to a face to facilitate the interpretation. The blue dots are the attribute positions of stimuli of 14 previous studies that used distance manipulations (Barton, Keenan & Bass, 2001, *British Journal of Psychology*, 92, 527-549; Bhatt, Bertin, Hayden & Reed, 2005, *Child Development*, 76, 169-181; Freire, Lee & Symons, 2000, *Perception*, 29, 159-170; Goffaux, Hault, Michel, Vuong & Rossion, 2005, *Perception*, 34, 77-86; Haig, 1984; Hayden, et al. 2007; Hosie, Ellis & Haig, 1988; Leder & Bruce, 1998, *Quarterly Journal of Experimental Psychology: Section A*, 51, 449-473; Leder & Bruce, 2000, *The Quarterly Journal of Experimental Psychology: Section A*, 53, 513-536; Leder, Candrian, Huber & Bruce, 2001, *Perception*, 30, 73-83; Le Grand, Mondloch, Maurer, & Brent, 2001; Rhodes, Brake & Atkinson, 1993, *Cognition*, 47, 25-57; Sergent, 1984, *British Journal of Psychology*, 75, 221-242; Tanaka, & Sengco, 1997).....34

- Figure 6. Leftmost column: In Experiment 2, feature masks—shown in translucent green—were bestfitted to the aligned annotations—represented by blue crosses. Rightmost column: In Experiment 2a, these feature masks were displaced according to the feature positions of another face of the same gender. Translucent green areas reproduce the feature masks of the leftmost column; translucent red areas represent the same feature masks after displacement; and translucent yellow areas represent the overlap between these two sets of feature masks38
- Figure 7. Sequence of events in two sample trials of our experiments. Top: In Experiment 2a, we asked whether human observers can use this real-world interattribute distance information, at different viewing distances, to resolve a matching-to-sample (ABX) task when interattribute distance is the only information available. Bottom: In Experiment 2b, we asked the complementary question: Can human observers use real-world cues other than interattribute distances such as attribute shapes and skin properties, at different viewing distances, to resolve an ABX.....40
- Figure 8. Mean proportion of correct face recognition in function of distance. Error bars represent one standard error. The dashed line represents performance when real-world interattribute distance is the only information available (Experiment 2a); and the solid line represents performance when only real-world cues other than interattribute distances such as attribute shapes and skin properties are available.....42
- Figure 9. (A) Visages composites utilisés dans les études sur l’effet composite (B) Visages de la section A présentées avec la technique Bubbles (Visages originaux; Rossion, 2008).67

LISTE DES ABBRÉVIATIONS

<i>FFA</i>	<i>Fusiform face area</i>
STS	Sulcus temporal supérieur
<i>OFA</i>	<i>Occipital face area</i>
EEG	Électroencéphalographie
MEG	Magnétoencéphalographie
PCA	<i>Principal component analysis</i>

REMERCIEMENTS

Je tiens spécialement à remercier mon directeur de recherche, Frédéric Gosselin, pour ses précieux conseils et son support constant. Via de nombreuses discussions d'une richesse intellectuelle remarquable, Frédéric a su susciter mon intérêt pour la recherche et ses conseils ont certainement été d'une aide inestimable d'un point de vue académique mais aussi d'un point de vue humain. Ce fut un réel privilège de pouvoir être l'un de ses étudiants et je le remercie énormément. J'aimerais également remercier les étudiants du laboratoire et plus particulièrement Cynthia Roy, Nicolas Dupuis-Roy, Sylvain Roy, Catherine Ethier-Macher, Daniel Fiset et Caroline Blais pour leur aide et leurs oreilles attentives lors des moments difficiles. De plus, j'aimerais remercier mon père Yves et mon frère Mathieu pour avoir cru en moi et pour leurs encouragements constant. Mes amis, Julien Tessier, Sophy Sansfaçon, Karel Bisson, Pierre-Alexis Gourde, Julien Lamirande, Krystel Richer, Guillaume Leblanc, Karine Pelletier, Julien Simard et Julie Giguère ont également été d'un support inestimable tout au long de mon parcours et je les remercie de leur patience et de leurs encouragements.

INTRODUCTION GÉNÉRALE

Le visage humain est l'un des stimuli visuel les plus riches en information socialement pertinente et l'efficacité de son traitement est d'une importance primordiale dans nos interactions quotidiennes. Il suffit d'un bref instant pour que le système visuel puisse extraire de l'information relative au sexe, à l'âge, à l'identité, à la race et à l'état émotionnel de la personne observée. Parmi ces différents aspects, la reconnaissance de l'identité renferme un intérêt certain de par notre remarquable aptitude à reconnaître rapidement et facilement une très grande quantité de visages et ce même dans les contextes les plus variés. Une myriade de démonstrations neurologiques et comportementales suggèrent effectivement qu'un mécanisme spécialisé serait recruté afin de traiter les visages par opposition aux autres objets de l'environnement (Kanwisher, 2000; Moscovitch, Winocur, & Behrmann, 1997; Yin, 1969). La nature de ce procédé est généralement comprise comme reposant principalement sur le traitement des distances entre les attributs internes du visage. L'étude centrale présentée dans ce mémoire vise à réévaluer cette position.

L'introduction de ce mémoire par article portera, dans un premier temps, sur la revue des démonstrations empiriques appuyant les théories principales de la reconnaissance des visages et, dans un second temps, sur le positionnement de la problématique abordée dans l'article du mémoire. La présentation de l'article central du mémoire suivra ainsi qu'une discussion plus élaborée de l'implication des résultats y étant rapportés pour le domaine de la reconnaissance des visages.

1. La localisation anatomique d'un processus spécialisé pour la reconnaissance des visages

Bien que des déficits affectant sélectivement la reconnaissance des visages aient été recensés par des études de cas de Hughlings Jackson et Charcot au dix-neuvième siècle, il fallut attendre le vingtième siècle avant que cette affection soit reconnue comme un trouble neurologique distinct. Le terme prosopagnosie a été utilisé pour la première fois en 1947 par un neurologue allemand, Joachim Bodamer afin de décrire la condition d'un jeune soldat ayant été atteint d'un projectile à la tête durant la guerre. Le terme prosopagnosie trouve son étimologie dans l'union des termes grecs *prosopon* et *agnosia* signifiant respectivement « visage » et « non-connaissance ». L'étude de cas du patient CK est fréquemment citée comme l'une des premières démonstrations cliniques à avoir suggéré que la prosopagnosie soit une affection distincte. Moscovitch et collègues (1997) ont rapporté que des suites d'un traumatisme crânien, le patient CK a présenté une agnosie visuelle complète tout en préservant totalement sa capacité à reconnaître les visages. Cette double dissociation fonctionnelle a été interprétée comme suggérant l'existence d'un processus de reconnaissance spécialisé pouvant être affecté tout en préservant la reconnaissance des autres objets de l'environnement.

Suivant la piste tracée par cette indication, plusieurs études d'imagerie ont également conclu à l'existence de certaines régions cérébrales répondant spécifiquement aux visages humains. Bien que certaines activations ont été observées dans le sulcus

temporal supérieur (STS) ainsi que dans l'*occipital face area* (OFA), l'effet le plus robuste est typiquement observé dans le gyrus fusiforme médian de l'hémisphère droit ayant été nommé pour cette raison la *Fusiform Face Area* (FFA) (Kanwisher, McDermott, & Chun, 1997). Cette activation spécifique de la FFA par les visages a également pu être observée par l'utilisation de l'électroencéphalographie (EEG) (Bentin et al., 1996), de la magnétoencéphalographie (MEG) (Liu, Higuchi, Marantz, & Kanwisher, 2000) ainsi que de l'enregistrement sous-dural de surface (Allison, Puce, Spencer & McCarthy, 1999; McCarthy, Puce, Belger & Allison, 1999; Puce, Allison & McCarthy, 1999). De plus, l'activité de cette région a même pu être observée lorsque les participants ont rapporté percevoir des visages plutôt qu'un vase dans la célèbre image vase-visages de Rubin (Andrews et al., 2002; Hasson, Hendler, Ben Bashat & Malach, 2001). Une telle région responsable spécifiquement du traitement des visages a également été observée dans le cerveau du macaque par l'enregistrement unicellulaire et ce, dans une région analogue à la FFA où 97% des neurones enregistrées présentaient une sélectivité aux visages (Tsao, Freiwald, Tootell & Livingstone, 2006).

1.1 FFA : Région d'expertise dans le traitement visuel ou spécialisée pour le traitement des visages?

Une certaine controverse demeure toutefois quant à la spécificité du traitement opéré dans la FFA ainsi qu'à la nature du traitement y étant effectué. Gauthier et

collègues (2000) ont entre autres suggéré que la FFA serait une aire corticale responsable de la spécialisation dans le traitement visuel et non spécifiquement réservée aux visages. Cette position a été appuyée par le fait que des experts de la reconnaissance visuelle d'oiseaux, de voitures et de « greebles » (stimuli artificiels visant à imiter le niveau de complexité visuelle des visages) recrutent également l'activité de la FFA pour reconnaître ces objets (Gauthier et al. 2000). Une revue complète des démonstrations appuyant chacune des positions est secondaire à l'objectif du présent manuscrit mais il est toutefois important de mentionner que la spécificité de la FFA pour la reconnaissance des visages fait actuellement l'objet de multiples débats dans la littérature. En plus des indices anatomiques, un certain nombre d'observations comportementales issues de la psychophysique viennent également appuyer cette hypothèse comme nous le verrons à la section suivante (voir aussi la section « La notion de distances inter-attributs et l'hypothèse holistique » de la discussion générale).

2. Les indications comportementales d'un processus cognitif spécialisé pour la reconnaissance des visages

Bien que très informatives sur le plan anatomique, les études précédemment citées sont demeurées peu utiles pour révéler la nature du processus de reconnaissance ainsi que l'information recrutée par le système visuel afin d'accomplir cette tâche. De façon complémentaire, des études comportementales ont utilisé une grande variété de

techniques afin d'aborder cette question. Le domaine a toutefois souffert d'une certaine confusion dans la définition et l'opérationnalisation de quelques concepts clés entraînant une grande diversité dans l'utilisation de la terminologie propre à la reconnaissance des visages. Ici, nous présenterons un portrait des différents indices empiriques appuyant les théories ainsi que la terminologie telle que décrite par deux revues récentes du domaine (Maurer, Le Grand & Mondloch, 2002; Rossion, 2008).

2.1 Les attributs internes du visage humain

Une très grande partie des études comportementales ayant porté sur les visages dans les années 70 et 80 ont débuté par tenter de déterminer les attributs internes du visages utilisés afin d'accomplir différents traitements visuels. Ces études ont permis d'établir que, parmi les attributs internes, la combinaison des yeux et des sourcils véhiculait l'information la plus utilisée pour accomplir la plupart des tâches de reconnaissance (Davies, Ellis, & Sheperd, 1977; Haig, 1985, 1986a, 1986b; Sheperd, Davies, & Ellis, 1981; Walker-Smith, Gale, & Findlay, 1977). Ces études reposaient essentiellement sur le masquage ou l'isolation d'attributs faciaux afin d'en déterminer l'utilité pour des fins de reconnaissance. Ces techniques ont toutefois rapidement fait l'objet de critiques puisqu'il était communément accepté que la présentation d'attributs internes isolés durant plusieurs milliers d'essais ne pouvait que mener à des conclusions impliquant ces dits attributs isolés. Certains chercheurs ont également fait remarquer que

ces techniques s'avéreraient peu informatives sur l'intégration et l'interdépendance des attributs internes. Afin de s'attarder à cet aspect du processus, un pan important de la recherche s'est orienté vers l'étude de l'effet d'inversion, de l'effet composite et de l'effet partie-tout. Par ailleurs, il est important de noter à ce stade que les résultats obtenus par le masquage et par l'isolation des attributs ont entre autres pu être corroborés plus récemment par des techniques psychophysiques telles que par la technique Bubbles (Gosselin & Schyns, 2001, 2005; Schyns, Bonnar & Gosselin, 2002) ainsi que par la technique de classification des réponses (Sekular, Gaspar, Gold & Bennett, 2004). Par opposition à l'isolation, ces techniques présentent l'avantage d'échantillonner la totalité du visage et donc de présenter à l'observateur des stimuli plus représentatifs de la reconnaissance des visages dans le monde réel. Nous y reviendrons à la section « Critiques du paradigme holistique/configural ».

3. L'hypothèse holistique/configurale

Des suites des études introduites à la section précédente, les chercheurs du domaine de la reconnaissance des visages ont étudié de façon exhaustive principalement trois effets visuels étant fréquemment conçus comme spécifique aux visages. Les conclusions tracées par ces observations ont mené à l'élaboration de l'hypothèse actuellement dominante dans le domaine; celle du traitement holistique/configural. La

section suivante brosse un portrait des indices issues de chacun de ces effets ainsi que de leur apport à la définition des processus holistiques/configuraux.

2.2.1 L'effet d'inversion

Une étude classique de Yin (1969) a été la première à rapporter un phénomène qui deviendra le pilier le plus important de l'hypothèse du traitement configural/holistique des visages; l'effet d'inversion des visages. L'auteur de cette étude a rapporté que l'inversion d'un visage mène à une diminution marquée de sa reconnaissance ainsi qu'à une augmentation des temps de réponses comparativement à une présentation dans le sens traditionnel d'observation (Yin, 1969). Cet effet semble être spécifique aux visages puisque d'autres objets mono-orientés (e.g. des chaises, des maisons et des avions) ne présentent habituellement pas un tel effet (Boutet, Collin, & Faubert, 2003; Leder & Carbon, 2006; Robbins & McKone, 2006). Une myriade d'études a suivi cette première observation et l'effet d'inversion a pu être observé pour des visages familiers et non-familiers (Collishaw & Hole, 2000) ainsi que dans une variété de conditions expérimentales (Pour une revue : Rossion, 2008; Rossion & Gauthier, 2002). L'effet d'inversion des visages a d'ailleurs été décrit comme l'un des phénomènes les plus robustes dans la littérature sur le traitement des visages.

L'un des aspects les plus notables de cet effet est que l'inversion semble affecter particulièrement la reconnaissance lorsque les distances entre les attributs internes

(sourcils, yeux, nez et bouche) ont été modifiées (figure 1 a) et ce, comparativement à des visages ayant encouru une modification des attributs eux-mêmes (figure 1 b) (Freire & Symons, 2000; Leder & Carbon, 2006; Le Grand, Mondloch, Maurer & Brent, 2001; Mondloch, Le Grand, & Maurer, 2002; Rhodes, Brake & Atkinson, 1993). À la lumière de ces résultats, l'inversion a été proposée comme interférante avec le traitement des distances euclidiennes entre les attributs du visage modifiant ainsi qualitativement le processus de reconnaissance (Diamond & Carey, 1986; Leder & Bruce, 1998, 2000; Rhodes, 1988). Certains chercheurs ont également proposé que l'effet d'inversion soit observé conséquemment à la perturbation du « traitement holistique » c'est-à-dire de « l'intégration simultanée des multiples attributs d'un visage dans une représentation perceptuelle unique » (Rossion, 2008 p.275; voir également Farah, Tanaka, & Drain, 1995; Sargent, 1984; Tanaka & Farah, 1993). Cette proposition repose sur l'hypothèse que l'inversion perturberait l'intégration des divers attributs internes en un seul et même percept et que le système visuel devrait donc utiliser une stratégie de traitement attribut par attribut pour reconnaître un visage inversé.

Bien qu'une certaine confusion persiste quant à la nature et à la définition des processus holistiques et configuraux, il est généralement admis que l'effet d'inversion soit expliqué par une altération *qualitative* du processus de reconnaissance en interférant avec le traitement des distances inter-attributs. L'effet d'inversion des visages est pour cette raison fréquemment cité comme la démonstration comportementale la plus

importante en faveur d'un processus spécialisé pour le traitement des visages (Yovel & Kanwisher, 2004).



Figure 1. (A) Visages créés par la manipulation des distances entre les attributs internes (B) Visages créés par le remplacement des attributs internes par ceux d'autres visages tout en préservant les distances inter-attributs (Mondloch et al., 2002).

2.2.2 L'effet composite

L'hypothèse « holistique » introduite ci-dessus est également appuyée empiriquement par l'effet composite des visages (Young, Hellawell, & Hay, 1987); une illusion observée lorsque la partie du haut d'un visage est perçue comme étant sensiblement différente lorsqu'elle est alignée avec des parties du bas issus de visages

différents (figure 2 a). Cette illusion a été interprétée comme indiquant que les attributs d'un visage ne peuvent pas être perçus isolément et que la perception de l'une des deux parties (par exemple, le haut du visage dans la figure 2a) dépend de l'autre partie (ici, le bas du visage) (Endo, Masame, & Maruyama, 1989; Goffaux & Rossion, 2006; Hole, 1994; Le Grand, Mondloch, Maurer, & Brent, 2004).

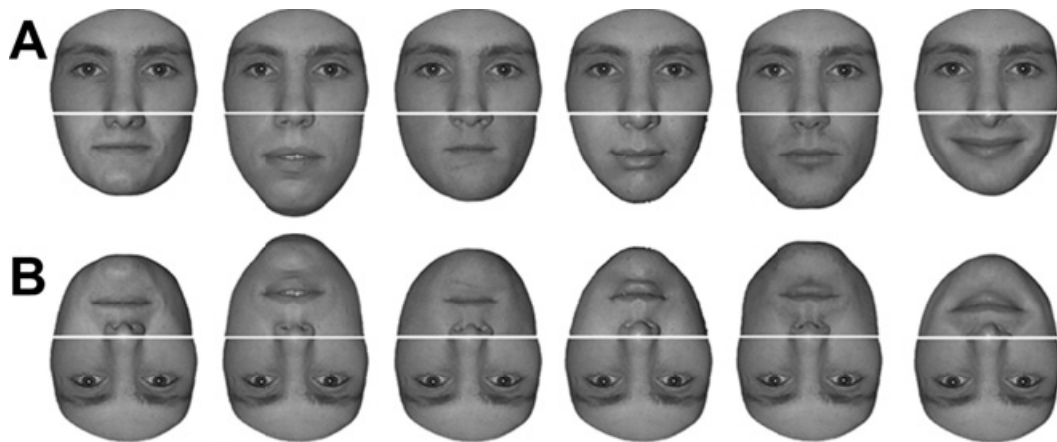


Figure 2. (A) Visages composites générés par l'utilisation du même haut de visage (au dessus de la ligne blanche) en combinaison avec plusieurs bas de visage différents (sous la ligne blanche). (B) Lorsque ces visages sont inversés l'effet composite disparaît (Rossion, 2008).

Une interprétation de cet effet propose que les attributs internes du visage soient si fortement intégrés entre eux lorsqu'un visage est traité par le système visuel qu'il est difficile de séparer cette représentation pour opérer une reconnaissance attribut par attribut (Hole, 1994). Il est intéressant de noter que l'effet disparaît complètement lorsque les deux parties d'un visage « composite » sont désalignées ou lorsque les

visages sont inversés (Figure 2b) supposément parce que ces deux manipulations perturberaient le traitement holistique du visage (e.g. Goffaux & Rossion, 2006).

2.2.3 L'effet partie-tout

L'effet partie-tout est un dernier type de manipulation ayant été utilisé pour tester l'hypothèse selon laquelle les visages seraient traités holistiquement (E.g. Davidoff & Donnelly, 1990; Tanaka & Farah, 1993). Cet effet est observé lorsque la reconnaissance d'attributs internes (e.g. les yeux) est facilitée du fait que l'attribut soit présenté avec les autres attributs présent lors de l'encodage (e.g. les yeux de Jim présentés avec le visage de Jim vs les yeux de Tom avec le visage de Jim) plutôt que lorsqu'il est présenté seul (e.g. Les yeux de Jim vs les yeux de Bob) (figure 3) (McKone, Kanwisher & Duchaine, 2007). Cet effet semble d'ailleurs être spécifique aux visages présentés dans l'orientation normale puisque des visages ayant les attributs internes mélangés, des visages inversés ainsi que des maisons ne présentaient pas une telle facilitation de la reconnaissance (Tanaka & Farah, 1993). Il est toutefois notable qu'un effet analogue existe pour l'identification d'une lettre isolée par rapport à la reconnaissance de cette lettre dans un mot—l'effet de supériorité du mot (Cattell, 1881; Reicher, 1969). Tanaka et Sengco (1997) ont observé que l'effet partie-tout est réduit lorsque les distances inter-attributs sont modifiées entre la phase d'apprentissage et la phase de test; ce qui

suggérerait une implication de ces distances dans le processus holistique de reconnaissance.

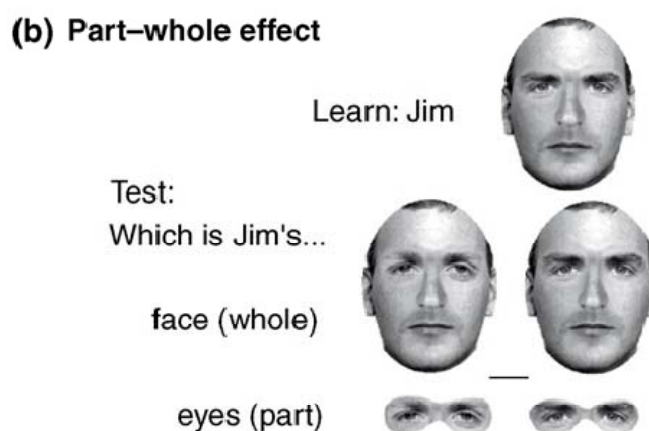


Figure 3. Illustration d'une tâche visant à étudier l'effet partie-tout (McKone, Kanwisher & Duchaine, 2006).

Il est important de noter que les termes holistique et configural ont été étroitement associés dans la littérature, si bien que plusieurs auteurs utilisent ces termes de façon interchangeable. Par exemple, dans une revue de la littérature, Maurer et collègues (2002) ont décrit le traitement configural comme « un phénomène impliquant la perception des relations entre les attributs d'un stimulus » (Traduction libre, p.255) et, selon ces auteurs, le traitement holistique (ici pris au sens de traitement de l'ensemble des attributs du visage) serait un sous-type de traitement configural c'est-à-dire qu'il reposerait sur une perception des relations entre les attributs. Rossion (2008) a proposé que le traitement holistique soit le processus général d'intégration des attributs en une représentation unique et que la manipulation de l'information configurale (ici décrite

comme les distances métriques entre les attributs) perturberait ce processus. Quoiqu'il en soit, il importe de souligner à cette étape que le processus holistique est typiquement conçu comme étant très associé au traitement des distances inter-attributs.

4. Critique de la conception et de l'utilisation des distances inter-attributs

L'hypothèse holistique/configural semble avoir été très étroitement associée à une prédominance du traitement des distances inter-attributs dans le processus de reconnaissance des visages. Cette idée semble être si répandue que selon plusieurs auteurs, la zeitgest contemporaine en reconnaissance des visages suggère que le processus de reconnaissance reposerait essentiellement sur le traitement des distances inter-attributs (e.g. Maurer, Le Grand & Mondloch, 2002). Il est surprenant de réaliser que peu de démonstrations empiriques indiquent que ces distances soient essentielles ou même utilisées lors de la reconnaissance de visages présentés à l'endroit. Cette perception semble prendre racine dans les travaux de Diamond et Carey (1977, 1986), Haig et collègues (1984; Hosie, Ellis, & Haig, 1988) et Rhodes (1988) bien qu'aucune de ces études n'ait évalué directement la question. Dans les prochains paragraphes, nous résumerons brièvement ces études.

Diamond et Carey (1977) ont observé qu'une modification dans la préférence des observateurs à utiliser les attributs internes du visage pour des fins de reconnaissance survient vers 10 ou 12 ans. Ces auteurs ont émis l'hypothèse que cette

modification comportementale pourrait être associée à une préférence chez les adultes pour l'utilisation des distances inter-attributs pour la reconnaissance des visages. Ils mirent cette hypothèse à l'épreuve dans une étude classique ayant évalué l'effet d'inversion de visages, de chiens et de paysages (Diamond & Carey, 1986). Les auteurs ont proposé que pour observer un effet d'inversion, il était nécessaire de pouvoir individualiser des images de la présentation sur la base de l'information relationnelle contenue entre les différentes parties des images. Toutefois, il est important de noter que les auteurs n'ont pas manipulé explicitement ces distances. L'utilité de cette information pour la reconnaissance des visages présentés à l'endroit n'a donc jamais été évaluée dans cette étude. De plus, il est intéressant de noter que les auteurs ont également observé un effet d'inversion pour les images de chiens et que l'amplitude de l'effet était corrélée à l'expertise des observateurs à la reconnaissance des images de la catégorie en question (e.g. expert des bergers allemands).

En utilisant une technique différente, Haig (1984) a tenté d'évaluer l'acuité humaine à discriminer des variations de distances entre les attributs du visage en mesurant la plus petite différence détectable par les observateurs (*just noticeable difference*). Les résultats de cette étude ont révélé que les observateurs humains sont aptes à reconnaître de très petites variations de distances parfois très près de la limite de l'acuité humaine pour certains attributs (e.g. monter la bouche). De plus, ce chercheur a montré que deux manipulations quantitativement similaires pouvaient entraîner des jugements différents selon les attributs et le sens du déplacement (e.g. bouger les yeux

vers l'intérieur ou vers l'extérieur). Ces manipulations ne fournissent toutefois aucun indice quant à l'utilisation de cette aptitude lors de la reconnaissance des visages.

Finalement, Rhode (1988) a utilisé la « *multidimensional scaling analysis* » afin d'évaluer l'apport d'un grand éventail d'éléments (e.g. âge, genre, attributs internes, distances inter-attributs...) au jugement de similarité entre deux visages. Les conclusions de cette étude ont indiqué que les attributs ainsi que les distances entre ces derniers étaient d'importance égale et que les « relations de second-ordre » (décrites comme les distances inter-attributs) ne pouvaient pas être comprises comme étant plus importantes que les attributs pour la reconnaissance » (Rhode, 1988).

Malgré cette dernière indication, l'idée que les distances entre les attributs sont primordiales pour la reconnaissance des visages est demeurée très centrale dans la littérature (e.g. Maurer et al., 2002). En suivant l'intuition de Diamond et Carey (1986), plusieurs auteurs ont explicitement manipulé les distances entre les attributs internes de visages utilisés dans des tâches expérimentales visant à étudier la prosopagnosie acquise (Barton, Press, Keenan, & O'Connor, 2002), la plasticité visuelle suivant le traitement de cataracte congénitale (Le Grand, Mondloch, Maurer & Brent, 2001) et, de façon plus systématique, l'effet d'inversion des visages.

Tel qu'indiqué précédemment, plusieurs chercheurs ayant étudié l'effet d'inversion des visages en sont venus à la conclusion que l'inversion perturbe essentiellement le traitement des distances entre les attributs internes (Barton, Keenan & O'Connor, 2001; Cabeza & Kato, 2000; Collishaw & Hole, 2000; Freire, Lee, &

Symons, 2000; Goffaux, 2008; Goffaux & Rossion, 2007; Le Grand et al., 2001; Leder & Bruce, 2000; Leder, Candrian, Huber, & Bruce, 2001; Leder & Carbon, 2006; Rhodes, Brake, & Atkinson, 1993; Rhodes, Hayward, & Winkler, 2007; Searcy & Bartlett, 1996; voir Rossion, 2008 pour une revue). Par contre, observer que l'inversion affecte d'avantage le traitement des distances inter-attributs que le traitement des attributs eux-mêmes ne constitue pas une preuve de l'utilité de ces distances lorsque les visages sont présentés à l'endroit. Au mieux, l'effet d'inversion des visages indique que cette difficulté est une simple conséquence de l'inversion (voir Rossion, 2008). Il est même possible que cette inaptitude à utiliser uniquement des distances inter-attributs pour reconnaître des visages inversés ne soit que le reflet de la capacité limitée du système visuel à traiter cette information même lorsque les visages sont présentés à l'endroit (voir la discussion de l'article central du mémoire). Il n'existe, à notre connaissance, aucune démonstration directe de l'utilité des distances inter-attributs pour la reconnaissance des visages présentés à l'endroit.

3.1 La notion de distance inter-attributs

Les manipulations expérimentales ayant été opérées sur les distances entre les attributs ont généralement été faites avec peu d'égards pour le réalisme ou, ce que les chercheurs en vision appellent, la « validité écologique » des stimuli utilisés. La majorité des chercheurs s'est bornée à altérer des photos « dans les limites du

plausible », c'est-à-dire guidés seulement par l'intuition, à l'aide d'outils informatiques permettant de sélectionner et de déplacer les attributs internes tel que Photoshop (Rhodes et al., 1993; Barton et al., 2001; Freire et al., 2000; Le Grand et al., 2001; Leder & Bruce, 2000; Leder et al., 2001; Leder & Carbon, 2006; Rhodes et al., 2007; Goffaux, 2008; Goffaux & Rossion, 2007). Par souci de réalisme, une minorité d'auteurs a utilisé les valeurs rapportées par une vaste étude anthropométrique (Farkas, 1981) comme ligne directrice (e.g. Le Grand et al., 2001; Mondloch et al., 2002, Mondloch, Geldart, Maurer, & Le Grand, 2003; Hayden, Bhatt, Reed, Corbly & Joseph, 2007). Toutefois, les statistiques rapportées par Farkas (1981) ne contiennent que les distances absolues moyennes (ainsi que leurs écart-types) entre les divers attributs internes. Modifier la distance entre deux attributs en fonction de la moyenne n'assure aucunement la création de stimuli réalistes puisqu'il est possible que les distances entre les attributs déplacés et les autres attributs deviennent irréalistes. Les distances entre les attributs internes contiennent beaucoup de covariance (voir la discussion générale de l'article). Est-ce réaliste, par exemple, d'augmenter la distance entre les yeux d'un visage pour atteindre la distance moyenne (selon Farkas, 1981) alors que le reste des distances inter-attributs de ce visage se retrouvent 2 écart-types sous leurs moyennes respectives? De plus, il est peu probable que le système visuel utilise des distances absolues pour reconnaître les visages puisque ceci nécessiterait une connaissance précise de la distance de présentation du visage observé. Effectivement, sans cette information un objet étant deux fois plus large qu'un autre pourrait occuper une même région sur la rétine que ce

dernier simplement en étant à une plus grande distance de présentation. Pour ces raisons, il est plus probable que les humains encodent des ratios de distances de façon à permettre une reconnaissance à plusieurs distances de présentation (Rhodes, 1988 ; pour une revue voir Tsao & Livingstone, 2008). Autrement dit, la taille absolue des visages est une source de variance dans les données de Farkas qui n'est probablement pas pertinente à la reconnaissance des visages.

5. Problématique

Il est donc légitime de se questionner sur la nature des variations de distances inter-attributs et sur l'information véhiculée par de telles variations. Sans cette connaissance, il est impossible de déterminer si les observateurs humains sont effectivement capables d'utiliser ces variations pour reconnaître leurs semblables. Il sera donc question, dans un premier temps, d'offrir un portrait précis de la variation des distances inter-attributs à l'intérieur du visage humain une fois la variance due à la taille absolue des visages éliminée. Pour ce faire, les attributs internes d'une grande quantité de visages seront annotés et chacun des visages sera « aligné » sur la position moyenne des attributs de tous les visages de la banque de façon à éliminer l'information associée à la taille absolue des visages. Des stimuli de visages dont les distances inter-attributs ont été modifiées venant d'études déjà publiées seront également annotés et alignés de la même façon pour évaluer leur réalisme.

Une fois l'annotation complétée, il sera alors possible de décrire l'étendue de la variation résiduelle des distances. Nous montrerons, ensuite, que ces variations résiduelles de distances contiennent de l'information pouvant servir à résoudre différentes tâches de reconnaissance de visages. Finalement, nous testerons deux hypothèses empiriquement : l'hypothèse suivant laquelle ces distances inter-attributs résiduelles sont suffisantes aux observateurs humains pour la reconnaissance des visages; et l'hypothèse suivant laquelle ces distances seraient nécessaires aux observateurs humains pour la reconnaissance des visages.

6. Contribution à l'article

À titre de premier auteur de l'article suivant, j'ai contribué à la réalisation expérimentale d'une idée originale de Frédéric Gosselin en opérant la cueillette des données, une partie des analyses ainsi que la rédaction d'un premier jet du manuscrit. De plus, j'ai participé avec Frédéric Gosselin à la modification du protocole expérimental initial suite à l'obtention de résultats préliminaires. Bruno Rossion, Philippe Schyns et Frédéric Gosselin ont contribué de façon très importante à la rédaction du manuscrit final en apportant de nombreuses idées d'analyses et en faisant plusieurs apports importants au texte de l'introduction et de la discussion générale. Frédéric Gosselin a également effectué plusieurs des analyses présentées dans le manuscrit. Tous les coauteurs consentent à ce que cet article fasse partie du présent mémoire, comme l'atteste le formulaire suivant.

Déclaration des coauteurs de l'étude

1. Identification de l'étudiant

Vincent Taschereau-Dumouchel
M.Sc. Psychologie (2-220-1-3)

2. Description de l'article

Auteurs : Vincent Taschereau-Dumouchel, Bruno Rossion, Philippe G. Schyns et Frédéric Gosselin.

Titre : Interattribute distances do not represent the identity of real-world faces
Soumis à la revue *Psychological Science* le 29 juillet 2009.

3. Déclaration des coauteurs

À titre de coauteur de l'article identifié ci-dessus, je suis d'accord pour que Vincent Taschereau-Dumouchel incluse cet article dans son mémoire de maîtrise qui a pour titre : *Les distances entre les attributs internes du visage humain*

Bruno Rossion

Coauteur	Signature	Date
----------	-----------	------

Phillippe G. Schyns

Coauteur	Signature	Date
----------	-----------	------

Frédéric Gosselin

Coauteur	Signature	Date
----------	-----------	------

ARTICLE

Running head: INTERATTRIBUTE DISTANCES IN HUMAN FACES

Interattribute distances do not represent the identity of
real-world faces

Vincent Taschereau-Dumouchel

Département de psychologie, Université de Montréal

Bruno Rossion

Unité Cognition & Développement, Faculté de Psychologie, Université catholique de
Louvain

Philippe G. Schyns

Department of Psychology, University of Glasgow

Frédéric Gosselin

Département de psychologie, Université de Montréal

Abstract

According to an influential view, based on studies of development and of the face inversion effect, human face recognition relies mainly on the treatment of the distances among internal facial features. However, there is surprisingly little evidence supporting this claim. Here, we first use a sample of 515 face photographs to estimate the face recognition information available in interattribute distances. We demonstrate that previous studies of interattribute distances generated faces that exaggerated by 400% this information compared to real-world faces. When interattribute distances are sampled from a real-world distribution, we show that human observers recognize faces poorly (with a maximum accuracy of 65%) across a broad range of viewing distances. In contrast, recognition performance is restored when observers only use the facial attributes of real-world faces. We conclude that facial cues other than interattribute distances such as attribute shapes and skin properties are the dominant information of face recognition mechanisms.

According to an influential view, human face processing rests mainly on *interattribute distances*¹ (e.g. inter-ocular distance, mouth-nose distance; Diamond & Carey, 1986; Carey, 1992; Maurer, Le Grand, & Mondloch, 2002). After briefly reviewing the origin of this claim, we will examine two points that have perhaps surprisingly been neglected so far: If face processing relies on interattribute distances then, surely, (1) real-world interattribute distances must contain useful information for face recognition; and (2) human observers must be more sensitive to these natural variations than to those of other facial cues.

The origin of the idea that relative distances between features are important for individual face processing can be traced back to the work of Haig (1984), and Diamond and Carey (1986). Haig (1984) moved the different features of a few unfamiliar faces independently by small amounts and measured the just noticeable differences of five observers for all manipulations (e.g. mouth-up, eyes inward) with respect to the original face. He noticed that the sensitivity of human adults to slight alterations in the positions of the features of a set of faces was quite good, at the limit of visual acuity for some alterations (e.g. mouth up), and that it varied greatly for different kinds of manipulations

¹ We intentionally avoided to use the expression "configuration" because it is ambiguous in the face recognition literature: It can refer to either to the relative distances between attributes (e.g. Maurer, Le Grand, & Mondloch, 2002), or to a way of processing the face ("configural processing", as used by Young et al., 1987)—"holistically" or as a gestalt. All face cues, including attribute shapes and skin properties, are "con \square gural" under the latter interpretation. 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. Maurer, Le Grand, & Mondloch, 2002). 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.

that were identical quantitatively (e.g. eyes inward vs. eyes outward). However, the ranges of these manipulations were arbitrary with respect to normal variations of feature positions in real life, and there was no assessment of the critical role of such manipulations in actual face identification tasks relative to featural changes.

Based on their developmental studies and their work on visual expertise with non-face objects, Diamond and Carey (1977; 1986) introduced the hypothesis that what makes faces special with respect to other object categories is the expert ability to distinguish among individuals of the category (i.e. different faces) based on what they so-called “second-order relational properties”, namely the idiosyncratic variations of distances between features. However, while these authors claimed that the ability to extract such second-order relational properties would be a key factor in accounting for the slow development of face processing abilities (Carey, 1992), they did not test this hypothesis in their paper or subsequent studies.

Studies of face inversion have also contributed to the idea that relative distances between attributes are fundamental for face recognition. Faces rotated by 180 degrees in the picture-plane induce important decreases in recognition accuracy and increasing response latencies (e.g. Hochberg & Galper, 1967). This impaired performance is disproportionately larger for faces in contrast to other mono-oriented objects such as houses and airplanes (Leder & Carbon, 2006; Robbins & McKone, 2006; Yin, 1969; for a review see Rossion & Gauthier, 2002). Thus face inversion has been used by researchers as a tool to isolate what’s special about upright face processing. It happens that the processing of interattribute distances is more affected by inversion than the

processing of the shape of attributes (Diamond & Carey, 1977, 1986; Haig, 1984; Hosie et al., 1988).

This last observation has been taken as supporting the view that relative *distances* between facial features are fundamental or most diagnostic for individuating upright faces (e.g. Diamond & Carey, 1977, 1986). However, two critical links are missing in the reasoning. First, there is no direct evidence that interattribute distances are diagnostic for upright face recognition. In fact, there is tentative evidence that interattribute distances might not be the main source of information for face recognition: Exaggerated interattribute distances do not impair recognition much (Tanaka & Sengco, 1997; Biederman & Kalocsai, 1997); interattribute distances are less useful in similarity judgments than attribute shape (Rhodes, 1988). Second, there is no direct evidence that a difficulty to process interattribute distances is the *cause* of the FIE. For example, this difficulty is predicted by Tanaka and Farah's (1991, 1993) proposal that face inversion leads to a loss of the ability to process the face as a gestalt or "holistically" (see Rossion, 2008, for a discussion).

To address this issue, a first question should be how much do faces vary in interattribute distances in the real world? Clearly, if there is little objective, real-world interattribute variation, there is little that the visual system could and should do with it. And if there is objective information, how much of it would human observers use for face recognition?

To address these questions, in Experiment 1 we estimated, from a sample of 515 full-frontal real-world Caucasian faces, how much information was objectively available to the human brain in relative interattribute distances for gender discrimination and face identification. We demonstrate that while there is objective interattribute distance information in faces, most previous studies have grossly exaggerated this information when testing it (on average 400%). In Experiment 2, we compared face recognition when interattribute distances are the only information source available (Experiment 2a) and unavailable (Experiment 2b), and we show that performance is much better in the latter case.

Experiment 1

Methods

Participants. Three female students (all 19 years old) from the Université de Montréal received course credits to annotate digital portraits on 20 internal facial feature landmarks (see Experiment 1, Procedure). The first author (22 years old) annotated faces from previous studies in which the distances between internal features had been altered. Participants had normal or corrected to normal vision.

Stimuli. A total of 515 Caucasian frontal-view real-world portraits presenting a neutral expression (256 females) were used. These faces came from multiple sources: the entire 300-face set of Dupuis-Roy, Fortin, Fiset and Gosselin (2009), 146 neutral faces from the Karolinska Directed Emotional Faces, the entire 40-face set of Arguin and Saumier (unpublished), the 16 neutral faces from Schyns and Oliva (1999), the

seven neutral faces from the CAFE set, and six faces from the Ekman and Friesen (1975) set. We also annotated 86 stimuli used in 14 previous studies in which interattribute distances had been manipulated “within the limit of plausibility” (for a list, see Figure 2).

Apparatus. The annotations were made on a Macintosh G5 computer running functions written for Matlab (Align Tools OSX, available at <http://mapageweb.umontreal.ca/gosselif/align%20tools%20OSX/>) using functions from the Psychophysics Toolbox. Stimuli were presented on a HP p1230 monitor at a resolution of 1920 x 1200 pixels with a 100 Hz refresh rate. The monitor luminance ranged from 1.30 to 80.9 cd/m².

Procedure. Participants were asked to place 20 points on specific landmarks of internal facial features with a computer mouse, one face at a time (see blue crosses in the leftmost column of Figure 1). These landmarks were chosen because they are easy to locate and allow a proper segmentation of the features. We increased the size of stimuli to match computer monitor resolution to ease the task of participants (Okada, von der Malsburg, & Akamatsu, 1999). Every participant annotated each of the 515 portraits in random order allowing us to estimate inter-subject annotation error; and every participant annotated 10% of these portraits twice allowing us to estimate intra-subject annotation error.

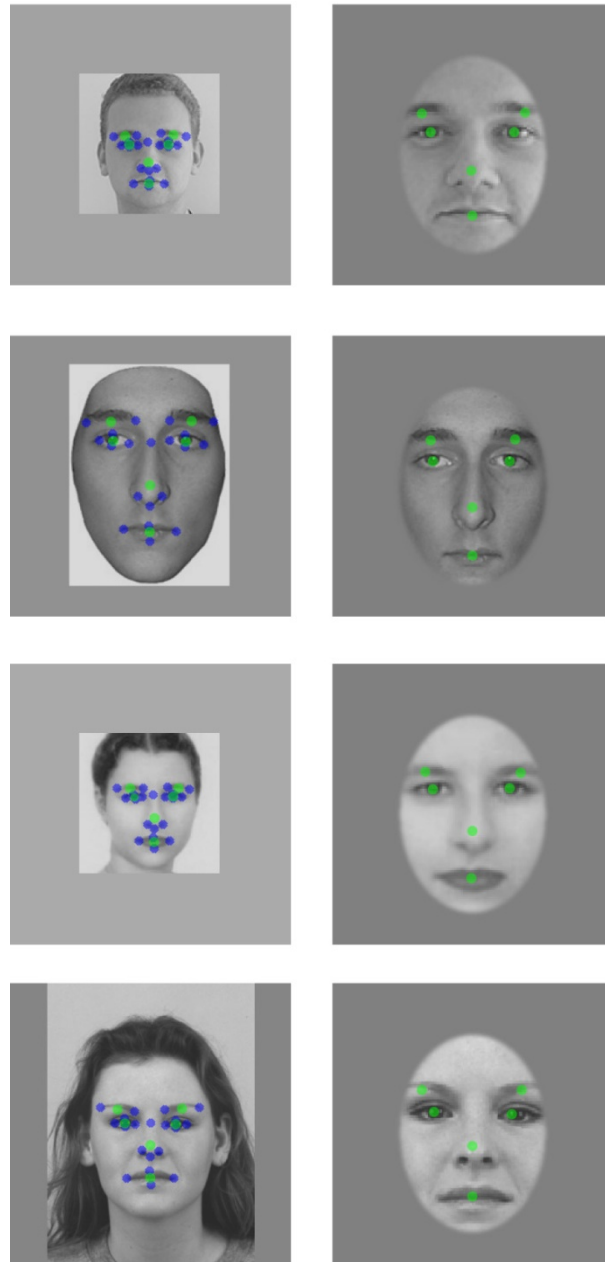


Figure 4. Leftmost column: Sample faces annotated in Experiment 1. The 20 blue crosses show, for these faces, the average annotations across participants. These 20 annotations were reduced to 6 attribute positions—green dots—by averaging the xy-coordinates of the annotations belonging to every attribute. Rightmost column: We translated, rotated, and scaled the attribute positions of each face to minimize the mean square of the difference between them and the average attribute positions across faces. The residual differences between aligned attribute positions—green dots—is the interattribute variance in the real-world.

Results

We reduced each set of 20 annotations to 6 feature positions by averaging the xy-coordinates of the annotations placed on landmarks belonging to the same facial feature (see green dots in the leftmost column of Figure 1) to disentangle attribute position from attribute shape. Indeed, to manipulate interattribute distances independently from attribute shape, whole attributes are typically cropped—including, in our case, all the pixels annotated by our observers on each of these attributes—and translated (e.g. Maurer, Le Grand, & Mondloch, 2002).

This 20-to-6 reduction also maximizes signal-to-noise ratio of attribute position. Assuming that annotation error is the same for the x- and y- dimensions and for all features (and systematic error aside), the signal-to-noise ratio of the measurements is estimated at 8.27 (total std=2.98, intra-annotating participant std=0.69 and inter-annotating participant std=0.98 per feature and per dimension for a mean interocular distance of 100 pixels). For all attributes except the eyebrows, four points were averaged, and thus signal-to-noise ratio of attribute position is twice that for individual annotations (i.e., 16.54); for the eyebrows, two annotations were averaged, and thus, signal-to-noise ratio was $\sqrt{2}$ that for individual annotations (i.e., 11.70). In sum, the signal-to-noise ratio of attribute position was high.

To estimate absolute interattribute distances, the brain would have to estimate absolute depth precisely; such absolute depth estimates are only possible at really close range, which is atypical of face identification distances. Thus it is usually assumed that

only *relative* interattribute distances are available to the brain for face identification (Rhodes, 1988).

We "relativized" interattribute distances by translating, rotating, and scaling the feature positions of each face to minimize the mean square of the difference between them and the average feature positions across faces (rotated so that the y-axis was the main facial axis; see the rightmost column of Figure 1; Ullman, 1989). The resulting interattribute distances are proportional to the ones obtained by dividing the interattribute distances of each face by its mean interattribute distance. However, only our alignment procedure provides an intuitive way of visualizing the *variance* of interattribute distances. The green clouds in Figure 2 represent the distributions of the aligned feature positions of real-world faces (they are smoothed with a Gaussian kernel of standard deviation equal to 0.015 the average interocular distance). The variance of each distribution reflects the contribution of the corresponding attribute to the overall interattribute distance variance in the real-world. (See Appendix for a description of the *covariance* between the aligned attributes.) Red lines represent one standard deviation of the aligned positions along their first and second components in a principal component analysis (PCA). As can be seen at a glance, the pairs of red lines on the eyes and eyebrows are of similar lengths, which means that the variance in the positions of these features is roughly the same at all orientations. However, the pairs of red lines on the nose and mouth are clearly of different lengths—the variance is mainly organized along the main facial axis for these features. Henceforth, we thus focus on the main facial axis and the axis orthogonal to it.



Figure 5. Distribution of post-alignment attribute positions (green clouds) of the 515 annotated faces, with standard-deviation-length eigenvectors (red segments) centered on the distributions, and overlaid to a face to facilitate the interpretation. The blue dots are the attribute positions of stimuli of 14 previous studies that used distance manipulations (Barton, Keenan & Bass, 2001, *British Journal of Psychology*, 92, 527-549; Bhatt, Bertin, Hayden & Reed, 2005, *Child Development*, 76, 169-181; Freire, Lee & Symons, 2000, *Perception*, 29, 159-170; Goffaux, Hault, Michel, Vuong & Rossion, 2005, *Perception*, 34, 77-86; Haig, 1984; Hayden, et al. 2007; Hosie, Ellis & Haig, 1988; Leder & Bruce, 1998, *Quarterly Journal of Experimental Psychology: Section A*, 51, 449-473; Leder & Bruce, 2000, *The Quarterly Journal of Experimental Psychology*:

Section A, 53, 513-536; Leder, Candrian, Huber & Bruce, 2001, *Perception*, 30, 73-83; Le Grand, Mondloch, Maurer, & Brent, 2001; Rhodes, Brake & Atkinson, 1993, *Cognition*, 47, 25-57; Sargent, 1984, *British Journal of Psychology*, 75, 221-242; Tanaka, & Sengco, 1997).

We also plotted the attribute positions from 14 studies that explicitly manipulated interattribute distances (blue dots). Most of these dots fall outside two standard deviations of the real-world, normal range of variations. Specifically, more than 73% of the experimental faces had at least one attribute falling more than two standard deviations away from the mean of at least one axis of the real-world faces (23% of the eyes on both axes, and 26% of the noses and 29% of the mouths on the y-axis). Thus, in most of these 14 studies, artificial interattribute distances were exaggerated compared to natural variations. What was the impact of this exaggeration on the information available for face recognition?

To answer this question, we performed two virtual experiments (see Supplementary material for details). In the first one, we repeatedly trained a model to identify one randomly selected set of aligned attribute positions (which is equivalent to interattribute distances but represented more efficiently) from 50% of the other sets of natural aligned attribute positions also randomly selected, and tested this model on the remaining sets of natural aligned attribute positions. Similarly, in the second virtual experiment, we repeatedly trained a model to identify one randomly selected set of *artificial* aligned attribute positions from 50% of the sets of *natural* aligned attribute positions also randomly selected, and tested them on the remaining sets of *natural* aligned attribute positions. In each case, we found how much noise was necessary for

the models to perform with a fixed sensitivity ($A' = 0.7$). The models trained to identify the artificial faces required 4 times more noise ($var = 400$ for a mean interocular distance of 100 pixels) than the ones trained to identify the real-world faces ($var = 100$ for a mean interocular distance of 100 pixels). This implies that the interattribute distances of the artificial faces convey 400% more information for identification than in real-world faces.

In sum, there is information in interattribute distances for processing in real-world faces. However, not nearly as much as the majority of past studies have assumed.

Experiment 2

In Experiment 2a, we asked whether human observers can use this real-world interattribute distance information, at different viewing distances, to resolve a matching-to-sample (ABX) task when interattribute distance is the only information available. And, in Experiment 2b, we asked the complementary question: Can human observers use real-world cues other than interattribute distances such as attribute shapes and skin properties, at different viewing distances, to resolve an ABX.

Methods

Participants. Sixteen observers (eight females and eight males; aged between 19 and 29 years of age; mean=22.8 years; std=2.5 years) participated in Experiment 2a; and ten different observers (five females and five males; aged between 21 and 31 years of

age; mean=23.9 years; std=3.28 years) participated in Experiment 2b. All observers had normal or corrected to normal vision.

Stimuli. We created 2,350 pairs of stimuli for each experiment. Base faces were those annotated in Experiment 1. First, we translated, rotated, and scaled all these faces and corresponding feature positions to minimize the mean square of the difference between them and the average feature positions across faces rescaled to an interocular distance of 50 pixels (or 1.4 cm). To create one stimulus pair in Experiment 2a, we randomly selected three faces of the same gender from the bank of 515 faces. Feature masks were best-fitted to the annotations of every internal features of one of these faces—the *feature face*—using affine transformations. The pixels covered by the features masks were then translated to the feature positions of the other two faces—the *distance faces*—producing a pair of face stimuli (see Figure 3—see Appendix for an alternative method for creating realistic interattribute distances). Pixels falling outside the feature models were bilinearly interpolated from the *feature face* using the `interp` Matlab function. This procedure ensures that face stimuli from a pair have identical internal features and only differ on the distances between these features.

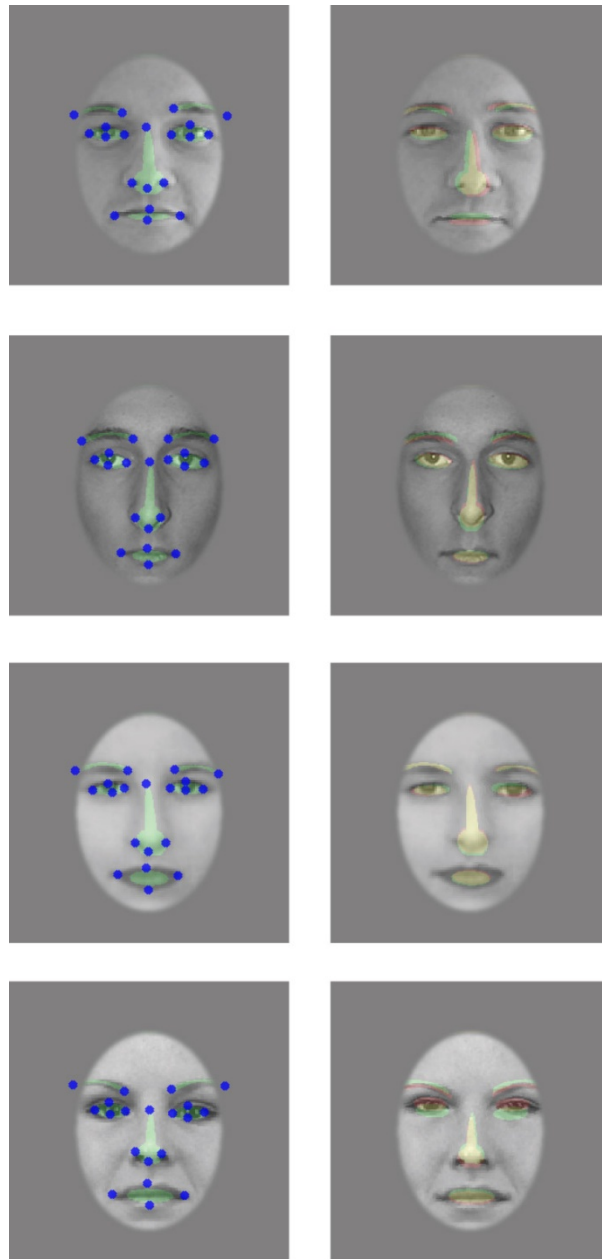


Figure 6. Leftmost column: In Experiment 2, feature masks—shown in translucid green—were bestfitted to the aligned annotations—represented by blue crosses. Rightmost column: In Experiment 2a, these feature masks were displaced according to the feature positions of another face of the same gender. Translucid green areas reproduce the feature masks of the leftmost column; translucid red areas represent the same feature masks after displacement; and translucid yellow areas represent the overlap between these two sets of feature masks.

In Experiment 2b we also randomly selected three faces of the same gender from the database. This time, however, we best-fitted feature models to the landmarks of the internal features of two of these faces—the *feature faces*—and the features were translated according to the feature positions of the third face—the *distance face*. The pixels falling outside the feature models were interpolated from the appropriate *feature face*. This procedure ensures that faces from a stimulus pair have identical interattribute distances but differ in cues other than interattribute distances, such as attribute shapes and skin properties.

All face stimuli were shown in grayscale, with equal luminance mean and variance, through a grey mask punctured by an elliptic aperture with a smooth edge (convolved with a Gaussian kernel with a standard deviation equal to 2 pixels) and with a horizontal diameter of 128 pixels and a vertical diameter of 186 pixels. This only revealed the inner facial features and their distances (for examples, see Figure 4).

Apparatus. Experiment 2 was performed on a Macintosh G5 running a computer script written for the Matlab environment using functions of the Psychophysics Toolbox. Stimuli were presented on a HP p1230 monitor at a resolution of 1920 x 1200 pixels at a refresh rate of 100Hz. The monitor luminance ranged from 1.30 to 80.9 cd/m².

Procedure. Participants completed 120 trials of their ABX task (the sequence of events in a trial is given in Figure 3) at each of five viewing distances in a randomized block design to equate the effect of learning. Viewing distances were equivalent to real-world viewing distances of 2, 3.4, 5.78, 9.82 and 16.7 m (which corresponds,

respectively, to average interocular widths of 1.79, 1.05, 0.62, 0.37, and 0.21 deg of visual angle). We used the interocular width average of 6.2 cm (mean for males=6.3 cm; and mean for females=6.1 cm) reported by Farkas (1981) to determine the equivalent real-world distances.

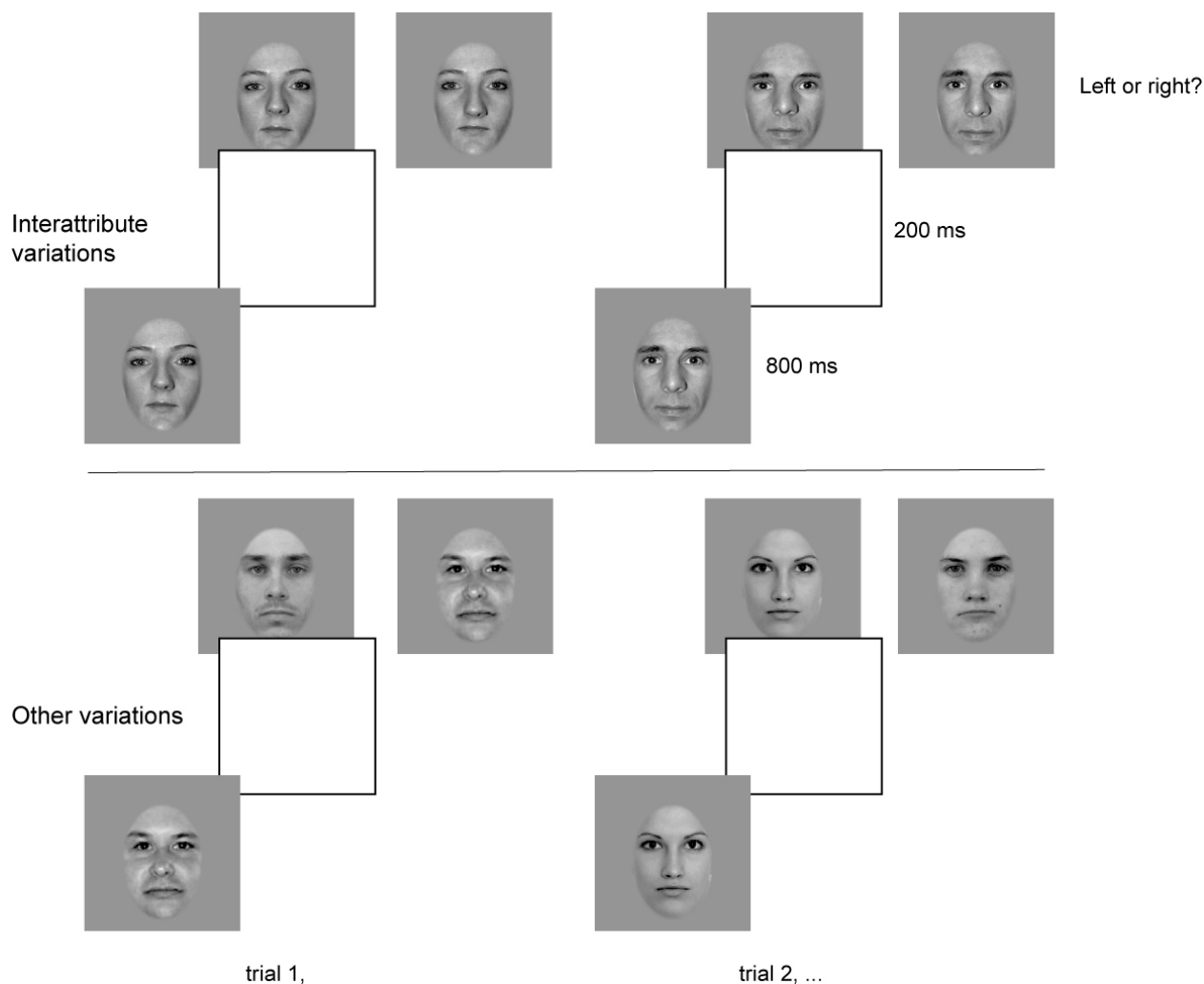


Figure 7. Sequence of events in two sample trials of our experiments. Top: In Experiment 2a, we asked whether human observers can use this real-world interattribute distance information, at different viewing distances, to resolve a matching-to-sample (ABX) task when interattribute distance is the only information available. Bottom: In Experiment 2b, we asked the complementary question: Can human observers use real-world cues other than interattribute distances such as attribute shapes and skin properties, at different viewing distances, to resolve an ABX.

On each trial, one stimulus from a pair (see Experiment 2, Stimuli) was randomly selected as the target. This target was then presented for 800 ms immediately followed by a blank presented for 200 ms immediately followed by the pair of stimuli presented side-by-side in a random order. Participants were asked to choose which face—on the left or on the right—was the target. No feedback was provided to the participants between trials.

Results

We submitted the results to a 2 x (5) mixed design ANOVA using *viewing distances* (2, 3.4, 5.78, 9.82 and 16.7 m) as a within-subjects factor and *group* (different vs. same interattribute distances) as a between-subjects factor. The test of within-subjects effect performed using the Greenhouse-Geisser correction revealed a significant interaction between viewing distances and groups ($F(2.4; 57.7)=4.89, p=.007, \eta^2=.17, p_{rep}=.96$). Contrasts of the between-subjects factor also revealed a significant difference of accuracy between both groups at all five viewing distances (all $F(1,24)>100, p<.00001, \eta^2>.80, p_{rep}\approx 1$). Observers who could solely rely on interattribute distances performed significantly lower at each of the five viewing distances (see Figure 5).

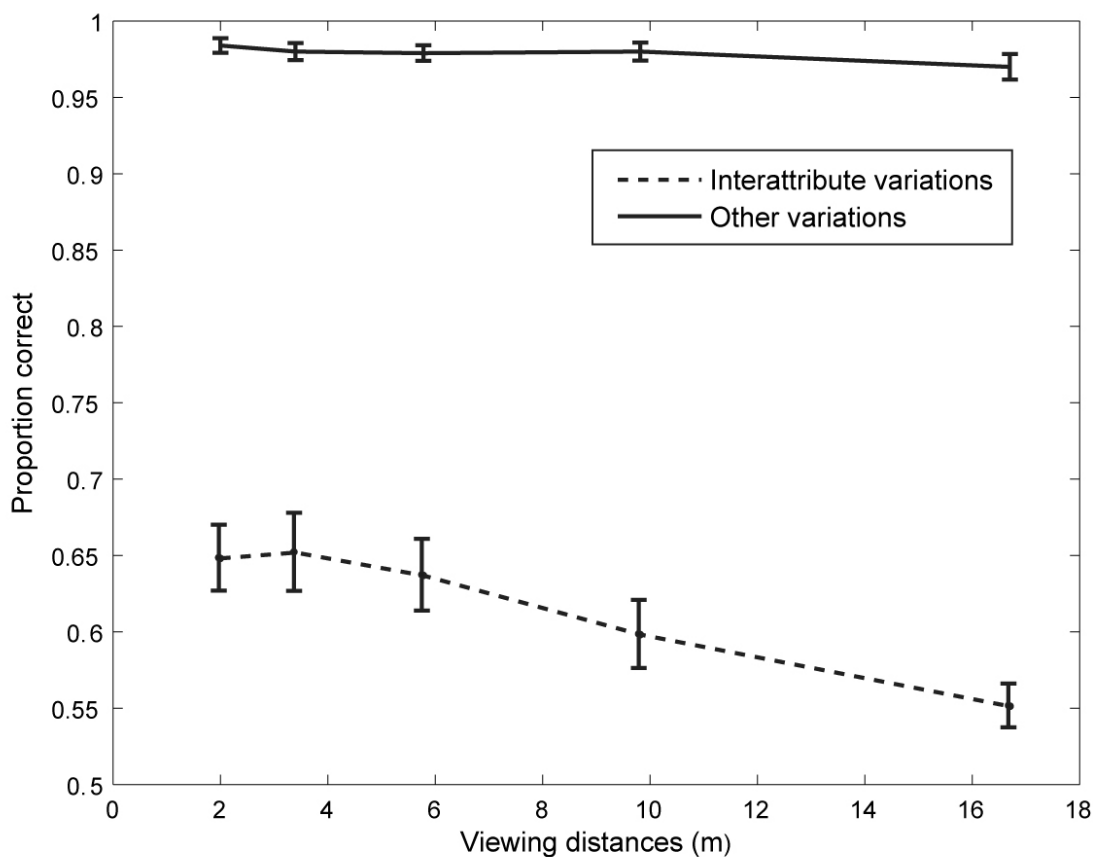


Figure 8. Mean proportion of correct face recognition in function of distance. Error bars represent one standard error. The dashed line represents performance when real-world interattribute distance is the only information available (Experiment 2a); and the solid line represents performance when only real-world cues other than interattribute distances such as attribute shapes and skin properties are available.

To test the effect of distance of presentation within each group, the data was separated and one-way ANOVAs were carried on each group independently. The within-subjects analysis revealed no differences of accuracy between any viewing points in the task where the interattribute distances were kept constant ($F(1.9,17.2)= 1.60, ns$). The same analysis revealed a significant difference between response accuracy as a function of distance in the group where the interattribute distances were different

($F(2.3,35)=10.51$, $p=.0001$, $\eta^2=.41$, $p_{rep}=.98$). A polynomial contrast revealed a significant linear relationship between response accuracy and distance of presentation when interattribute distances is the sole information available to perform the discrimination ($F(1, 15)=24.41$, $p=.0001$, $\eta^2=.62$, $p_{rep}=.98$). The group averages in this task indicated a decreasing accuracy with increasing distances (*nearest*: mean= 64.74%, std=8.6; *furthest*: mean=55.1%, std=5.77). Figure 5 displays the mean proportions and standard errors of correct responses as a function of viewing distances.

A 2 x (5) mixed ANOVA with *viewing distances* (2, 3.4, 5.78, 9.82 and 16.7 m) as a within-subjects factor and *groups* (different vs. same interattribute distances) as a between-subjects factor revealed a main effect of groups on response time ($F(1,24)=23.81$, $p <.00001$, $\eta^2=.50$, $p_{rep}=.99$). Same interattribute distances (mean=0.995 s, std=.22) elicited significantly faster reaction times than different interattribute distances (mean=1.94 s, std=.73).

General Discussion

In Experiment 1, we asked whether relative distances between real-world internal facial features contain enough information for face categorizations (identity and gender). We carried out a series of simulations on these faces to assess the information available in their residual interattribute distances. We found that real-world interattribute distances did in fact contain information useful to resolve face identification.

In Experiment 2a, we examined whether human observers could use real-world interattribute distance information to resolve a matching-to-sample (ABX) task when

this is the only information available. In Experiment 2b—the exact reciprocal of Experiment 2a—we asked if human observers could use information other than interattribute distances, such as attribute shapes and skin reflectance properties, to resolve an ABX task. Results of the Experiment 2a indicated that human observers perform poorly when required to recognize faces solely on the basis of real-world interattribute distances (best accuracy = 65% correct); whereas results of Experiment 2b showed that they perform close to perfection when required to recognize faces on the basis of real-world information other than interattribute distances such as attribute shapes and skin properties (e.g. O’Toole, Vetter & Blanz, 1999) at all tested viewing distances (equivalent to 2 to more than 16 m in the real-world).

One reason that may explain why the majority of researchers overestimated the importance of interattribute distances for face recognition is the use of grotesque face stimuli. We have computed that on average face stimuli in the interattribute distance literature convey 400% more information for identification than real-world faces. In an attempt to create realistic face stimuli, a subset of these researchers (e.g. Le Grand et al., 2001; Mondloch et al., 2002, 2003; Hayden, Bhatt, Reed, Corbly & Joseph, 2007) altered the interocular or the nose-to-mouth distance within a reasonable number of standard deviations from the mean according to the anthropometric norms of Farkas (1981). However, this effort is insufficient because the Farkas statistics are contaminated by an undesirable source of variance—the absolute size of faces—and they do not contain covariance information—e.g., the fact that the eyebrows have a tendency to follow the eyes. Rather, in our experiments, we have sampled interattribute

distances from real-world distributions. Alternatively, the method presented in the Appendix of this article can be used.

The results of Experiment 2a are all the more remarkable that they provide an *upper-bound* on the usefulness of interattribute distances for real-world face recognition. Our ABX task, which requires the identification of one recently viewed face among two face stimuli, is much easier than real-life face identification, which requires typically the comparison of hundreds of memorized faces with one face stimulus. To be fair, however, interattribute distances could gain some importance for human observers with increasing familiarity (e.g. Ge, Luo, Nishimura & Lee, 2003). Furthermore, no noise was added to the interattribute distances of our stimuli; real-life interattribute distances are contaminated by several sources of noise—facial movements, foreshortening, shadows, and so on. Finally, the interattribute distance information of our stimuli slightly overestimated real-life interattribute distance information because of unavoidable annotation errors.

In conclusion, facial cues other than interattribute distances such as attribute shapes and skin properties are the dominant information of face recognition mechanisms in the real-world.

References

- Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology, 58*, 71-78.
- Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions: Biological Sciences, 302*, 1203-1219.
- Bruce, V., & Young, A. W. (1998). *In the eye of the beholder: The science of face perception*. Oxford, UK: Oxford University Press.
- Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. *Psychological Science, 11*, 419–433.
- Carey, S., 1992. Becoming a face expert. *Philosophical Transaction of the Royal Society of London, 335*, 95-103.
- Diamond, R., & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology, 23*, 1-22.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise, *Journal of Experimental Psychology: General, 115*(2), 107-117.
- Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision, 9*(2),10, 1-8.

- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.
- Farah M J, Tanaka J W, Drain H M, (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, 21, 628-634
- Farkas, L.G. (1981). *Anthropometry of the head and the face*. New York : Elsevier.
- Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *Journal of Cognitive Neurosciences*, 11, 349–370.
- Ge, L., Luo, J. , Nishimura, M., & Lee, K. (2003). Lasting impression of Chairman Mao: Hyper-fidelity of familiar face memory. *Perception*, 32, 601-614.
- Haig, N.D. (1984). The effect of feature displacement on face recognition. *Perception*, 13, 502-512.
- 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, 85–98.
- Hochberg, J., & Galper, R. (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 perception of well known faces, *Perception*, 17, 461-474.
- Ingvalson, E. M., & Wenger, M. J. (2005). A strong test of the dual-mode hypothesis. *Perception & Psychophysics*, 67(1), 14-35.
- Leder, H., & Carbon, C. C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology*, 97, 19–29.
- Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature* 410, 890 (Correction: *Nature* 412, 786).
- Maurer, D., Le Grand, R., & Mondloch, C.J. (2002). The many faces of configural processing. *Trends in cognitive sciences*, 6, 255-260.
- Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of experimental child psychology*, 86, 67-84.
- Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, 31(5), 553-566.
- Okada, K., von der Malsburg, C., & Akamatsu, S. (1999). A Pose-Invariant Face Recognition System using Linear PCMAP Model. *Proceedings of IEICE workshop of Human Information Processing*, (HIP99-48), pages 7-12, Okinawa, November 1999.

- O'Toole, A.J., Vetter, T., & Blanz, V. (1999). Two-dimensional reflectance and three-dimensional shape contributions to recognition of faces across viewpoint. *Vision Research*, 39, 3145-3155.
- Rhodes, G. (1988). Looking at faces – 1st order and 2nd order features as determinants of facial appearance. *Perception*, 17, 43–63.
- Rhodes, G., Hayward, W. G., & Winkler, C. (2007). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin Review*, 13, 499-505.
- Robbins, R., & McKone, E. (2006). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, 103, 34-79.
- Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, 128, 274-289.
- Rossion, B. & Gauthier I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, 1, 62-74.
- Rotshtein, P., Geng, J. J., Driver, J., & Dolan, R. J. (2007). Role of features and second-order spatial-relations in face discrimination, face recognition, and individual face skills: Behavioral and fMRI data, *Journal of Cognitive Neuroscience*, 19 (9), 1435-1452.

- Schyns, P. G., & Oliva, A. (1999). Dr Angry and Mr Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.
- Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 904–915.
- Shepherd, J., Davies, G., & Ellis, H. (1981). Studies of cue saliency. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 105–131). London: Academic Press.
- Tanaka, J. W., & Farah, M. J. (1991). Second-order relational properties and the inversion effect: Testing a theory of face perception. *Perception and Psychophysics*, *50*, 367–372.
- Tanaka, J. W., & Farah, M.J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology: Section A*, *46*, 225-245.
- Tanaka, J.W., & Sengco, J.A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, *25*, 583-59.
- Ullman, S. (1989) Aligning pictorial descriptions: an approach to object recognition. *Cognition*, *32*, 193–254
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.

Young, A. M., Hellawell, D. & Hay, D. C. (1987). Configural information in face perception. *Perception*, 10, 747-759.

Appendix: How to create realistic interattribute distances

The most straightforward method—the one we opted for in Experiment 2a—consists in sampling interattribute distances from a real-world distribution. This solution has the advantage of preserving all interattribute distance information; but it has the disadvantage of being clumsy. In this appendix, we sketch an alternative method, which is a good compromise.

We can simulate the variance and covariance of the interattribute distances of our aligned female face set with μ_{female} the mean of the xy-coordinates of those features:

−39.10
−39.67
−16.40
−16.50
32.96
91.56
−56.86
56.84
−50.23
50.37
0.32
−0.50

(y-coordinates of left eyebrow, right eyebrow, left eye, right eye, nose, and mouth, followed by the x-coordinates of the same; the upper left quadrant being negative for both x and y coordinates) and $\mathbf{K}_{\text{female}}$ their covariance matrix

5.67 2.42 -1.96 -2.15 -2.01 -2.00 -2.76 3.45 -0.85 2.39 -0.76 -1.19
 2.42 6.37 -2.96 -2.74 -1.78 -1.25 -3.68 2.85

5.67 2.42 -1.96 -2.15 -2.01 -2.00 -2.76 3.45 -0.85 2.39 -0.76 -1.19
 2.42 6.37 -2.96 -2.74 -1.78 -1.25 -3.68 2.85 -1.61 -0.14 1.38 1.30
 -1.96 -2.96 5.13 3.87 -2.76 -1.21 0.35 0.33 -0.77 1.23 -0.74 -0.53
 -2.15 -2.74 3.87 5.79 -3.16 -1.31 -0.40 -0.49 -1.33 1.48 0.39 0.24
 -2.01 -1.78 -2.76 -3.16 10.77 -1.03 2.87 -2.26 1.49 -1.93 -0.36 0.08
 -2.00 -1.25 -1.21 -1.31 -1.03 6.90 3.58 -4.04 3.05 -2.97 0.16 0.18
 -2.76 -3.68 0.35 -0.40 2.87 3.58 6.68 -5.11 -0.11 -0.45 -0.83 -0.25
 3.45 2.85 0.33 -0.49 -2.26 -4.04 -5.11 7.22 -0.84 -0.89 0.06 -0.27
 -0.85 -1.61 -0.77 -1.33 1.49 3.05 -0.11 -0.84 5.42 -2.79 -1.26 -0.25
 2.39 -0.14 1.23 1.48 -1.93 -2.97 -0.45 -0.89 -2.79 6.72 -1.74 -0.61
 -0.76 1.38 -0.74 0.39 -0.36 0.16 -0.83 0.06 -1.26 -1.74 3.89 -0.27
 -1.19 1.30 -0.53 0.24 0.08 0.18 -0.25 -0.27 -0.25 -0.61 -0.27 1.73

via the following transformation of a Gaussian noise vector \mathbf{w} :

$$\mathbf{x}_{female} = E_{female} \Lambda_{female}^{1/2} \mathbf{w} + \boldsymbol{\mu}_{female},$$

where E_{female} is the orthogonal matrix of eigenvectors of \mathbf{K}_{female} and where Λ_{female} is the diagonal matrix of eigenvalues of \mathbf{K}_{female} . Likewise, we can simulate the variance and

covariance of the interattribute distances of our aligned male face set with μ_{male} :

-37.90
-38.46
-17.83
-18.11
31.80
93.46
-56.25
56.31
-49.86
49.86
0.57
-0.62

and K_{male} :

9.38	5.89	-3.98	-5.03	-3.31	-2.74	-3.35	4.47	-2.56	4.29	-1.99	-1.06
5.89	10.63	-5.11	-4.98	-3.43	-2.67	-5.82	4.61	-3.14	1.40	1.16	1.69
-3.98	-5.11	7.29	5.53	-2.39	-1.37	0.46	0.54	-0.12	0.97	-1.03	-0.62
-5.03	-4.98	5.53	7.50	-2.21	-0.88	0.20	-0.47	-0.68	-0.46	1.14	0.41
-3.31	-3.43	-2.39	-2.21	11.40	-0.04	3.17	-3.17	2.98	-3.08	0.17	-0.30
-2.74	-2.67	-1.37	-0.88	-0.04	7.74	5.22	-5.90	3.45	-3.04	0.56	-0.07
-3.35	-5.82	0.46	0.20	3.17	5.22	8.79	-6.95	0.59	-0.72	-0.57	-0.82
4.47	4.61	0.54	-0.47	-3.17	-5.90	-6.95	8.63	-2.06	0.88	-1.05	0.44
-2.56	-3.14	-0.12	-0.68	2.98	3.45	0.59	-2.06	6.50	-3.47	-1.15	-0.30
4.29	1.40	0.97	-0.46	-3.08	-3.04	-0.72	0.88	-3.47	6.92	-2.48	-0.93
-1.99	1.16	-1.03	1.14	0.17	0.56	-0.57	-1.05	-1.15	-2.48	5.32	-0.18
-1.06	1.69	-0.62	0.41	-0.30	-0.07	-0.82	0.44	-0.30	-0.93	-0.18	1.97

A Matlab function (i.e., `create_feature_pts`) implementing this method is freely available at <http://mapageweb.umontreal.ca/gosselif/align%20tools%20OSX/>).

Supplementary material: Virtual experiments

In the first virtual experiment, we trained a minimum squared-error linear classifier at identifying one randomly selected target natural set of aligned attribute positions from 50% of the other sets of natural aligned attribute positions also randomly selected by matrix pseudo-inverse:

$$\mathbf{b} = (\mathbf{X}_1^T \mathbf{X}_1)^{-1} \mathbf{X}_1^T \mathbf{y}_1,$$

with \mathbf{b} , the regression coefficients of the classifier. \mathbf{X}_1 is the set of training vectors:

$$\begin{bmatrix} \mathbf{x}_{\text{target}} & 1 \\ \dots & 1 \\ \mathbf{x}_{\text{target}} & 1 \\ \mathbf{x}_{\text{distracter}1} & 1 \\ \dots & 1 \\ \mathbf{x}_{\text{distracter}257} & 1 \end{bmatrix} + \begin{bmatrix} N(0, \sigma) & \dots & N(0, \sigma) & 0 \\ \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & 0 \\ N(0, \sigma) & \dots & N(0, \sigma) & 0 \end{bmatrix};$$

with \mathbf{x}_i , a training vector corresponding to the set of aligned attribute coordinates of face i :

$$[x_{\text{left eyebrow}} \quad x_{\text{right eyebrow}} \quad x_{\text{left eye}} \quad x_{\text{right eye}} \quad x_{\text{nose}} \quad x_{\text{mouth}} \quad y_{\text{left eyebrow}} \quad y_{\text{right eyebrow}} \quad y_{\text{left eye}} \quad y_{\text{right eye}} \quad y_{\text{nose}} \quad y_{\text{mouth}}]$$

where x and y are x-axis and y-axis coordinates, respectively; and with $N(0, \sigma)$, a Gaussian random variable of mean 0 and of standard deviation σ . And \mathbf{y}_1 is the category membership of the training vectors:

$$\begin{bmatrix} 1 \\ \dots \\ 1 \\ -1 \\ \dots \\ -1 \end{bmatrix}.$$

We tested this model on the remaining sets of natural aligned attribute positions:

$$\mathbf{y}_2' = f(\mathbf{b}\mathbf{X}_2),$$

with f , a Heaviside step function—if the value exceeds c the function outputs 1 and otherwise it outputs -1. \mathbf{X}_2 is the set of test vectors:

$$\begin{bmatrix} \mathbf{x}_{\text{target}} & 1 \\ \dots & 1 \\ \mathbf{x}_{\text{target}} & 1 \\ \mathbf{x}_{\text{distracter}_{258}} & 1 \\ \dots & 1 \\ \mathbf{x}_{\text{distracter}_{515}} & 1 \end{bmatrix} + \begin{bmatrix} N(0, \sigma) & \dots & N(0, \sigma) & 0 \\ \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & 0 \\ N(0, \sigma) & \dots & N(0, \sigma) & 0 \end{bmatrix}.$$

We computed the hit ($y_2'=\text{target}$ and $y_2=\text{target}$) and false alarm ($y_2'=\text{target}$ and $y_2=\text{distracter}$) rates in function c to obtain a Response Operating Curve (ROC). Our measure of sensitivity was the area under the ROC curve (A'). We varied the quantity of noise (σ). For each level of noise we did 1,000 repetitions in order to obtain a stable estimate of A' . We interpolated the quantity of noise required for the classifier to perform with an $A'=0.7$.

Similarly, in the second virtual experiment, we trained a linear classifier at identifying one randomly selected set of *artificial* aligned attribute positions from 50% of the sets of natural aligned attribute positions also randomly selected, and tested the classifier on the remaining sets of natural aligned attribute positions. Again, we found the quantity of noise required for the classifier to perform with an $A'=0.7$.

The classifiers trained to identify the artificial faces required 4 times more noise ($\sigma^2=400$ for a mean interocular distance of 100 pixels) than the ones trained to identify

the real-world faces ($\sigma^2=100$ for a mean interocular distance of 100 pixels). This means that the interattribute distances of the artificial faces convey 400% more information for identification than in real-world faces.

DISCUSSION GÉNÉRALE

L'article central du mémoire fait principalement trois apports au domaine de la reconnaissance des visages. Premièrement, la description des distances inter-attributs en considérant leurs covariances représente, à notre connaissance, la première description détaillée de cette source d'information.

Deuxièmement, les résultats présentés dans ce manuscrit consistent la première évaluation directe de l'utilisation de distances inter-attributs réalistes pour la reconnaissance des visages. Nous avons montré que cette source d'information n'est ni suffisante, ni nécessaire aux observateurs humains pour la reconnaissance des visages.

Troisièmement, la technique présentée en appendice de l'article (voir aussi la fonction Matlab reproduite en appendice de ce mémoire) permettra la création de visages ayant des distances inter-attributs réalistes pouvant être utilisés par les études futures. L'utilisation de cette technique pourra certainement limiter les problèmes de « validité écologique » engendrés par des manipulations arbitraires des distances inter-attributs. Comme il a été rapporté dans l'expérience 1 de l'article, les stimuli créés par des manipulations intuitivement réalistes ou en utilisant des lignes directrices reposant sur des valeurs absolues communiquent en moyenne 37 fois plus d'information que les stimuli réels. Nous recommandons donc fortement l'utilisation de cette technique.

1. La notion de distances inter-attributs et l'hypothèse holistique

Les conclusions de l'article central du mémoire sont en contradiction directe avec l'hypothèse suivant laquelle la reconnaissance des visages reposerait

essentiellement sur les distances inter-attributs; hypothèse pourtant admise par la majorité des chercheurs dans le domaine. Comme, nous l'avons expliqué dans l'introduction du mémoire, il persiste une certaine confusion quant à la nature et à la définition des processus holistiques et configuraux. Il est donc difficile, à ce stade de la présentation, de déterminer dans quelle mesure l'intégrité de l'hypothèse holistique a été atteinte par nos expériences. Toutefois, en plus des résultats présentés dans l'introduction de ce mémoire, un certain nombre d'études récentes ont remis en question d'un point de vue comportementale cette hypothèse. Ces indices proviennent essentiellement de la réévaluation de l'effet d'inversion des visages, de l'effet composite et de l'effet partie-tout à l'aide des nouvelles techniques psychophysiques telles que la technique de classification des réponses (Ahumada & Lovell, 1971; Eckstein & Ahumada, 2002) ainsi que la technique Bubbles (Gosselin & Schyns, 2001). La section qui suit offre un portrait de ces démonstrations suivit d'une brève discussion des critiques qui leurs ont été adressées.

1.1 Discussion de l'effet d'inversion

1.1.1 La technique de classification

Afin d'évaluer l'information utilisée lors de l'observation de visages à l'endroit et à l'envers, Sekuler et collègues (2004) ont utilisé la technique de classification des réponses (Ahumada & Lovell, 1971; Eckstein & Ahumada, 2002). Brièvement, cette

technique suppose que certains processus traitent différents aspects d'un visage et que le résultat phénoménologique de ces processus (e.g. la reconnaissance d'un visage plutôt qu'un autre) peut être utilisé pour résoudre une certaine tâche par l'observateur (e.g. répondre à une tâche de discrimination). Les processus en question sont donc considérés comme étant des filtres pouvant être soit linéaires, quasi-linéaires ou non-linéaires. Selon ces postulats, il est donc possible dans une tâche de reconnaissance visuelle, d'ajouter du bruit aux images cibles et à l'aide des réponses du participant d'évaluer les régions du stimulus où la présence de bruit affecte significativement sa reconnaissance. Gold et collègues (2000) ont décrit cette technique comme révélant les « champs récepteurs comportementaux » d'une tâche visuelle. En utilisant cette technique, Sekuler et collègues (2004) ont observé que l'information véhiculée par la région des yeux et des sourcils était linéairement corrélée avec la discrimination des visages à l'endroit mais aussi avec celle des visages à l'envers. De plus, les auteurs ont pu démontrer que l'apport de processus non-linéaires était négligeable dans cette tâche. Ces observations sont en accord avec plusieurs études précédentes ayant démontré l'utilité de la région des yeux et des sourcils pour la reconnaissance des visages à l'endroit (voir la section : « Les attributs internes du visage humain »). Comme les corrélations étaient simplement plus importantes pour les visages à l'endroit que pour les visages à l'envers, ces résultats suggèrent que la différence entre l'information utilisée dans chaque condition serait uniquement de nature *quantitative* (voir aussi Riesenhuber, Jarudi, Gilad, & Sinha, 2004; Yovel & Kanwisher, 2004; Gaspar, Bennett, & Sekuler, 2008). Cette observation est par conséquent en contradiction directe avec l'interprétation holistique de l'effet

d'inversion puisque cette hypothèse propose que l'inversion d'un visage entraîne une altération *qualitative* du processus de reconnaissance.

1.1.2 La technique Bubbles

Bien que la conclusion de Sekuler et collègues (2004) ait été controversée dans la littérature (voir Rossion, 2008), d'autres démonstrations ont également appuyé cette position. Par exemple, Willenbockel et collègues (sous-presse) ont récemment utilisé une variante de la technique Bubbles (Gosselin & Schyns, 2001) pour déterminer les fréquences spatiales précisément utilisées pour la reconnaissance de visages présentés à l'endroit ou à l'envers. Cette technique a l'avantage de pouvoir déterminer plus précisément les fréquences spatiales utilisées durant une tâche que les autres techniques habituellement utilisées telle que les filtres passe-bas et les filtres passe-haut (e.g. Boutet, Collin & Faubert, 2003; Goffaux & Rossion, 2006) et la méthode de masquage de la bande critique (Solomon & Pelli, 1994 ; Gaspar, et al., 2009). Les résultats de cette étude ont indiqué que la reconnaissance des visages à l'endroit et à l'envers nécessite l'information véhiculée par les mêmes fréquences spatiales (Willenbockel et al., sous-presse ; voir aussi Gaspar et al., 2009). Les auteurs ont donc conclu, tout comme Sekuler et collègues (2004), qu'à la lumière de ces résultats, l'explication la plus parcimonieuse de l'effet d'inversion semble être l'hypothèse d'une altération *quantitative* du processus de reconnaissance.

1.1.3 Les distances verticales vs les distances horizontales

Un dernier aspect de l'effet d'inversion ayant été proposé comme étant indicateur du traitement holistique est la sélectivité de l'effet d'inversion des visages pour les manipulations verticales des distances inter-attributs. Il a effectivement été rapporté que l'effet d'inversion serait sélectif aux manipulations de distances verticales alors que les manipulations des distances horizontales n'entraîneraient aucun effet d'inversion (Goffaux & Rossion, 2007; voir aussi Yovel & Kanwisher, 2004). Les auteurs ont expliqué ce phénomène par le fait que l'inversion forcerait le système visuel à traiter un visage élément par élément plutôt que de façon intégrée (holistique). Cette interprétation repose sur l'idée que la manipulation des distances horizontales entraîne une altération de distances métriques plus petites (e.g. distance entre un œil et le bord du visage) que la manipulation des distances verticales (e.g. la distance entre un œil et le nez). Ces auteurs proposent que les modifications de petites distances peuvent mieux être perçues lors de l'inversion puisque la reconnaissance se fera attribut par attribut compliquant ainsi le traitement des distances nécessitant le traitement de plus d'un attribut.

À la lumière des résultats rapportés dans notre manuscrit, il est toutefois possible d'offrir une hypothèse alternative plus parcimonieuse. Nous avons démontré dans la première expérience qu'un modèle linéaire des moindres carrés pouvait apprendre à utiliser l'information véhiculée par les distances horizontales des yeux pour identifier des visages. La revisite des articles précédents a permis de révéler que les seules

distances testées dans ces études ont effectivement été les distances horizontales et verticales des yeux (Goffaux & Rossion, 2007; Yovel & Kanwisher, 2004). Il est envisageable que cette observation soit associée à une exagération des distances horizontales au-delà des variations réelles. Les auteurs ont d'ailleurs reconnu le manque de réalisme des modifications opérées et que ces dernières ont été effectuées de façon à obtenir une performance comparable à celle de la tâche de reconnaissance des attributs (Goffaux & Rossion, 2007; section stimuli).

La sélectivité à l'orientation n'est donc pas forcément un exemple d'altération *qualitative* du traitement des visages inversés tel que rapporté par les auteurs. Effectivement, observer qu'un ajout d'information irréaliste élimine l'effet d'inversion n'implique pas l'existence d'une altération du processus de reconnaissance pour tous les visages normaux inversés. Ces observations pourraient simplement indiquer qu'il est possible d'exagérer des distances jusqu'à ce que l'observateur adopte une stratégie visuelle atypique reposant sur cette exagération. Dans ce cas précis, l'exagération minimalement nécessaire est peut être moindre pour les distances horizontales puisqu'elles contiennent d'avantage d'information permettant l'identification des visages par un observateur idéal (voir article du mémoire). Cette hypothèse alternative mérite certainement une étude plus approfondie (voir la section « Les pistes d'études futures »).

1.2 Discussion de l'effet composite

L'hypothèse holistique/configurale a également été ébranlée par la récente démonstration d'une absence d'association entre la force de l'effet composite et la performance d'individus à une tâche d'identification de visages (Konar, Bennett, & Sekular, 2007, 2008 ; voir aussi, Michel, Rossion, Han, Chung, Caldara, 2006). Ces résultats sont difficiles à expliquer d'après le modèle holistique puisque ce paradigme suppose que l'effet composite des visages soit informatif sur la nature du processus de reconnaissance. L'effet composite est d'ailleurs fréquemment utilisé comme la meilleure façon d'étudier le processus holistique dans les diverses études d'imagerie (e.g. Jacques & Rossion, 2009; de Heering & Rossion, 2008). Or, une absence d'association entre la performance à une tâche d'identification et la force de l'effet composite suggère que le mécanisme impliqué dans cet effet n'est probablement pas relié à l'identification des visages. Cette conclusion remet donc en question certains fondements de l'hypothèse configural/holistique du traitement des visages ou du moins certains moyens utilisés pour l'étudier.

1.3 Critique de l'étude psychophysique du processus holistique

Il a été proposé par certains chercheurs que les techniques psychophysiques telles que la technique Bubbles et la méthode de classification entravent le traitement holistique en interférant avec la nature intégrative du processus. Cette critique est

essentiellement fondée sur la supposition que l'observation de l'information parcellaire utilisée dans ces diverses techniques recruterait des processus perceptifs distincts (i.e. le traitement local des attributs) de ceux opérant lors de la reconnaissance traditionnelle de visages (i.e. l'intégration des attributs internes en une représentation unique).

L'argumentaire de cette critique prend racine dans le débat ayant été suscité par les études de masquage et d'isolation des parties (voir la section « Les attributs internes du visage humain » de l'introduction générale). Des suites de ce débat, il était généralement accepté dans la littérature que les techniques reposant uniquement sur la reconnaissance des attributs internes ne pouvaient qu'aboutir à des conclusions en ce sens (Rossion, 2008). Il a donc été proposé que la nature intégrative du traitement holistique ne puisse également pas être étudiée par les méthodes psychophysiques. Il est toutefois important de remarquer qu'il est possible d'observer l'effet composite - l'effet généralement considéré comme le reflet le plus direct du processus intégratif holistique – dans des stimuli morcelés comme ceux utilisés lors de l'application de la technique Bubbles. La figure 9 illustre cette situation. Notons également que les techniques de classification de réponses—Bubbles et la « reverse corrélation »—reposent sur la régression multiple linéaire et doivent donc utiliser un régresseur variable. Il est possible d'imaginer une tâche où ce régresseur—la réponse d'intérêt—serait la présence ou l'absence de l'effet composite et où les stimuli mèneraient parfois à l'observation d'un effet composite (e.g., figure 4b) et parfois à l'absence de cet effet. En somme, les techniques de classification de réponses pourraient être utilisées pour étudier l'effet composite. De plus, il est essentiel de noter qu'il est souvent possible d'estimer l'apport de processus non-

linéaires aux tâches évaluées par les techniques psychophysiques et que de cette façon l'existence d'un tel processus devrait pouvoir être décelé de cette façon. Par suite, la position voulant que les techniques psychophysiques soient à priori insensibles aux processus holistiques est fautive dans la mesure où l'effet composite est un indice du traitement holistique.

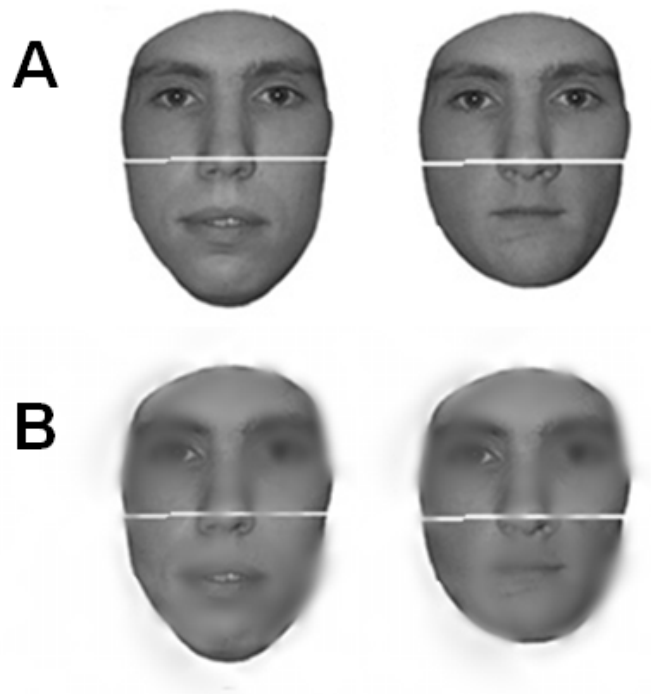


Figure 9. (A) Visages composites utilisés dans les études sur l'effet composite (B) Visages de la section A présentés avec la technique Bubbles (Visages originaux; Rossion, 2008).

2. Les limites de l'étude

L'un des facteurs limitant les conclusions de l'étude est l'impossibilité de proposer une alternative au traitement holistique/configural à la lumière des résultats ici

présentés. Il est effectivement possible d'affirmer que l'information contenue dans les distances ne semble pas jouer un rôle primordial dans la reconnaissance des visages sans toutefois pouvoir définir précisément l'information effectivement pertinente. Cette limitation est essentiellement liée au fait que la distance inter-attribut est le seul paramètre à avoir été explicitement manipulé. Par exemple, les stimuli de l'expérience 2b ont des attributs internes variables mais il demeure difficile de conclure à leur seule contribution à la reconnaissance des visages puisque leur variation unique n'a pas été volontairement contrôlée. Il est certainement possible d'avancer l'hypothèse de leur utilisation puisqu'il s'agit de la seule autre information véhiculant théoriquement suffisamment d'informations pour permettre une aussi grande individualisation des visages. Toutefois, la nature du processus exact ne pourra pas être davantage élucidé par les résultats ici présentés. Il est possible que certains attributs de la peau (e.g. grain de beauté) ou des différences de contraste au niveau de la peau puisse également influencer la reconnaissance.

Qui plus est, nous ne pouvons quantifier l'importance relative des distances inter-attributs et des autres sources d'information principalement pour deux raisons. Premièrement, les performances rapportées dans l'article sont soit près du plafond (Expérience 2b) ou soit près du plancher (Expérience 2a) ce qui limite la comparaison des deux conditions à chacune des distances de présentation. Deuxièmement, les tâches ABX des deux expériences sont potentiellement de difficultés différentes (i.e., les sujets sont-ils meilleurs en 2b parce que la tâche est plus facile ou parce qu'ils encodent ces informations plutôt que celles de 2a?). Il serait donc pertinent comme nous allons le

proposer dans la section suivante d'utiliser une mesure d'efficacité qui tient compte de la difficulté de la tâche (voir Gold, Bennett & Sekuler, 2004) et qui permet de comparer des tâches de difficulté différentes (e.g., ABX à partir des distances inter-attributs et ABX à partir des autres sources d'information).

3. Pistes pour des études futures

3.1 L'efficacité des distances inter-attributs et des autres sources d'information

Bien que les proportions de bonnes réponses et les temps de réponses soient informatifs, il serait intéressant d'obtenir une mesure d'efficacité pour déterminer si les observateurs humains sont intrinsèquement meilleurs dans une condition (i.e. utiliser les distances) que dans une autre (i.e. utiliser les attributs internes). Pour ce faire, il est possible d'imaginer une nouvelle expérience où il serait possible d'introduire du bruit gaussien soit dans l'information des distances ou dans celle des attributs de façon à évaluer la sensibilité des observateurs à l'aide du d' (d-prime). Le d' est une mesure reposant sur la distance entre les moyennes des distributions du signal et du bruit (« noise ») en unité d'écart-type. En psychophysique, la mesure est fréquemment calculée à partir des vrais positifs (« Hit ») et des faux positifs (« False alarm ») dans une tâche quelconque ($d' = Z(\text{taux de vrais positifs}) - Z(\text{taux de faux positifs})$). Cette mesure reflète la spécificité de la réponse puisqu'elle indique à la fois l'aptitude des observateurs à déterminer la bonne réponse mais aussi leur aptitude à ne pas détecter les

fausses réponses comme étant vraie. Par exemple, dans notre situation, un premier visage cible pourrait être présenté suivi d'un visage ayant subi une altération des distances inter-attributs selon une courbe gaussienne représentative des distances réelles (voir l'expérience 1 de l'article). Le bruit gaussien ici ajouté nous permettrait d'évaluer la sensibilité des observateurs à utiliser les variations de distances inter-attributs. En utilisant ce même procédé pour les attributs internes (introduction de bruit gaussien en faisant varier le contraste des pixels), il serait possible d'évaluer la sensibilité des observateurs dans chacune des tâches. L'efficacité (F) des observateurs pourrait alors être calculée dans chacune des conditions (distances ou contrastes) par le d' des observateurs standardisé à l'aide de la difficulté de la tâche telle que déterminée par le d' de l'observateur idéal (e.g. $F_{\text{distances}} = (d'_{\text{humain}}/d'_{\text{observateur idéal}})^2$).

L'observateur idéal (voir expérience 1) est un modèle informatique disposant de la totalité de l'information véhiculée par les images pour résoudre la tâche. En bruitant un paramètre de cette tâche (e.g. les distances inter-attributs) il est ainsi possible d'obtenir une mesure de la performance « idéale » dans cette condition en fonction du bruit introduit. L'observateur idéal est ici utilisé comme mesure objective de la difficulté de la tâche puisque cette indice est ni affecté par le bruit interne, ni par le bruit perceptif inhérent à la performance de l'observateur humain. Cette standardisation permettrait donc de comparer adéquatement les deux conditions (i.e. variations des distances ou variation du contraste) tout en contrôlant pour la difficulté de chacune des tâches.

3.2 L'effet d'inversion et le réalisme des distances inter-attributs

Il a été suggéré précédemment que l'effet d'inversion observé préférentiellement pour les distances inter-attributs ne soit que le reflet de l'inaptitude du système visuel à utiliser ces distances lorsque les visages sont présentés à l'endroit. Il serait possible de tester cette hypothèse à l'aide de variations horizontales ou verticales de la position des yeux (voir la section «Les distances verticales vs les distances horizontales»). Le modèle linéaire des moindres carrés utilisé dans l'expérience 1 a effectivement indiqué que l'information des distances horizontales entre les yeux serait la distance inter-attribut la plus utile pour reconnaître un visage. Il a également été suggéré par des études précédentes que les variations de distances horizontales mèneraient à une identification correcte des visages inversés c'est-à-dire à l'élimination de l'effet d'inversion (Goffaux & Rossion, 2007; Yovel & Kanwisher, 2004). Afin d'expliquer ces résultats, nous posons l'hypothèse que l'absence d'effet d'inversion observée lors de la manipulation des distances horizontales soit engendrée par l'utilisation d'information interoculaire irréaliste étant potentiellement très diagnostique pour l'identification des visages. Pour mettre cette hypothèse à l'épreuve plusieurs manipulations pourraient être effectuées.

Premièrement, il serait possible d'annoter les visages des études précédentes (Goffaux & Rossion, 2007; Yovel & Kanwisher, 2004), de les aligner selon la technique présentée dans la première section de l'article et de tester l'efficacité du modèle linéaire des moindres carrés à identifier chacune de ces modifications (verticale et horizontale). Nous posons l'hypothèse que les stimuli ayant encouru des manipulations des distances horizontales seront plus facilement identifiés que ceux ayant encouru des manipulations

des distances verticales. Il pourrait également être intéressant d'évaluer le réalisme de ces distances en comparant ces stimuli avec la distribution des distances inter-attributs réelles décrite à l'expérience 1.

Deuxièmement, il serait possible d'imaginer une condition expérimentale où les distances verticales des yeux seraient si exagérées qu'elles représenteraient un outil discriminatif aussi intéressant que les distances horizontales pour un modèle linéaire des moindres carrés. Ceci étant rendu équivalent, il serait possible de faire une tâche d'inversion similaire à celle de Goffaux et Rossion (2007) mais en ayant normalisé la quantité d'information pertinente à la reconnaissance dans chacune des conditions (distances horizontales ou distances verticales). Si notre hypothèse est vraie, aucun effet d'inversion ne devrait être observé dans l'une ou l'autre des conditions.

Troisièmement, si l'hypothèse voulant que le traitement des distances verticales soit affecté par l'inversion est vraie, alors cet effet devrait pouvoir être observé en utilisant des stimuli respectant les distances inter-attributs du monde réel. Si cet effet est effectivement sélectif à l'orientation des manipulations alors il sera encore possible d'observer une différence entre les conditions lorsque les distances horizontales ou verticales subiront des modifications réelles. Par contre, si cet effet est essentiellement engendré par l'irréalisme des distances utilisées et par la nature diagnostique de l'information horizontale alors l'utilisation de distances réelles devrait faire disparaître l'effet observé pour les distances horizontales. Cette hypothèse repose sur le fait que les distances interoculaires réelles semblent être peu utilisées par les observateurs humains pour accomplir une tâche de reconnaissance des visages à l'endroit (voir expérience 2).

Références

- Ahumada, A., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, *49*, 1751-1756.
- Allison, T., Puce, A., Spencer, D. D. & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and nonface stimuli. *Cerebral Cortex*, *9*, 415–430.
- Andrews, T. J., Schiuppeck, D., Homfray, D., Matthews, P., & Blakemore, C. (2002). Activity in the fusiform gyrus predicts conscious perception of Rubin's vase-face illusion. *Neuroimage*, *17*, 890-901.
- Barton, J. J., 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.
- Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, *58*, 71–78.
- Bentin, S., Allison, T., Puce, A., Perez, E. & McCarthy, G. (1996) Electrophysiological studies of face perceptions in humans. *Journal of cognitive Neurosciences*. *8*, 551–565.
- Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics*, *65*, 1078-1093.
- Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural

- and configural processing to face recognition. *Psychological Science*, *11*, 419–433.
- Cattell, James McKeen. (1886a). The time taken up by cerebral operations, Parts 1 & 2. *Mind*, *11*, 220-242.
- Collishaw, S. M., & Hole, J. C. (2000). Featural and configural processes in the recognition of faces of different familiarity. *Perception*, *29*, 893-909.
- Davies, G., Ellis, H., & Shepherd, J. (1977). Cue saliency in faces as assessed by the “photofit” technique. *Perception*, *6*, 263–269.
- Diamond, R., & Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology*, *23*, 1-22.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology General*, *115*, 107–117.
- Davidoff, J., & Donnelly, N. (1990). Object superiority : A comparison of complete and part probes. *Acta Psychologica*, *73*(3), 225-243.
- de Heering, A. & Rossion, B. (2008). Prolonged visual experience in adulthood modulates holistic face perception. *PLOS One*, *3*(5), e2317.
- Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision*, *2*(1), i.
- Endo, M., Masame, K., & Maruyama, K. (1989). Interference from configural of a schematic face onto the recognition of its constituent parts. *Tohoku Psychologica Folia*, *48*, 97-106.
- Farah, M. I., Tanaka, J. N., & Drain, M. (1995). What causes the face inversion effect.

- Journal of Experimental Psychology: Human Perception and Performance*, 21, 628-634.
- Farkas, L.G. (1981). *Anthropometry of the head and the face*. New York : Elsevier.
- 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, 159-170.
- Gaspar, C. M., Bennett, P..J., Sekuler, A. B. (2008). The effects of face inversion and contrast-reversal on efficiency and internal noise. *Vision Research*, 48, 1084-1095.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neurosciences*, 3, 191-197.
- Goffaux, V. (2008). The horizontal and vertical relations in upright faces are transmitted by different spatial frequency ranges, *Acta Psychologica*, 128, 119-126.
- Goffaux, V., & Rossion, B. (2006). Faces are “spatial” – Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1023–1039.
- Goffaux, V., & Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 995–1002.
- Gold, J. M., Murray, R. F., & Bennett, P. J., & Sekular, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology*, 10, 663-666.

- 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. (2005). Bubbles: A user's guide. In L. Gershkoff-Stowe & D.H. Rakison (Eds.) *Building Object Categories in Developmental Time* (pp. 91-106). Hillsdale, NJ: Erlbaum.
- Haig, N.D. (1984). The effect of feature displacement on face recognition. *Perception*, *13*, 502-512.
- Haig, N.D. (1985). How faces differ – A new comparative technique. *Perception*, *14*, 601-615.
- Haig, N. D. (1986a). High-resolution facial feature saliency mapping. *Perception*, *15*, 373-386.
- Haig, N.D. (1986b). Exploring recognition with interchanged facial features, *Perception*, *15*, 235-247.
- 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*, 85–98.
- Hasson, U., Hendler, T., Ben Bashat, D. & Malach, R. (2001). Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *Journal of Cognitive Neurosciences*, *13*, 744-753.
- Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, *23*, 65–74.

- Jacques, C., & Rossion, B. (2009). The initial representation of individual faces in the right occipito-temporal cortex is holistic: electrophysiological evidence from the composite face illusion. *Journal of Vision, 9*(6), 1–16.
- Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neurosciences, 3*, 759-763.
- Kanwisher, N., McDermott, J. & Chun, M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neurosciences, 17*, 4302–4311.
- Konar, Y., Bennett, P. J., & Sekuler, A. B. (2007). The composite face effect is not correlated with face identification accuracy. *Journal of Vision, 7*(9), 501.
- Konar, Y., Bennett, P. J., & Sekuler, A. B. (2008). The composite face effect is still not correlated with face identification accuracy. *Journal of Vision, 8*(6), 891.
- Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *The Quarterly Journal of Experimental Psychology: Section A, 51*, 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*, 513–536.
- Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30*, 73–83.
- Leder, H., & Carbon, C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology, 97*, 19–29.
- Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual

- experience and face processing. *Nature*, *410*, 890.
- Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science*, *15*, 762–768.
- Liu, J., Higuchi, M., Marantz, A. & Kanwisher, N. (2000) The selectivity of the occipitotemporal M170 for faces. *Neuroreport*, *11*, 337–341.
- Maurer, D., Le Grand, R., & Mondloch, C.J. (2002). The many faces of configural processing. *Trends in cognitive sciences*, *6*, 255-260.
- McCarthy, G., Puce, A., Belger, A. & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of facespecific potentials generated in occipitotemporal cortex. *Cerebral Cortex*, *9*, 431–444.
- McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces?. *Trends In Cognitive Sciences*, *11(1)*, 8-15.
- Michel, C., Rossion, B., Han, J., Chung, C-S., Caldara, R. (2006). Holistic processing is finely tuned for faces of our own race. *Psychological Science*, *17*, 608-615.
- Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology*, *86*, 67–84.
- Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*, 553–566.
- Moscovitch, M., Winocur, G. & Behrmann, M. (1997). What is special about face recognition? nineteen experiments on a person with visual object agnosia and

- dyslexia but normal face recognition. *Journal of Cognitive Neurosciences*, 9, 555–604.
- Puce, A., Allison, T. & McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex* 9, 445–458.
- Reicher, G.M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 275-280.
- 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, 25–57.
- Rhodes, G., Hayward, W. G., & Winkler, C. (2007). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin Review*, 13, 499–505.
- Riesenhuber, M., Jarudi, I., Gilad, S., & Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings in Biological Sciences*, 271 Suppl 6, S448–S450.
- 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, 274-289.
- Rossion, B., & Gauthier, I. (2002). How Does the Brain Process Upright and Inverted

Faces? *Behavioral and Cognitive Neuroscience Reviews*, 1, 62–74.

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13, 402-209.

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 904–915.

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 process underlying face recognition. *British Journal of Psychology*, 75, 221-242.

Shepherd, J., Davies, G., & Ellis, H. (1981). Studies of cue saliency. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 105–131). London: Academic Press.

Solomon, J. A. & Pelli, D. G. (1994). The visual channel that mediates letter identification. *Nature*, 369, 395-397.

Tanaka, J. W., & Farah, M.J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology: Section A*, 46, 225-245.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, 25, 583–592.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B., & Livingstone, M. S. (2006). A cortical

- region consisting entirely of face-selective cells. *Science*, *311*, 670.
- Tsao, D. Y., & Levingston, M. S. (2008). Mechanisms of face perception. *Annual Review in Neurosciences*, *31*, 411-437.
- Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception*, *6*, 323-326.
- Willenbockel, V., Fiset, D., Chauvin, A., Blais, C. Arguin, M., Tanaka, J., Bub, D. & Gosselin, F. (sous-presses). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance*.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.
- Young, A. M., Hellawell, D., & Hay, D. C. (1987). Configural information in face perception. *Perception*, *10*, 747–759.
- Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, *44*, 889–898.

Appendice

```

function [ptsCoor, average] = create_feature_pts(eye_distance, gender)

% CREATE_FEATURE_PTS produces random facial interattribute positions
% (ptsCoor) which have the same the 1st and 2nd moment properties than
% a set of 256 Caucasian full-frontal neutral female faces (when the
% gender argument is set to 'female') and a set of 259 Caucasian full-
% frontal male faces (when the gender argument is set to 'male'). The
% argument eye_distance sets the average interocular distance to the
% specified value (in pixels) for the specified gender. The outputs of
% the function have three columns which correspond, from left to right,
% to the x and y coordinates of the features' center of gravity and the
% constant 1 (i.e., ready for affine transformation); and six rows
% which correspond, from top to bottom, to the left and right eyebrows,
% the left and right eyes, the nose and the mouth.
%
%      E.g., [ptsCoor, average] = create_feature_pts(100, 'female');
%
% If you use this function in your research, please, cite us:
%      Taschereau-Dumouchel, V., Rossion, B., Schyns, P. G. &
%      Gosselin, F. (submitted).
%      Interattribute distances do not represent the identity of real-
%      world faces.
%
%      Frederic Gosselin, 20/07/2009

if strcmp(gender, 'female')

K = [
    5.6728    2.4208   -1.9567   -2.1514   -2.0117   -2.0001   -2.7591
    3.4470   -0.8536    2.3932   -0.7643   -1.1941;
    2.4208    6.3717   -2.9613   -2.7379   -1.7752   -1.2489   -3.6848
    2.8534   -1.6068   -0.1394    1.3782    1.3000;
    -1.9567   -2.9613    5.1269    3.8729   -2.7646   -1.2119    0.3488
    0.3320   -0.7691    1.2269   -0.7370   -0.5294;
    -2.1514   -2.7379    3.8729    5.7882   -3.1649   -1.3135   -0.4014
    -0.4928   -1.3342    1.4818    0.3875    0.2373;
    -2.0117   -1.7752   -2.7646   -3.1649   10.7671   -1.0330    2.8720
    -2.2573    1.4931   -1.9305   -0.3638    0.0824;
    -2.0001   -1.2489   -1.2119   -1.3135   -1.0330    6.8985    3.5782
    -4.0396    3.0525   -2.9650    0.1594    0.1824;
    -2.7591   -3.6848    0.3488   -0.4014    2.8720    3.5782    6.6832
    -5.1114   -0.1101   -0.4549   -0.8288   -0.2451;
    3.4470    2.8534   -0.3320    0.3320   -0.4928   -2.2573   -4.0396   -5.1114
    7.2174   -0.8403   -0.8854    0.0639   -0.2686;
    -0.8536   -1.6068   -0.7691   -1.3342    1.4931    3.0525   -0.1101
    -0.8403    5.4230   -2.7940   -1.2557   -0.2471;
    2.3932   -0.1394    1.2269    1.4818   -1.9305   -2.9650   -0.4549
    -0.8854   -2.7940    6.7206   -1.7404   -0.6059;
    -0.7643    1.3782   -0.7370    0.3875   -0.3638    0.1594   -0.8288
    0.0639   -1.2557   -1.7404    3.8905   -0.2686;

```

```

    -1.1941    1.3000   -0.5294    0.2373    0.0824    0.1824   -0.2451
-0.2686   -0.2471   -0.6059   -0.2686    1.7255;];

```

```

mu = [
-39.0977
-39.6680
-16.3984
-16.5039
 32.9609
 91.5586
-56.8633
 56.8398
-50.2266
 50.3711
  0.3242
-0.5000];

```

```

elseif strcmp(gender, 'male')

```

```

K = [
    9.3821    5.8934   -3.9812   -5.0283   -3.3100   -2.7383   -3.3478
4.4743   -2.5557    4.2929   -1.9856   -1.0599;
    5.8934   10.6292   -5.1095   -4.9762   -3.4337   -2.6718   -5.8212
4.6136   -3.1431    1.4031    1.1628    1.6862;
   -3.9812   -5.1095    7.2863    5.5258   -2.3929   -1.3673    0.4643
0.5383   -0.1233    0.9695   -1.0332   -0.6173;
   -5.0283   -4.9762    5.5258    7.4999   -2.2117   -0.8765    0.1976
-0.4746   -0.6791   -0.4574    1.1395    0.4054;
   -3.3100   -3.4337   -2.3929   -2.2117   11.4014   -0.0353    3.1703
-3.1672    2.9768   -3.0831    0.1694   -0.3036;
   -2.7383   -2.6718   -1.3673   -0.8765   -0.0353    7.7377    5.2204
-5.9043    3.4493   -3.0426    0.5581   -0.0737;
   -3.3478   -5.8212    0.4643    0.1976    3.1703    5.2204    8.7934
-6.9542    0.5922   -0.7220   -0.5692   -0.8155;
    4.4743    4.6136    0.5383   -0.4746   -3.1672   -5.9043   -6.9542
8.6314   -2.0646    0.8771   -1.0548    0.4384;
   -2.5557   -3.1431   -0.1233   -0.6791    2.9768    3.4493    0.5922
-2.0646    6.4972   -3.4729   -1.1473   -0.2955;
    4.2929    1.4031    0.9695   -0.4574   -3.0831   -3.0426   -0.7220
0.8771   -3.4729    6.9214   -2.4762   -0.9341;
   -1.9856    1.1628   -1.0332    1.1395    0.1694    0.5581   -0.5692
-1.0548   -1.1473   -2.4762    5.3234   -0.1783;
   -1.0599    1.6862   -0.6173    0.4054   -0.3036   -0.0737   -0.8155
0.4384   -0.2955   -0.9341   -0.1783    1.9726;];

```

```

mu = [
-37.9035
-38.4595
-17.8340
-18.1081
 31.7992
 93.4595

```

```
-56.2510
 56.3050
-49.8649
 49.8571
  0.5714
 -0.6216
];

else
    error('Second argument must be ''female'' or ''male''.')
end

[E, lambda] = eig(K);

scaling = eye_distance / (mu(10) - mu(9));
temp_pts = (E * sqrt(lambda) * randn(12,1) + mu);
ptsCoor = ones(6,3);
ptsCoor(:,1:2) = [temp_pts(1:6)'; temp_pts(7:12)'];]*scaling;
average = ones(6,3);
average(:,1:2) = [mu(1:6)'; mu(7:12)'];]*scaling
```

