



**HAL**  
open science

# Modélisation de la beauté et étude de son influence sur les comportements sociaux d'un primate non-humain

Sonia Tiew

► **To cite this version:**

Sonia Tiew. Modélisation de la beauté et étude de son influence sur les comportements sociaux d'un primate non-humain. Biodiversité et Ecologie. Université de Montpellier, 2023. Français. NNT : 2023UMONG124 . tel-04865977

**HAL Id: tel-04865977**

**<https://theses.hal.science/tel-04865977v1>**

Submitted on 6 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Ecologie et Biodiversité

École doctorale GAIA

Unité de recherche CEFE UMR 5175

Modélisation de la beauté et étude de son  
influence sur les comportements sociaux d'un  
primate non-humain

Présentée par Sonia TIEO

Le 01/12/2023

Sous la direction de Marie CHARPENTIER  
et Julien RENOULT

Devant le jury composé de

Marie CHARPENTIER, DR, CNRS, ISEM

Julien RENOULT, CR CNRS, CEFE

Blandine DOLIGEZ, DR, CNRS, LBBE

Olivier PASCALIS, DR, CNRS, LPNC

Elise HUCHARD, DR, CNRS, ISEM

Gwenaël KAMINSKI, MCF, Université de Toulouse, CLLE

Directrice de thèse

Directeur de thèse

Rapportrice

Rapporteur

Examinatrice

Examineur



UNIVERSITÉ  
DE MONTPELLIER



*“This is definitively the darkest timeline.”*

*Abed Nadir, Community*



# Table des matières

<b>INTRODUCTION GENERALE</b>	<b>5</b>
I. Cadre philosophique de la beauté .....	14
II. Apport des sciences cognitives .....	18
A. Origines et concept de la théorie de la fluence.....	18
B. Un cas particulier de la fluence : la prototypicalité.....	22
III. L'angle évolutionniste : origines et conséquences de la beauté .....	24
A. Théories évolutionnistes de la beauté .....	24
B. La fluence: un exemple de biais perceptif.....	27
C. Conséquences sociales et sexuelles de la beauté.....	28
IV. Modéliser la Fluence .....	30
A. Approches classiques pour quantifier la fluence.....	30
B. Modéliser la fluence : approche générale.....	32
1. CNNs comme modèle de perception .....	33
2. Métriques de la fluence : l'efficacité et l'efficacité .....	43
V. Objectifs de la thèse.....	45
<b>CHAPITRE 1</b>	<b>39</b>
<b>CHAPITRE 2</b>	<b>59</b>
<b>CHAPITRE 3</b>	<b>111</b>
<b>CHAPITRE 4</b>	<b>134</b>
<b>DISCUSSION</b>	<b>159</b>
I. La fluence comme déterminant de la beauté : Modélisation à travers la prototypicalité et prédiction de l'attractivité .....	170
A. Prototypicalité et Féminité.....	170
1. Définitions et mesures de la féminité.....	170
2. Corrélation entre féminité et attractivité .....	173
3. Sélection de caractéristiques faciales VS approche « automatisée » basée sur les CNNs.....	174
4. Féminité : disparités intra-groupes.....	176
B. De la Féminité comme Prototype à la Typicalité Statistique : Vers une Perspective Élargie.....	178
1. Typicalité statistique à tous les niveaux de perception .....	178
2. Modulation du prototype : adaptabilité basée sur l'expérience .....	180
3. Retour à la fluence : interprétation sous les notions d'efficacy et d'efficiency.....	180
II. Féminité : Conséquences socio-sexuelles et Origine Biologique et Evolutive .....	182

A. Impact de la féminité sur les comportements des mandrills.....	182
B. Préférences des Mandrills pour les Femelles Moins Féminines : Origine évolutive ...	187
C. Origine de la variation de la féminité des visages chez les mandrills juvéniles.....	190



---

# REMERCIEMENTS

---

Comment mettre les mots sur toute ma gratitude sans que cela sonne comme une carte de remerciement pré-écrite ?

Je vais commencer par les personnes qui m'ont embarquée dans cette aventure (pas si tranquille) qu'est la thèse. Julien et Marie, je veux vous remercier à la fois pour votre rigueur et votre exigence, qui m'ont incitée à aller toujours plus loin et à assurer un travail de qualité, et également pour votre écoute, patience et empathie, qui ont su rendre mon parcours de thèse bien plus agréable. Julien, tu m'as offert cette opportunité et j'en suis sincèrement reconnaissante. Dès le début, tu m'as impressionnée par ta capacité à jongler avec mille idées différentes et ta curiosité pour tant de disciplines différentes. Tu fais des connexions à la fois déroutantes et fascinantes. Tout ça m'a constamment motivée et, au fond, poussée à donner le meilleur de moi-même pour être 'à la hauteur' ! J'ai énormément appris à tes côtés, dans un domaine qui m'était totalement étranger au début. Marie, tu m'as montré qu'en tant que femme dans la recherche, on peut être forte, brillante, avoir une belle carrière sans sacrifier la qualité des relations humaines. Tu es toujours disponible et tu crées une atmosphère où l'on sait que l'on peut parler de tout avec toi. Merci !

Je tiens à exprimer ma gratitude envers tous ceux qui ont contribué, de près ou de loin, à ce projet. Une thèse n'est jamais le travail d'une seule personne ; il s'agit d'un ensemble de données collectées, triées et analysées, mais également d'un partage d'informations et de savoirs. Cela inclut le travail des chercheurs, des ingénieurs, des stagiaires et des assistants de terrain. Un merci pour le travail des stagiaires Melvin et Roland qui ont largement contribué aux résultats. Je remercie François Rousset pour son aide en statistiques. En me rendant sur le terrain Mandrills, j'ai pris conscience de la complexité de la collecte de données, donc Merci aux assistants de terrain qui ont recueilli des données ; pour ceux que j'ai eu la chance de rencontrer (Anna, Marianne, Jeff, Pascal) et pour ceux que je n'ai pas eu le plaisir de croiser, sachez que votre travail est précieux. Un gros merci à Jules et Berta: vous m'avez supporté sur le terrain avec tant de bienveillance et de gentillesse, malgré mes deux pieds gauches. C'est toujours un plaisir de vous retrouver loin de la forêt et de ses alentours (sans avoir à jeter un œil



par-dessus notre épaule en craignant une charge de buffle surprise) pour trinquer avec quelque chose de nettement plus agréable que la Regal !

Je tiens à remercier mes bougs, Erwan et Nico, mes petits co-bureaux. J'ai eu trop de chance d'avoir commencé ma thèse avec vous ! Entre nos discussions (qui, évidemment, ont toujours été scientifiques hors des pauses café, n'est-ce pas ?), nos rires qui résonnaient dans tout le couloir, et Erwan avec ses interrogations existentielles à n'en plus finir... Merci les gars d'avoir été là, pour râler mais aussi pour booster le moral quand les choses se compliquent ! Vous avez rendu la thèse amusante ! Un merci spécial à Nico, ce joyeux luron toujours accueillant et prêt à tendre la main (surtout pour les trucs administratifs). C'est formidable d'avoir un ami comme toi sur qui on se repose et compter. Merci, même si ma manière d'exprimer ma gratitude n'est pas aussi démonstrative que la tienne (tmtc).

Merci aussi à toutes les personnes de l'équipe du CEFÉ et de l'ISEM (et même ceux qui ne sont pas dans l'équipe, d'ailleurs) ! Fiou, en 3 ans, il y a eu tellement de visages qui ont défilé et je suis certaine d'en oublier quelques-uns (alors un merci général s'adresse également à ceux que je ne mentionnerais pas). Et un gros merci pour toutes ces réunions, où j'ai pu plonger dans vos projets ultra-passionnants ! Non seulement j'ai engrangé une tonne de connaissances scientifiques, mais en plus, j'ai adoré nos pauses café et déj' ensemble (et aussi les soirées au bois de Montmaur, aux bars ou à la plage): David, Rita, Pietro, Amelie, Aude, Lisa, Nikos, Constance, Tracy, Megan, Laura ... Un merci spécifique à Axelle pour avoir été un pilier moral (t'es un peu une sorte de grande sœur modèle halala), et pour tous ces doux moments partagés ensemble. (Et un merci à Romain aussi, hein ? On ne voudrait pas que bichon soit vexé, il fait partie intégrante de ces beaux moments également !)

Je remercie les membres de mon jury, Elise Huchard, Olivier Pascalis, Blandine Doligez, et Gwenaël Kaminski, pour avoir accepté d'évaluer mon travail. Je remercie également les membres de mon comité de thèse, Anne Charmantier, Pierre-Olivier Antoine, et Rufin Van Rullen, pour leurs précieux conseils et leurs questions pertinentes tout au long de ce parcours.

Un gros merci aux potes de Montpellier. Quand je suis arrivée ici, je ne connaissais quasi personne, et franchement, chaque personne qui a partagé des moments sympas avec moi mérite un merci géant ! Entre les soirées techno, les aprèms pétanque, les

soirées à la Barbote avec une partie de flipper entre chaque verre de bières, les soirées jeux et séries à la maison, les anniversaires surprises, les danses dans la cuisine du Crous, les brunchs du dimanche (où, avouons-le, on passait plus de temps à chercher un spot pas trop bondé et à faire la queue, mais c'était top quand même). Merci à Edouard, Steven, Pierre, Gaby, Theo, Zaineb, Gwenaël, Xiao, Emmanuel, Louise pour tous ces moments.

Un big up à mes "mésaventurières" de la science, Soumaiya et Amélia, mes potes depuis la L3. On a plongé ensemble dans ce monde fou de la recherche, et même si j'ai pris un autre chemin après le Magister (la paillasse, très peu pour moi !!!) pour aller foirefouiller avec des ordinateurs, nos appels interminables, nos lamentations, nos fous rires et notre soutien mutuel ont été un vrai support moral ! C'est toujours un crève-cœur de ne pas se voir aussi souvent qu'on le voudrait avec nos thèses qui nous accaparent, mais chaque retrouvaille est pépite (clin d'œil à Arles quoi ?)

Et vous, mes copines du lycée, pas question de vous oublier ! Surtout toi, Dania, ouah, plus de 12 ans d'amitié, merci d'être cette pote en or sur qui je peux compter aveuglément. On a tant grandi depuis l'époque où on regardait Naruto au collège ! Et merci à vous, Naima, Imen, et Océane, de m'avoir sorti la tête de ma thèse ! Purée, nos vacances, toujours aussi folles et mémorables, dans des coins paumés ou non, avec AirBNB tous plus loufoques les uns que les autres (Douchy, la cave qui rappelle Charles Péguy, la fête à la fraise, les vlogs du Lidl), heureusement qu'il y avait le fameux jacuzzi!

Enfin, un immense merci à ma famille pour son soutien sans faille. À mon père, qui avec sa force de volonté inébranlable et sa capacité à naviguer à travers les changements avec tant d'assurance, m'a toujours inspirée. Pour ma mère, cette citation de Deborah Levy me vient à l'esprit : « Il faut de l'habileté, du temps, de la dévotion et de l'empathie pour créer un foyer où tout le monde se sent bien. Être l'architecte du bien-être des autres est un acte de générosité immense. Bien des gens pensent que cette tâche est réservée aux femmes. Par conséquent, on utilise toute sorte de mots pour minimiser cet effort monumental. » Ton dévouement et ton amour ne sont pas minimisés à mes yeux, Maman, et je te suis éternellement reconnaissante. Et à mon frère Vincent, ton imagination sans fin et ta bonne humeur ont toujours été une lumière dans ma vie. Cám ơn à vous tous (et sans oublier les petits Titi et Toto) !

Un merci final pour David, merci de faire partie de ma vie et merci d'être le coéquipier le plus solide avec qui je pourrais affronter toutes les aventures. (oulà je me crois trop dans un shonen). Merci de m'aider autant que tu le fais, d'être si attentif et présent dans mes moments de doute et d'angoisse. Ta gentillesse, ta bienveillance et ton humour rendent mon quotidien tellement plus léger et amusant ! Comme a dit un grand sage un jour « Il n'y a pas de lézards sous le rocher. » dans notre relation !



---

# INTRODUCTION GENERALE

---



© Produit par DALL·E

La beauté, à la fois énigmatique et familière, est une mélodie silencieuse qui résonne en écho dans l'esprit humain. Elle se faufile entre les pages de la littérature, s'infiltré dans les toiles des artistes et suscite les débats passionnés des philosophes. Elle est à la fois le miroir de notre culture et le reflet d'universaux qui transcendent les frontières et le temps. Si autrefois, nous nous sommes contentés de l'admirer, aujourd'hui, les avancées en neuroesthétique et en sciences cognitives ont élargi notre horizon, mettant en lumière son rôle crucial dans la perception esthétique. Parmi les nombreuses hypothèses sur les mécanismes biologiques de la beauté, la théorie de la fluence émerge comme un domaine essentiel d'investigation, explorant comment la facilité avec laquelle notre cerveau traite un objet influence notre appréciation esthétique de celui-ci. Par la suite, la biologie a également jeté un regard sur les causes évolutives de la beauté, illustrant comment elle peut conférer des avantages dans les interactions sociales et la reproduction au sein de nombreuses espèces. Ces démarches mettent en exergue une interconnexion profonde entre la biologie et les sciences cognitives dans la quête de déchiffrer le mystère de la beauté.

Cette thèse se propose d'explorer le concept de beauté et d'analyser ses implications biologiques et sociales, en utilisant un primate non-humain comme organisme modèle. Dans le premier chapitre nous introduirons notre modèle principal d'étude, les mandrills, ainsi que la base de données utilisée dans le cadre de cette thèse. Dans, le second chapitre, nous explorerons la féminité faciale comme un des critères de beauté et nous nous intéresserons aux enjeux sociaux et sexuels de cette féminité chez les femelles mandrills. Le troisième chapitre se consacrera à l'étude de la féminisation des jeunes mâles mandrills en explorant les théories sous-jacentes à cette féminisation et ses manifestations comportementales. Enfin, le quatrième chapitre, à caractère méthodologique, se concentrera sur la quantification de la fluence : nous comparerons la capacité de deux métriques de fluence mesurées à partir de réseaux de neurones artificiels, la typicalité statistique et la sparsité neuronale, à prédire la beauté faciale.

## I. Cadre philosophique de la beauté

Dans cette quête pour définir et comprendre la beauté, deux grands courants de pensées divisent les philosophes : le courant objectiviste et le courant subjectiviste.

L'esthétique objectiviste trouve ses racines dans l'Antiquité, incarnée par des penseurs tels que Platon. Pour lui, la beauté est une essence universelle, constante et inaltérable, nichée au cœur des objets eux-mêmes — des colonnes stoïques des temples grecs aux mélodies harmonieuses des compositions classiques. Cette beauté ne demande qu'à être reconnue par le spectateur attentif, comme l'exprime Platon dans le *Banquet* :

*« Cette beauté est éternelle, inaltérable, inchangeable ; elle ne se dégrade ni en apparaissant ici ou là, maintenant ou alors ; elle ne croît ni ne décroît ; elle n'est pas belle d'un côté et laide de l'autre, ni belle en tel temps, en tel lieu, sous tel rapport, dans telle situation ; elle n'est pas belle pour les uns et laide pour les autres. » - Banquet, Platon*

Aristote, marchant sur les traces de son mentor, a mis en avant l'ordre et la structure — telle que la symétrie, la proportion et l'harmonie—comme piliers d'une tragédie réussie. Pour lui, la beauté résidait dans des qualités universelles. Cette perspective esthétique a été affinée par de nombreux penseurs, dont certains se sont tournés vers des concepts mathématiques. L'influence pythagoricienne est palpable dans la pensée objectiviste. Les Pythagoriciens, convaincus que la beauté s'ancrait dans des relations numériques, ont mis en avant des concepts tels que le nombre d'or. Cet idéal mathématique est incarné par des chefs-d'œuvre architecturaux comme le Parthénon, qui repose sur une symétrie parfaite et l'utilisation du nombre d'or. Selon Beardsley (Beardsley, 1981), l'appréciation esthétique est ancrée dans les caractéristiques objectives d'un objet qui guident un « jugement de valeur » conforme à des règles édictées. Pour illustrer ce point, prenons les vers de Racine : leur beauté découle des normes strictes de l'alexandrin. Ainsi, la connaissance et le respect de ces règles poétiques sont essentiels. L'appréciation de l'art serait-il donc l'apanage d'une élite capable de décrypter les rimes, la structure des phrases, ou encore de reconnaître le nombre d'or dans la suite de Fibonacci ? Pendant longtemps, ce courant objectiviste soutenait que pour apprécier pleinement la beauté, les spectateurs devaient avoir une connaissance approfondie des conventions artistiques.

Tout en reconnaissant l'importance des règles traditionnelles, l'esthétique subjectiviste offre une perspective distincte, axée sur la question du goût individuel. Elle privilégie l'expérience personnelle, mettant en avant l'émotion et la perception propre à chaque individu dans l'appréciation esthétique. Comme le souligne David Hume dans *Des normes du goût*, la beauté ne réside pas intrinsèquement dans l'objet observé, mais émerge de l'esprit qui le contemple.

*« La beauté n'est pas une qualité inhérente aux choses elles-mêmes, elle existe seulement dans l'esprit qui la contemple » - Hume*

Kant joue un rôle crucial dans la théorie subjectiviste de la beauté. Dans son essai *Critique de la faculté de juger* datant de 1790, il explore la tension entre la subjectivité individuelle et un idéal universel. Lorsqu'il avance l'idée que :

*« Le beau est ce qui plaît universellement sans concept » - Kant*

le début de la phrase « *Le beau est ce qui plaît* » de Kant met l'accent sur le caractère intrinsèquement subjectif du jugement esthétique. Quand on dit que quelque chose est beau parce qu'il « *plaît* », cela renvoie à une expérience personnelle et à une réaction émotionnelle propre à chaque individu. Dans cette perspective, le beau c'est ce qui plaît à une personne, indépendamment des autres. Cependant, avec l'adverbe « *universellement* », Kant introduit une forme de contradiction. Alors que « *ce qui plaît* » est propre à l'individu, « *universellement* » suggère une acceptation ou une reconnaissance collective de cette beauté. Autrement dit, bien que la beauté soit une expérience personnelle, elle contient également une prétention à l'universalité. Ainsi, bien que le jugement de beauté provienne d'un sentiment intime et personnel, il comporte aussi une attente que ce sentiment soit partagé par tous. Cette universalité présumée est d'autant plus intrigante que Kant y adjoint l'idée du « *sans concept* ». Selon lui, la beauté ne peut être définie ou justifiée rationnellement. On ne peut pas argumenter ou prouver la beauté d'une œuvre ou d'un objet, car elle réside dans l'émotion qu'elle évoque, et non dans une liste de critères objectifs.

Au cours de la seconde moitié du XXe siècle, l'émergence d'une approche interactionniste va unifier ces courants de pensées. L'essence de ce courant repose sur une idée fondamentale : l'expérience esthétique naît de l'interaction dynamique entre l'objet (qu'il soit artistique ou autre) et son observateur.



Tout d'abord, John Dewey, dans *Art as Experience* (Dewey, 1934), soutient que l'esthétique n'est pas un domaine élitiste, séparé de la vie ordinaire ou réservé à quelques initiés. Pour lui, l'art transcende les musées, se manifestant dans les moments les plus banals de la vie. Cependant, Dewey insiste sur l'engagement total dans ces expériences, affirmant que leur essence réside non seulement dans la perception de la beauté, mais aussi dans l'immersion totale dans le moment.

Jean-Marie Schaeffer, dans *L'Expérience esthétique*, élabore et approfondit la perspective introduite par Dewey (Schaeffer, 2015). Tout comme ce dernier, Schaeffer s'écarte des conceptions traditionnelles qui cantonnent l'esthétique à des œuvres d'art canoniques ou à des expériences purement subjectives. Néanmoins, là où Schaeffer ajoute une dimension unique, c'est dans sa minutieuse exploration des mécanismes cognitifs et attentionnels à l'œuvre dans l'expérience esthétique. Selon lui, l'esthétique ne découle pas simplement d'une interaction, mais fait appel à des processus cognitifs profonds, avec l'attention en tant qu'acteur central. Au lieu d'une attention pragmatique orientée vers un but, celle déployée dans une expérience esthétique est sans contrainte, exploratoire et non dirigée. Schaeffer va plus loin en introduisant le concept de « calculateur hédonique », selon lequel notre cerveau évalue constamment le plaisir ou le déplaisir liés à une expérience, guidant ainsi notre engagement envers elle.

*« L'expérience esthétique est une expérience humaine de base, et plus précisément une expérience attentionnée exploitant nos ressources cognitives et émotives communes. » - L'Expérience esthétique, Jean-Marie Schaeffer*

Illustrons ceci avec le tableau (Figure 1) présentant une table garnie. À première vue, le spectateur pourrait vouloir saisir un verre de vin et une tarte. Toutefois, c'est la lumière jouant sur les verres vides ou la texture de la nappe qui captent réellement l'attention. Ces détails, bien qu'accessoires dans un contexte utilitaire, deviennent fondamentaux dans notre expérience esthétique.



Figure 1 | Heda, Willem Claeszoon - Breakfast Table with Blackberry Pie

Le rôle prépondérant accordé à l'attention lors de notre interaction avec une œuvre nous conduit-il à un état similaire à l'hypnose ? Les études menées par la neurologue Marie-Noëlle Metz-Lutz semblent indiquer des parallèles (Metz-Lutz et al., 2010). En analysant les réactions cérébrales durant une représentation théâtrale, la chercheuse a constaté que des émotions intenses tendent à accélérer le rythme cardiaque, celui-ci ralentit lors des scènes les plus prenantes. De surcroît, le prénuceus, une région du cerveau liée à la perception émotionnelle, se désactive, un phénomène biologique également observé dans des expériences d'hypnose. Là où notre cerveau trie et régule naturellement les informations face à des situations complexes (Mushtaq et al., 2011), en état d'hypnose, notre discernement est réduit, entravant notre capacité à hiérarchiser les informations. Dans le cadre de l'expérience esthétique, les priorités habituelles de notre perception seraient alors suspendues, donnant une importance inattendue à ce qui est généralement considéré comme "superflu" (la nappe du tableau), sans hiérarchie établie.

Ainsi, tandis que la philosophie traditionnelle s'attache à explorer les fondements et l'essence de l'expérience esthétique, la neuroesthétique et les sciences

cognitives s'efforcent de déchiffrer les mécanismes neuronaux qui la sous-tendent. Certains penseurs comme Schaeffer ont joué un rôle essentiel en tentant de relier ces disciplines et en cherchant à établir une vision intégrative de la beauté qui prendraient en compte les bases biologiques révélées par l'esthétique empirique. Ainsi, plongeons plus en profondeur pour découvrir ce que ces sciences de l'esthétique ont à nous offrir.

## II. Apport des sciences cognitives

### A. Origines et concept de la théorie de la fluence

L'empirisme en esthétique, profondément enraciné dans les travaux avant-gardistes de Gustav Fechner en 1876 avec *Vorschule der Ästhetik*, a été élargi par les recherches contemporaines (Fechner, 1876). À l'opposé de l'esthétique philosophique, cette approche cherche à étudier des réactions quantifiables, comme le plaisir ressenti. Les méthodologies dans ce domaine nécessitent par exemple que les participants sélectionnent parmi une série de stimuli visuels ceux qu'ils jugent esthétiquement plaisants. Ces techniques ont posé les bases pour le développement de la neuroesthétique, un champ qui, comme son nom l'indique, se concentre sur l'étude des bases neuronales de l'appréciation esthétique à travers des techniques d'imagerie cérébrale. Conjugée à la psychologie de l'esthétique, ces disciplines forment ce qu'on nomme « esthétique expérimentale », ou « empirical aesthetics » en anglais.

Au sein de ce vaste champ, un jalon particulièrement important a été l'introduction de la théorie de la **fluence perceptive** principalement développée par Reber, Schwarz, et Winkielman en 1998, puis étendue en 2004. La fluence perceptive peut être définie comme la facilité avec laquelle une information est traitée perceptivement.

Dans leur étude pionnière de 1998, Reber, Schwarz et Winkielman démontrent empiriquement comment la fluence perceptive affecte les jugements esthétiques (R. Reber et al., 1998). Dans une première expérience, des participants ont évalué 20 dessins de représentations neutres, comme des chevaux ou des avions. Les dessins et leur arrière-plan étaient délibérément altérés pour rendre leur reconnaissance plus difficile. Chaque dessin était précédé d'une amorce visuelle (*primer*), exposée brièvement pendant 25 millisecondes. Ces amorces étaient des contours correspondant au dessin cible ("amorce correspondante") ou des contours appartenant

à un autre dessin ("amorce non correspondante"). Les dessins précédés par des amorces correspondantes ont été non seulement reconnus plus rapidement, mais aussi jugés plus attractifs que ceux précédés par des amorces non correspondantes (Figure 2).

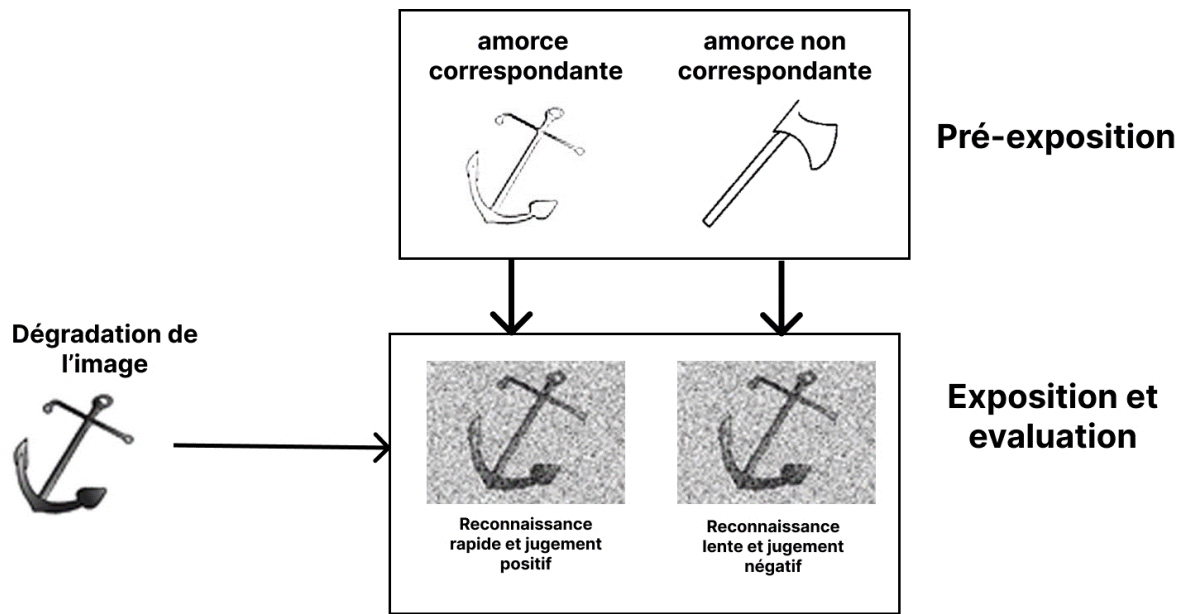


Figure 2 | Schéma de l'expérience 1 de Reber et al. 1998. J'ai créé ce schéma en m'inspirant des illustrations qu'ils ont utilisées, provenant de la base de données d'images de Snodgrass and Vanderwart (1980).

Avant cette étude pionnière, c'était l'effet d'exposition étudié par Bornstein & D'Agostino en 1992 qui avait introduit l'idée que la facilité avec laquelle on traite un objet pourrait influencer le plaisir esthétique (Bornstein & D'Agostino, 1992). Les auteurs avaient avancé l'idée que la familiarité accrue avec un objet, due à des expositions répétées, pourrait faciliter sa perception, donnant lieu à une sensation de fluidité associée à une réaction positive. Cependant, l'étude de 1998 (R. Reber et al., 1998) va consolider cette hypothèse, montrant que n'importe quelle condition améliorant la fluidité perceptive pouvait renforcer l'attractivité visuelle, même si l'objet était vu pour la première fois. Plusieurs autres études, en jouant sur des paramètres telles que l'amorçage (Albrecht & Carbon, 2014), le contraste (Reber, Winkielman, and Schwarz 1998) ou encore la durée de visualisation du stimulus (Forster et al., 2015) ont renforcé l'idée que la fluidité dans la perception conduisait à des évaluations esthétiques plus positives.

**Les fondement du cadre théorique:** Il aura fallu attendre 2004 pour que ces même auteurs -Reber, Schwarz et Winkielman- érigent un cadre théorique autour de la notion de fluence, plaçant celle-ci au cœur de l'expérience esthétique à travers une optique interactionniste (R. Reber et al., 2004). Dans leur article, ils mentionnent explicitement l'expérience esthétique comme une interaction entre les propriétés du stimulus et les processus cognitifs et affectifs de la personne qui la perçoit. Les préférences esthétiques sont dictées par la facilité avec laquelle un observateur peut traiter un objet donné.

Dans le paysage de la recherche sur l'expérience esthétique, cet article constitue un véritable tournant en offrant un cadre unificateur qui rassemble une diversité de variables auparavant étudiées séparément. Ces variables s'articulent autour de deux pôles principaux : les caractéristiques objectives du stimulus et les expériences passées de l'observateur. Pour ce qui est des caractéristiques intrinsèques des stimuli, des propriétés telles que la simplicité, la symétrie grâce à sa structure prévisible, et le contraste marqué entre la figure et le fond facilitent le traitement cognitif et renforcent leur attrait esthétique. Pour le côté subjectif de l'observateur, ses expériences antérieures, comme l'exposition répétée à un stimulus, peuvent améliorer la facilité du traitement. Pour affiner davantage notre compréhension, je développerai en détail l'exemple de la prototypicalité (*B. Un cas particulier de la fluence : la prototypicalité*).

Avant de poursuivre, il est essentiel de clarifier quelques termes et distinctions, comme proposé par les auteurs de l'article. Bien que la littérature distingue la fluence perceptive de la fluence conceptuelle (définies dans l'encadré 1), les auteurs affirment que ces deux types de fluence se manifestent de manière équivalente dans l'expérience vécue: elles conduisent à un sentiment généralisé de fluidité, ou de facilité dans le traitement de l'information, et influencent de manière similaire nos jugements. Les auteurs regroupent ces deux aspects sous le terme *visual fluency*. De même, j'utiliserai simplement le mot **fluence** dans le cadre de cette thèse, qui traite toutefois essentiellement de la fluence perceptive.

Sur le plan sémantique, la fluence se divise également en deux types : objective et subjective (voir Box 1). La fluence objective est mesurée par des critères tangibles comme la vitesse et la précision du traitement cognitif. À l'inverse, la fluence subjective relève de notre ressenti pendant le traitement, comme une impression de facilité ou de

faible effort. Généralement, ces deux formes de fluence vont de pair, même si des exceptions existent. Par exemple, sous l'effet de l'alcool, on pourrait ressentir une fluence subjective élevée, même si nos performances objectives ne suivent pas. Dans cette thèse, je me suis bornée à la fluence objective.

#### Encadré 1 | Définitions de la fluence

**Fluence objective :** C'est la mesure quantifiable de la facilité du traitement cognitif d'un stimulus. Elle se manifeste par une rapidité de traitement, une efficacité et une consommation minimale de ressources mentales, sans pour autant que le sujet soit conscient de cette facilité.

**Fluence subjective :** C'est l'expérience consciente et subjective qu'un individu a de la facilité du traitement d'un stimulus. Elle se traduit par une sensation de facilité, de rapidité, et de faible effort lors de l'interprétation ou de la compréhension d'un stimulus.


**Fluence perceptive ou « ce qui plaît aux sens » :** Elle concerne les opérations cognitives de bas niveaux, principalement axées sur les caractéristiques visibles et directement perceptibles d'un stimulus. Ces opérations sont généralement influencées par des éléments comme la répétition d'un stimulus, son contraste, sa durée d'apparition, et autres facteurs visuels.

**Fluence conceptuelle :** Elle fait référence à la facilité avec laquelle un stimulus est intégré ou relié aux structures de connaissance sémantique préexistantes d'un individu. Cette facilité est souvent affectée par des éléments tels que la familiarité, la prototypicalité, les associations sémantiques. Lors de la résolution d'un problème, c'est cette fluidité conceptuelle qui émerge, ce qui pourrait être assimilé au "plaisir épistémique" de l'eureka, ressenti lorsque ce qui semblait embrouillé et chaotique dans nos circuits neuronaux devient soudainement clair et fluide.

## B. Un cas particulier de la fluence : la prototypicalité

Un **prototype** est un représentant d'une catégorie perceptive (Mervis & Rosch, 1981). En psychologie cognitive, la catégorisation d'objets ou d'entités est réalisée à partir des prototypes à travers un processus d'abstraction. Le prototype d'une catégorie serait une sorte de « tendance centrale », parfois conceptualisée comme la moyenne arithmétique (Posner, 1970) ou le mode (Neumann, 1977) de tous les exemples de cette catégorie auxquels un individu a été exposés. Par exemple, si on vous demande de penser à un « oiseau », l'image qui vous vient à l'esprit est probablement plus proche d'un moineau que d'une autruche, car le moineau est plus prototypique de la catégorie « oiseau » pour beaucoup de gens (Figure 3).

### A Prototypical "Bird" Image (for an Urban American)



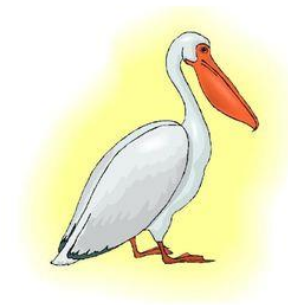
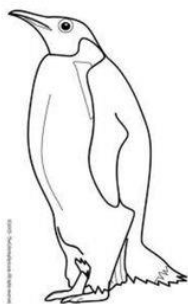
© 2007 Thomson Higher Education

- The prototypical "bird" image is similar to the average of many bird images.

### Not Prototypical "Bird" Image



Psych 355, Miyamoto, Spr '18



Basic Claim of Prototype Theory

9

Figure 3 | Exemples d'oiseaux prototypiques (Haut) vs. non prototypiques (Bas)

Des études montrent que les prototypes sont catégorisés plus rapidement et nécessitent moins de ressources neuronales pour leur traitement (Posner & Keele, 1968; P. J. Reber et al., 1998). Dans ce contexte, il est crucial de noter le travail de Winkielman et collaborateurs (2006), qui souligne que cette fluidité cognitive accrue

ne conduit pas seulement à une facilité de traitement, mais également à une réponse émotionnelle positive (Winkielman et al., 2006). Ce cadre de « **fluidité hédonique** » propose que les prototypes sont attractifs en partie parce qu'ils sont « fluents », c'est-à-dire facilement reconnaissables et traitables.

Par exemple, les travaux de Langlois et Roggman (1990) ont montré qu'un visage prototypique, construit en moyennant plusieurs visages, est souvent jugé plus attractif que les visages individuels qui l'ont formé ((Langlois & Roggman, 1990) ; Figure 4). L'effet de l'attractivité des visages prototypiques a été corroboré dans de nombreuses études (Little et al., 2011; Gillian Rhodes, 2006). Il est important de noter que cet effet demeure significatif même lorsque d'autres variables potentiellement influentes pour l'attractivité, comme la symétrie faciale sont contrôlées (Gillian Rhodes et al., 1999; Valentine et al., 2004). Les études sur les nourrissons montrent également que ces derniers sont capables de former un prototype de visages après une exposition limitée (Walton & Bower, 1993) et les apprécient (Damon et al., 2017; Quinn et al., 2010). La fluence, suggérée par Winkielman comme explication de l'attractivité des visages prototypiques, a été source de critiques, particulier concernant les méthodes employées pour créer un visage prototypique. À titre illustratif, la Figure 4 montre que la peau sur l'image composite est lissée, ce qui est couramment associé à des signes de bonne santé. Par conséquent, l'attractivité du visage prototypique pourrait non pas découler de sa fluence, mais plutôt du fait qu'il véhicule des indices de bonne santé. Ajouté à cela plus le nombre de visages réels utilisés pour créer ce visage composite est élevé, plus ce dernier est jugé attractif.



[Research from: Langlois & Roggman (1990). Photos courtesy of Dr. Marin Gruendl, [www.beautycheck.de](http://www.beautycheck.de)]

Figure 4 Préférence pour le visage moyen : En réalisant une moyenne numérique de photos de visages, on obtient un visage composite.



### III. L'angle évolutionniste : origines et conséquences de la beauté

#### A. Théories évolutionnistes de la beauté

Jusqu'à présent, la beauté a été principalement abordée à travers le prisme de la philosophie, des sciences cognitives et de la psychologie, toujours axé sur l'expérience humaine. Bien que Darwin ait postulé un sens esthétique (« *sense of beauty* ») chez certains animaux, la plupart des évolutionnistes réduisent l'esthétique à des traits simplement attractifs, considérés comme des indicateurs de qualité (Andersson, 1994; David M. Buss & Barnes, 1986). Dans cette optique, il n'y a tout simplement pas de place pour une attitude esthétique et la théorie des indicateurs de qualité et de ces dérivées telle que le Principe du Handicap d'Amotz Zahavi (Zahavi, 1997) a largement influencé le domaine pour expliquer les traits attractifs. Cette théorie suggère que des traits coûteux à produire, comme la queue extravagante du paon, servent de signaux « honnêtes » de la qualité d'un individu. Des recherches sur la symétrie faciale chez les humains confortent cette idée, indiquant que des visages symétriques pourraient signaler une meilleure homéostasie durant le développement, reflétant une meilleure qualité génétique (Gillian Rhodes et al., 2007). Pour en revenir à la prototypicalité (distance au prototype) des visages, leur préférence peut s'expliquer par cette théorie : Les visages éloignés d'un prototype pourraient signaler des anomalies génétiques ou développementales (Thornhill & Møller, 2007). Les visages moyens sont en effet souvent perçus comme plus sains (G. Rhodes et al., 2001). La recherche de Zebrowitz et Rhodes (2004) a également souligné que les visages distinctifs étaient associés à une perception de santé moindre (Zebrowitz & Rhodes, 2004). La féminité, dans ce contexte, peut être interprétée comme un cas particulier de prototypicalité, où le prototype est défini spécifiquement pour la catégorie « femme ». Bien que les visages féminins moyennés soient généralement valorisés (A. L. Jones & Jaeger, 2019), cette assertion est sujette à controverse. En effet, les traits féminins, envisagés sous un angle dimorphique - c'est-à-dire en mettant l'accent sur les différences morphologiques entre les sexes -, sont souvent privilégiés en tant qu'indicateurs des qualités inhérentes à celles qui les possèdent. Ces traits sont fréquemment associés à des taux élevés d'œstrogènes et, par conséquent, considérés

comme liés à la fertilité et à des compétences maternelles chez les femmes qui présentent ces traits. (R. Reber et al., 1998).

Néanmoins, la théorie des indicateurs de qualité ne permet pas d'expliquer toutes les préférences. Par exemple, elle n'explique pas pourquoi des poulets montrent une préférence pour des visages humains féminins prototypiques (ceux jugés les plus attractifs par des humains) (Ghirlanda et al., 2002). Si une préférence généralisée pour la prototypicalité est observée dans des contextes qui n'ont pas de conséquences directes sur la valeur sélective individuelle (comme le cas du poulet préférant des visages de femmes prototypiques), cela pourrait suggérer que cette préférence découle d'un mécanisme cognitif plus général. Il pourrait s'agir d'un effet secondaire d'une autre adaptation pour traiter plus rapidement l'information visuelle, plutôt que d'une stratégie spécifiquement conçue pour la sélection de partenaires reproductifs de qualité (D. M. Buss et al., 1998). C'est ce que soutiennent les recherches de Halberstadt et Rhodes sur la prototypicalité. Ils ont exploré des catégories sans liens apparents avec la valeur sélective individuelle telles que les chiens, les montres et les oiseaux (J. Halberstadt & Rhodes, 2000). Leurs résultats ont indiqué une forte attraction pour des objets prototypiques (ici la moyenne) suggérant que cette préférence pourrait être enracinée dans des mécanismes cognitifs universels. Néanmoins des critiques ont été faites sur ces travaux, des chercheurs ont proposé que l'effet de familiarité pourrait se confondre avec la prototypicalité. Autrement dit, ce n'est pas nécessairement le caractère prototypique des oiseaux, des poissons ou des montres qui les rend attractifs, mais plutôt une familiarité ou une ressemblance avec des entités fréquemment rencontrées ou perçues qui pourraient susciter une préférence. Ainsi, dans une étude ultérieure, les auteurs ont ajouté un contrôle pour la familiarité subjective (Jamin Halberstadt & Rhodes, 2003) sur d'autres catégories de stimuli. Les résultats ont indiqué que l'attrait pour le prototype demeure même après avoir contrôlé pour la familiarité subjective pour les oiseaux et les poissons (mais pas pour les voitures). Cela suggère que plusieurs mécanismes contribuent à l'attraction pour le prototype. Un de ces mécanismes pourrait être une préférence générale pour la familiarité et l'autre une attraction intrinsèque pour le prototype en lui-même. Pour contrôler le facteur de la familiarité dans l'évaluation de l'attractivité des stimuli, une étude menée par Winkielman et al. en 2006 a exposé les participants à des ensembles de points arbitrairement classés en deux catégories, désignées comme « Acks » et « Blubs »

(Winkielman et al., 2006). Les participants ont été testés ultérieurement à l'aide d'un ensemble mixte de stimuli, incluant des motifs déjà présentés et des prototypes jamais exposés. Les données ont révélé une corrélation positive entre la prototypicalité (distance au prototype) et la rapidité de la classification effectuée par les participants. Cette fluidité de classification prédisait, à son tour, une évaluation positive des stimuli. Ce résultat souligne que l'attrance pour la prototypicalité n'est pas uniquement liée à des objets ou des situations familières ou ayant des conséquences en termes de valeur sélective individuelle. Elle semble aussi s'étendre à des stimuli arbitraires, ce qui renforce l'idée que des mécanismes cognitifs plus généraux pourraient être à l'œuvre. Richard Prum propose de revenir à la notion darwinienne d'une esthétique pure, où certains traits évoluent simplement pour le plaisir qu'ils procurent, sans refléter nécessairement une qualité génétique (Prum, 2018). Il suggère que la beauté peut avoir une valeur intrinsèque pour l'observateur, indépendamment de tout avantage adaptatif.

La théorie de l'exploitation sensorielle a ouvert de nouvelles perspectives sur la compréhension de l'évolution des traits attractifs (Ryan & Rand, 1990). Selon cette théorie, des traits particuliers peuvent évoluer s'ils « exploitent » des préférences sensorielles préexistantes. Par exemple, certaines espèces, influencées par des préférences pour des couleurs ou des motifs spécifiques (e.g. la couleur rouge des fruits consommés par les primates de l'Ancien Monde), peuvent voir l'évolution de traits associés (e.g. la préférence pour les couleurs vives de leurs partenaires sexuels). Prenez les interactions entre les abeilles et les fleurs : les abeilles montrent une préférence pour des signaux floraux exagérés (en termes de taille et de contraste), même si ces traits ne reflètent pas la qualité de la fleur (Naug & Arathi, 2007). Ces biais pourraient être exploités par des plantes qui, au cours de l'évolution, vont augmenter de taille et de contraste. Chez plusieurs espèces de poissons Goodeinae, les mâles arborent une bande jaune à l'extrémité de leur queue, simulant l'apparence d'un ver, une source de nourriture pour ces poissons. Selon Garcia et Ramirez (2005), cette caractéristique exploite le système sensoriel des femelles en quête de nourriture (les vers), les attirant ainsi vers les mâles. (Garcia & Ramirez, 2005).

## B. La fluence: un exemple de biais perceptif

En intégrant la notion de fluence (issue des sciences cognitives) à la biologie évolutive, celle-ci apparaît comme un mécanisme sous-jacent de l'exploitation sensorielle. Les signaux qui tirent parti de cette fluence, en optimisant le traitement neurocognitif du perceuteur pour se rendre attractif, pourraient ainsi bénéficier d'un avantage sélectif (J. P. Renoult & Mendelson, 2019). L'attrait des humains pour les images reflétant les statistiques spatiales des scènes naturelles en est une manifestation. L'étude de Redies et al. (2007) établit un lien entre l'esthétique de l'art et la beauté des scènes naturelles, en suggérant que ces deux éléments partagent des propriétés statistiques fractales et invariantes à l'échelle (Redies et al., 2007). Cette ressemblance repose sur l'idée que le système visuel traite de manière optimale les propriétés statistiques des stimuli naturels grâce à un codage optimal et économe - donc fluent- (voir partie IV.B.2). Les auteurs suggèrent que l'artiste, dans sa création, vise à induire une résonance particulière dans le système visuel de l'observateur, une résonance qui découle de l'adaptation de ce dernier aux stimuli naturels, conduisant ainsi à une appréciation esthétique optimale. À l'instar des artistes qui reproduisent les caractéristiques spatiales des paysages naturels pour rendre leurs créations plus séduisantes, un modèle basé sur l'exploitation sensorielle intégrant la fluence prévoit que les organismes évoluent en adaptant leurs signaux de communication pour qu'ils reflètent les statistiques propres à leur environnement. Si cette hypothèse s'avère correcte, alors les signaux de différentes espèces devraient refléter les caractéristiques spécifiques de leur habitat. La prédiction selon laquelle les statistiques visuelles des ornements correspondront à celles de leur habitat a été récemment renforcée par une étude (Hulse et al., 2020). En explorant les pentes de Fourier des habitats de diverses espèces parmi dix types de poissons dards (*Etheostoma* spp.), en lien avec leurs comportements de parade nuptiale, il a été révélé que les motifs visuels des mâles (un signal sexuel attractif) se conformaient aux caractéristiques visuelles de leur environnement. Cette correspondance n'a pas été identifiée chez les femelles, ce qui suggère que les mâles ont adapté leurs motifs visuels pour refléter les caractéristiques visuelles de leur habitat spécifique dans le but d'attirer les femelles. Ces découvertes appuient l'hypothèse de l'exploitation sensorielle, reconnaissant ainsi la fluence dans le traitement de l'information comme une force majeure dans l'évolution des signaux.

## C. Conséquences sociales et sexuelles de la beauté

Quelque soit l'origine évolutive du lien entre fluence et attractivité, l'impact des préférences esthétiques, en particulier concernant la prototypicalité de sexe des visages, sur les interactions sociales et sexuelles demeure un domaine d'étude complexe. Bien que les recherches en laboratoire mettent en lumière une préférence pour les prototypes, et notamment pour la féminité comme exemple de prototypicalité, il est essentiel de distinguer une préférence visuelle des comportements réels qu'elle pourrait engendrer.

Les outils de recherche en laboratoire, tels que les tests de choix (Chen & Risen, 2010; Fechner, 2016) ou encore le temps passé à regarder des images (Winters et al., 2015), test à l'origine destiné aux nourrissons (Fantz, 1958) et ensuite adapté pour d'autres espèces (Wilson et al., 2021), montrent leurs limites lorsqu'il s'agit d'interpréter ces préférences dans un contexte naturel. Prenons pour exemple les femelles macaques rhésus : bien qu'en laboratoire, elles montrent une attirance pour les visages masculins rouges (Waitt et al., 2003) , elles manifestent sur le terrain une préférence pour les mâles au teint plus foncé (Dubuc et al., 2014). Cette discordance est également observée chez les humains, qui peuvent verbaliser cette distinction : bien que certains hommes, indépendamment de leur orientation sexuelle, évaluent de manière cohérente l'attractivité d'un visage de femme, leurs choix ou préférences réels en matière de partenaires peuvent diverger (Kranz & Ishai, 2006).

Il est donc primordial d'effectuer des observations directes en milieu naturel pour mieux appréhender les préférences. L'étude des conséquences sociales et sexuelles de la beauté, considérée sous un angle évolutif, nous donne un aperçu des raisons pour lesquelles certaines préférences esthétiques émergent chez certaines espèces.

Se pose alors une question centrale : les individus présentant des traits esthétiquement préférés sont-ils systématiquement choisis comme partenaires ou bénéficient-ils d'une meilleure intégration sociale ? La réponse est loin d'être tranchée. Les deux hypothèses principales évoquées sont actuellement débattues : l'hypothèse de la fluence (comme cas particulier de l'exploitation de biais sensoriels) et celle des

indicateurs de qualité. L'hypothèse de la fluence propose que la beauté, devrait **systematiquement** offrir des avantages dans les interactions, dans n'importe quel contexte, dès lors que l'information est traitée facilement par le système cognitif. En revanche, la théorie des indicateurs suggère que la beauté pourrait signaler d'autres attributs bénéfiques, et donc qu'**un traitement facile de l'information n'est pas systematiquement associé à un avantage**. Chez l'espèce humaine, si la beauté peut offrir certains avantages, elle n'est pas exempte d'inconvénients. Par exemple, dans certaines circonstances, des femmes hautement féminines (un estimateur de prototypicalité et donc beauté) pourraient être considérées comme moins compétentes dans un contexte professionnel par exemple (Sczesny et al., 2006). Les hommes pourraient également hésiter à choisir les femmes les plus attractives et féminines, craignant qu'elles soient moins fidèles et qu'ils se retrouvent à élever un enfant qui n'est pas biologiquement le leur (Little et al., 2014). Cette dynamique est également présente chez d'autres espèces animales, comme les rosélins domestiques où l'on note que les mâles jugés moins attractifs interagissent plus fréquemment avec différents partenaires sociaux comparés à ceux perçus comme plus attractifs (Oh & Badyaev, 2010). Enfin, pour traiter la question, il est essentiel de souligner que la réaction d'un individu face à un stimulus visuel ne se limite pas à sa simple préférence. Plusieurs facteurs, internes tels que la faim ou le statut hormonal, ou externes comme l'environnement, la présence de prédateurs ou les ressources, peuvent influencer cette réponse.

## IV. Modéliser la Fluence

### A. Approches classiques pour quantifier la fluence

La quantification de la fluence constitue un enjeu majeur dans l'étude des préférences et des jugements esthétiques. Les premières études dans ce domaine ont cherché à identifier les propriétés intrinsèques des stimuli—telles que la simplicité, la symétrie, le contraste et la luminosité—qui pourraient contribuer à la facilité de traitement perceptif et, par conséquent, augmenter l'appréciation esthétique de ces stimuli.

Des méthodes variées ont été employées pour mesurer la fluence. Etant donné que la fluence se définit par la facilité relative de traitement, elle peut donc être évaluée par des méthodes mesurant la rapidité ou la difficulté de traitement. Une méthode courante consiste à enregistrer **le temps de réponse aux stimuli**. Par exemple, dans l'étude présentée précédemment (R. Reber et al., 1998), le temps de réponse pour reconnaître des stimuli correspond à un indicateur de fluidité. Autre exemple, Albrecht et Carbon (2014) ont exposés des participants à une série d'images, dont certaines avaient été rendues fluentes (méthode similaire que celle présentée en Figure 2) (Albrecht & Carbon, 2014). Les participants devaient rapidement juger si l'image était à l'envers ou correctement orientée en appuyant sur une touche du clavier. Le temps pris pour prendre cette décision est enregistré comme un indicateur de fluidité. De manière similaire, le temps d'entrer une évaluation de préférence (Vrana & Van den Bergh, 1995) a été utilisé comme mesure de fluidité de traitement.

Cependant, évaluer la fluence uniquement en fonction du temps de réponse présente des limites. Tout d'abord, cette méthode ignore l'écart entre la difficulté anticipée et la difficulté réellement rencontrée lors du traitement d'un stimulus. En d'autres termes, si une tâche est plus facile ou plus difficile que prévu, cette discordance ne serait pas prise en compte dans une mesure basée uniquement sur le temps de réponse. Certaines de ces études ont tenté de pallier ces limites en intégrant des évaluations subjectives. Des échelles **d'auto-évaluation** ont été utilisées pour mesurer divers aspects de la fluence, tels que la difficulté auto-rapportée (Gomez & Torelli, 2015), la complexité (Westerman et al., 2015), la facilité (Dragojevic & Giles, 2016), ou même le sentiment de fluence lui-même (Forster et al., 2015). En dépit de ces avancées, le manque de standardisation des termes entre les différents articles

créée un certain flou conceptuel. En effet, c'est ce que montre une revue (Graf et al., 2018) en recensant 120 articles consacrés à la quantification de la fluence subjective : plusieurs mesures existantes incluent des termes qui ne sont pas strictement associés à la définition la plus couramment acceptée de la fluence. Des mots comme « organisé » ou « complexe », qui se rapportent plutôt aux sources de la fluence, ou « attractif » et « accrocheur », qui sont plutôt des conséquences de celle-ci, sont parfois inclus. Ce travail est d'autant plus significatif qu'il est l'un des premiers à évaluer et à quantifier la cohérence dans les mesures de fluence subjective. L'étude démontre également que même une échelle simplifiée, allant de « difficile » à « facile », peut constituer un outil efficace pour évaluer différents aspects de la fluence.

Outre les approches subjectives, la recherche a également bénéficié de méthodes plus objectives plus sophistiquées comme l'Imagerie par Résonance Magnétique fonctionnelle (IRMf). Dans ce contexte, une étude a mesuré l'activité cérébrale des participants pendant qu'ils étaient exposés à des proverbes familiers et non familiers (Bohrn et al., 2012). Leurs résultats révèlent que les proverbes moins familiers suscitent une activité cérébrale accrue dans les zones liées à la compréhension de texte, suggérant un effort cognitif plus important par rapport aux proverbes bien connus. Des études physiologiques sur la fluence ont aussi apporté une pierre à l'édifice : Winkielman, Berntson et Cacioppo ont montré que les stimuli facilement traitables provoquent des réponses électromyographiques plus fortes dans les muscles de la joue (Winkielman et al., 2007), un indicateur de plaisir. Une autre étude indique également que les réponses électromyographiques sont d'autant plus fortes avec une exposition répétée à un stimulus, signe d'un traitement plus aisé (Harmon-Jones & Allen, 2001). Ces méthodes objectives ouvrent la voie à des mesures plus précises de la fluence, enrichissant ainsi notre compréhension au-delà des évaluations subjectives et autodéclarées.

En résumé, bien que des avancées notables aient été réalisées dans la quantification neuronale de la fluence, il subsiste des failles importantes. Les méthodes objectives indirectes couramment employées s'avèrent limitées, car elles se concentrent principalement sur les propriétés intrinsèques des stimuli, en supposant qu'elles sont façonnées par des préférences dictées par la fluence. Cette focalisation sur les stimuli omet d'intégrer l'interaction cruciale entre l'observateur et les stimuli, conformément à la théorie de la fluence. De plus, même si certaines de ces lacunes



peuvent être partiellement comblées par des évaluations subjectives, ces dernières ne sont pas sans défauts, en raison notamment de l'ambiguïté des termes employés et de leur inapplicabilité à d'autres espèces que la nôtre. Quant aux mesures de la fluence basées sur des réponses d'observateurs, obtenues par des techniques avancées d'analyse d'activité cérébrale, elles peuvent également être biaisées. En effet, ces mesures tendent à refléter une combinaison de la fluence et d'autres processus sensorimoteurs, ajoutant une couche de complexité à leur interprétation. Une approche pour surmonter certaines des limites mentionnées précédemment serait d'adopter une modélisation de la fluence centrée sur l'interaction entre le stimulus et l'observateur. En effet, l'encodage neuronal de l'information visuelle offre une description précise de cette interaction. Il ne s'agit pas seulement des caractéristiques intrinsèques des stimuli, mais aussi de la manière dont le système visuel les interprète. Dès lors, il est pertinent d'étudier la fluence à travers des modèles d'encodage de l'information ou, en d'autres termes, des modèles de perception.

## B. Modéliser la fluence : approche générale

La fluence, en tant que concept, est profondément enracinée dans l'interaction entre le stimulus et la perception de l'observateur. Par conséquent, une première étape cruciale pour modéliser la fluence est d'adopter un modèle de perception approprié qui reflète cette interaction. L'adoption d'un tel modèle permettra de comprendre comment différents stimuli sont traités par le système perceptif. Une fois ce modèle en place, la prochaine étape consisterait à définir une métrique précise qui évalue à quel point l'information est traitée facilement par la perception.

Il est crucial de souligner que cette thèse se focalisera principalement sur la perception visuelle de la fluence, bien qu'elle puisse être étudiée dans divers domaines sensoriels, et ce, pour plusieurs raisons. Premièrement, c'est dans le domaine visuel que la majorité des preuves expérimentales de la fluence ont été recueillies. Deuxièmement, c'est également dans ce domaine que les modèles de perception sont les plus développés, notamment grâce à l'avènement des Réseaux Neuronaux Convolutifs (CNN). Ces modèles, qui imitent la manière dont le cerveau traite l'information visuelle, offrent une opportunité unique de décomposer et de comprendre la fluence à un niveau granulaire.

## 1. CNNs comme modèle de perception

Les réseaux neuronaux convolutifs profonds, soit CNN (pour Convolutional Neural Network), s'inscrivent à l'intersection entre la neuroscience et l'intelligence artificielle. Largement utilisés comme outil pour des tâches de détection d'objets, de classification d'images, voire de traitement du langage naturel, ils s'avèrent également être un outil puissant pour modéliser la vision des primates. Historiquement, ces CNN sont nés de la cristallisation d'idées inspirées par la complexité du cerveau humain et rendues possibles grâce aux avancées en informatique.

Hubel et Wiesel ont marqué une étape cruciale dans notre compréhension du cortex visuel des mammifères (Hubel & Wiesel, 1960). À travers leurs études sur les chats, ils ont identifié deux catégories distinctes de cellules dans le cortex visuel primaire V1 (voir le schéma du traitement visuel par les différentes aires visuelles du cerveau en Figure 5). Les « cellules simples » se sont avérées sensibles à des barres lumineuses d'orientations précises, leur activité fluctuant en fonction de l'angle de ces barres. Les « cellules complexes », quant à elles, se sont montrées moins rigides dans leur réponse. Elles peuvent réagir à une orientation préférée, dans différentes positions spatiales. Hubel et Wiesel ont donc proposé que ces cellules complexes sont probablement alimentées par plusieurs cellules simples, présentant une même orientation préférée mais des localisations spatiales légèrement différentes.

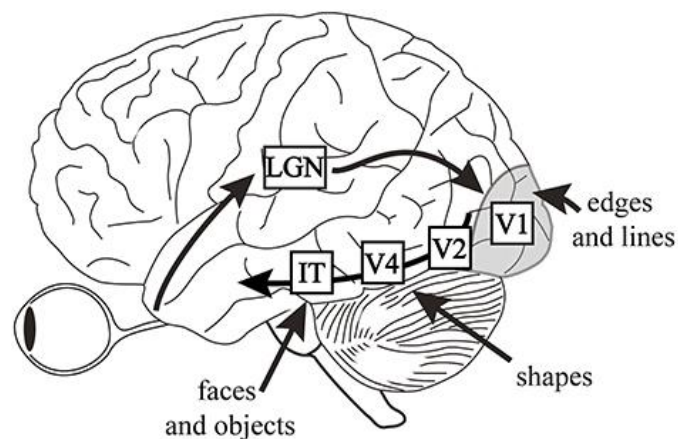


Figure 5 | Schéma de la chaîne ventrale du traitement de l'information visuelle dans le cerveau des mammifères. Cette information circule à travers différents étages du système visuel, la rétine, le corps géniculé latéral (LGN), les cortex primaires (V1), secondaires (V2), quaternaire (V4) et le lobe infero-temporal (IT), avant d'être traitée par le cortex pré-frontal (Felleman et Van Essen, 1991).

En 1980, Fukushima a formalisé ces observations biologiques dans un modèle qu'il a nommé le Neocognitron (Fukushima, 1980). Ce modèle informatique précurseur des CNN repose sur deux types de cellules artificielles, les S-cellules et les C-cellules, qui imitent les fonctions des cellules simples et complexes. Les S-cellules analysent des zones spécifiques d'une image grâce à une grille de poids 2D, tandis que les C-cellules apportent une réaction non-linéaire basée sur plusieurs S-cellules provenant de la même grille mais de positions différentes. Le modèle répète ces deux étapes de S-cellules et C-cellules à travers de multiples couches. Ainsi, il crée une structure hiérarchique qui ne se limite pas à la simulation du cortex visuel primaire (V1). Il englobe également les mécanismes plus avancés du trajet visuel ventral, offrant une représentation plus complète de la vision humaine. De plus, ce modèle montre que le réseau est « auto-organisé », en d'autres termes, les poids des cellules sont plastiques et s'ajustent suite à une exposition répétée à des images (non labellisées).

Dans les années 1990, ce modèle a ouvert la voie à d'autres architectures hiérarchiques, telles que le modèle HMAX. Ce dernier a introduit des opérations mathématiques plus simples, comme l'opération « max » (maximum d'une liste de valeur), pour calculer les réponses des C-cellules, ce qui a rendu le modèle plus robuste aux variations dans les images. Ces modèles ont permis de faire des comparaisons directes avec les performances humaines dans des expériences de psychophysique visuelle. Une corrélation a par exemple été établie entre ces modèles hiérarchiques et les premiers 100 à 150 ms du traitement visuel. Ils ont également été adaptés pour capturer les réponses des neurones V4 à des stimuli de formes complexes (Cadieu et al., 2007; Riesenhuber & Poggio, 1999; Serre et al., 2007).

L'héritage de ces modèles pionniers est encore vivace dans les modèles de CNNs modernes. LeCun, souvent considéré comme l'architecte des CNNs, a expressément cité le Neocognitron comme une source d'inspiration pour le développement de modèles plus complexes. Les CNNs comportent donc des opérations clés qui s'inspirent des travaux précédents. La première opération dite de convolution, commence par appliquer différents filtres à une image initiale, créant ce que l'on appelle des « cartes de caractéristiques » (*feature maps*) semblables aux grilles de S-cellules dans le Neocognitron (Figure 6). La deuxième opération de « max-pooling »

permet de résumer l'information de ces *features maps* (semblables à des cellules complexes (Figure 6)). Après plusieurs itérations de ces étapes qu'on peut qualifier de « features learning », soit d'apprentissage de caractéristiques, des couches entièrement connectées non-convolutives sont ajoutées. Si on considère que ce CNN doit effectuer une tâche de classification, la dernière couche du réseau contient autant d'unités qu'il y a de catégories à classer et détermine à quelle catégorie appartient l'image (schéma Figure 7). Cette synthèse offre une compréhension intuitive du mécanisme sous-jacent des CNNs (quelques détails sont apportés dans l'encadré technique 2).

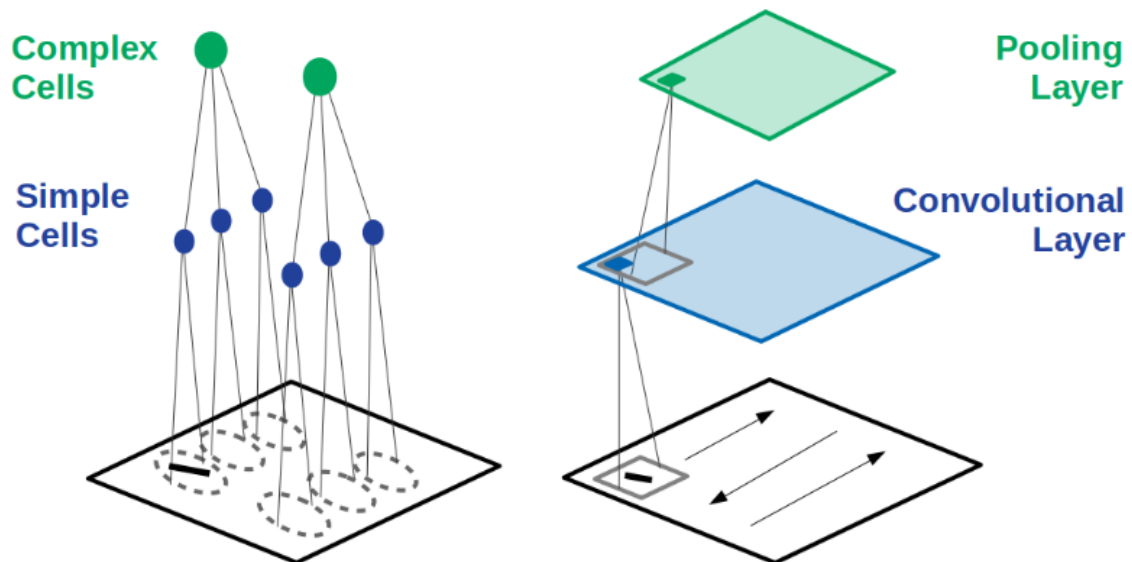


Figure 6 | La relation entre les composants du système visuel (schéma de gauche) et les opérations de base d'un réseau neuronal convolutif (schéma de droite). Hubel et Wiesel ont découvert que les cellules simples (à gauche, en bleu) ont des emplacements préférés dans l'image (ovales en pointillés) où elles réagissent le plus fortement à des barres d'orientation particulière. Les cellules complexes (en vert) reçoivent des entrées de nombreuses cellules simples et ont donc des réponses plus spatialement invariantes. Ces opérations sont reproduites dans un réseau neuronal convolutif (schéma de droite). La première couche convolutive (en bleu) est produite en appliquant une opération de convolution à l'image. Plus précisément, l'application d'un petit filtre (boîte grise) à chaque emplacement de l'image crée une carte de caractéristiques. Une couche convolutive a autant de cartes de caractéristiques que de filtres appliqués (une seule est présentée ici). Prendre l'activation maximale dans une petite section de chaque carte de caractéristiques (boîte grise) réduit la taille de l'image et conduit à des réponses semblables à celles des cellules complexes (en vert). Ceci est connu sous le nom d'opération de "max-pooling" et les cartes de caractéristiques réduites produites par celle-ci constituent une couche de pooling. Figure tirée de l'article de Grace W. Lindsay (2021).

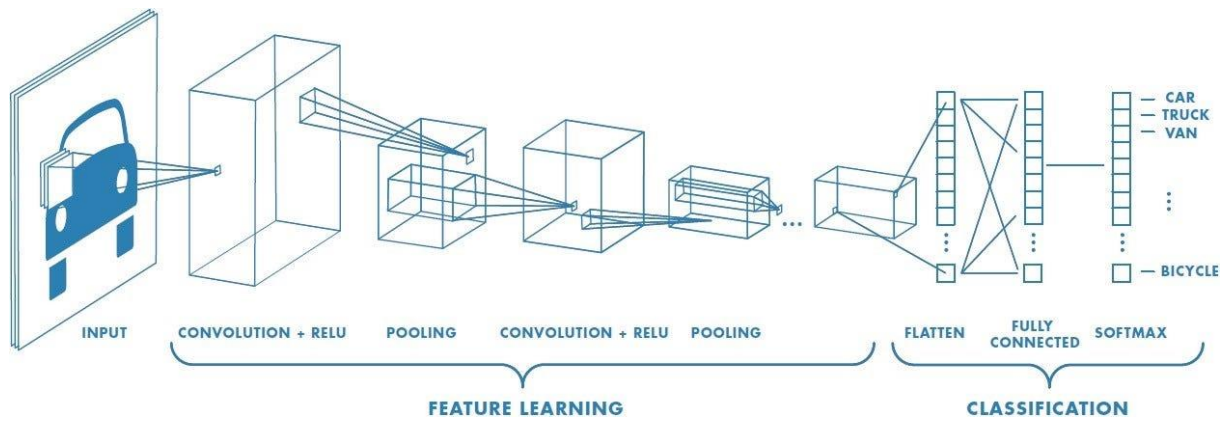


Figure 7 |Schéma d'un CNN qui doit deviner la catégorie de véhicule de l'image en entrée (ici, une voiture).(Prabhu, 2018)

La première démonstration majeure du pouvoir des CNN a eu lieu en 1989, lorsque LeCun et son équipe ont démontré qu'un petit CNN pouvait classer des chiffres écrits à la main (LeCun et al., 1989). Mais le véritable tournant pour les CNN a eu lieu en 2012 avec AlexNet, un réseau à 8 couches, qui a pulvérisé les records sur le défi ImageNet qui consiste à classer plus d'un million d'images issues du monde réel (*ImageNet large scale visual recognition competition 2012 (ILSVRC2012)*, 2012).

Depuis cette avancée majeure, le domaine des CNN a connu une effervescence scientifique, avec l'examen de diverses architectures et une multitude de paramètres ajustables. Ces derniers incluent la profondeur du réseau, l'agencement des couches de pooling, et le nombre de cartes de caractéristiques par couche, entre autres. L'ambition principale est d'améliorer les performances sur les jeu de données standards en vision par ordinateur. Paradoxalement, l'alignement avec les systèmes biologiques est devenu moins central dans ces investigations, rendant la recherche sur l'application des CNNs dans ce contexte de plus en plus spécialisée et marginale.

*À quel moment les CNN ont-ils été reconnus comme étant alignés avec le système visuel biologique et sur quelles preuves cette reconnaissance repose-t-elle ?*

Si l'on doit synthétiser en termes simples l'efficacité des réseaux de neurones convolutifs (CNN) pour modéliser le système visuel humain, comme l'évoque Lindsay, on mettrait en avant les trois aspects suivant : **(1)** une architecture hiérarchique semblable, **(2)** une activité neuronale corrélée et **(3)** des performances en tâches visuelles équivalentes à celles des humains (Lindsay, 2021).

Pour débiter avec la première caractéristique **(1)**, l'architecture hiérarchique des CNN trouve des échos directs dans la structure du système visuel humain. Cette structure permet aux CNN de traiter l'information visuelle de manière séquentielle et incrémentielle. Comme les signaux visuels humains sont d'abord traités par la rétine, les images introduites dans un CNN sont d'abord normalisées et décomposées en trois canaux de couleur, tels que le Rouge, le Vert et le Bleu (RGB). Ensuite, l'information sur les motifs de couleur passe successivement par le Corps Géniculé Latéral (NGL), les cortex visuels primaire (V1), secondaire (V2) et quaternaire (V4), le cortex temporal inférieur (IT), et finalement le cortex préfrontal qui relie la perception à la mémoire et aux actions (Felleman & Van Essen, 1991) (voir Figure 5). La représentation des objets émerge de manière hiérarchique à mesure que cette information circule, avec des caractéristiques visuelles simples (par exemple, l'orientation et le bord) codées dans le cortex visuel précoce (V1, V2), et des caractéristiques visuelles plus complexes codées dans l'IT comme la reconnaissance d'objets et de scènes (voir les deux études précédentes, mais aussi (Cadena et al., 2019; A. Dubey et al., 2016; Kriegeskorte, 2015)). Cette augmentation de la complexité des caractéristiques peut être directement observée grâce aux techniques de visualisation disponibles dans les CNN (voir Figure 8).

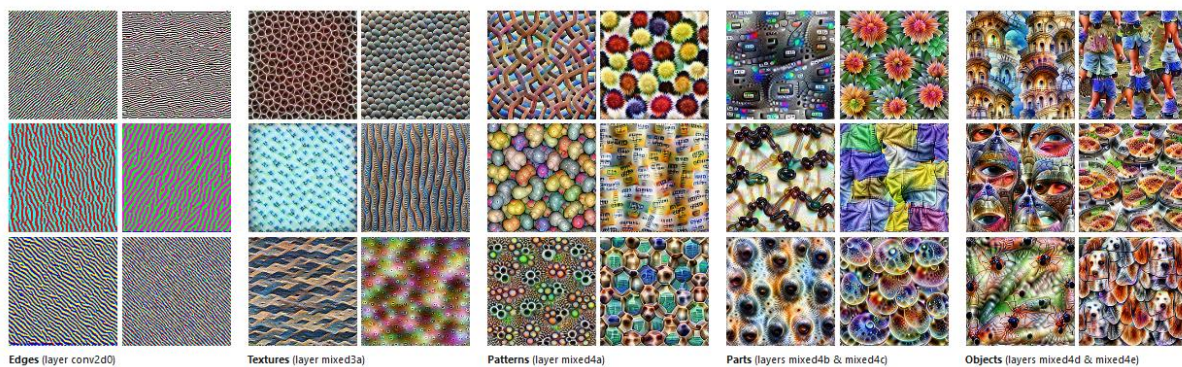


Figure 8 | Progression des caractéristiques (« features ») apprises par le réseau GoogLeNet. Les premières couches détectent des contours simples, suivis par la reconnaissance de textures et de motifs (patterns) de plus en plus complexes dans les couches intermédiaires, pour finalement aboutir à l'identification complète de l'objet dans les couches finales. Les images sont tirées d'un site interactif où des informations supplémentaires sont disponibles : <https://distill.pub/2017/feature-visualization/appendix/>.

Cette similitude avec la perception biologique est non seulement structurelle mais également fonctionnelle, ce qui nous mène au deuxième pilier en faveur de l'adéquation des CNN comme modèles du système visuel humain et de primates non-humains: la corrélation entre leur activité neuronale et celle observée dans les systèmes biologiques **(2)**. Un regain d'intérêt récent pour les réseaux de neurones artificiels parmi les neuroscientifiques est en grande partie dû à la découverte que ces réseaux pouvaient reproduire la représentation de l'information visuelle le long du flux visuel ventral. Plus précisément, lorsqu'on présente la même image à des CNN et à des animaux, l'activité des unités artificielles peut être utilisée pour prédire celle des neurones réels et ce, avec une précision supérieure à celle de toute autre méthode antérieure. Des études notables ont servi de catalyseurs dans ce rapprochement. L'étude de Yamins et al. (2014) a exploré différentes architectures de CNN pour évaluer leur capacité à prédire l'activité des neurones du cortex inférotemporal (IT) des primates non-humains, qui représente le dernier maillon de la chaîne dans la voie visuelle ventrale (Figure 5), à partir des modèles de régression linéaire (Yamins et al., 2014). Étonnamment, même sans avoir été calibrée spécifiquement pour les données neuronales, la dernière couche de sortie du modèle s'est avérée être un excellent prédicteur des réponses neuronales dans le cortex IT face à des images naturelles complexes. De plus, les dernières couches intermédiaires du réseau étaient efficaces pour prédire les réponses dans le cortex V4, une région visuelle intermédiaire qui alimente principalement le cortex IT (Figure 5). Plusieurs études (Cadena et al., 2019; Cichy et al., 2016; Güçlü & van Gerven, 2015; Seeliger et al., 2018; Seibert et al., 2016) ont utilisé différentes méthodes (l'IRM fonctionnelle et/ou magnétoencéphalographie) et diverses bases de données pour confirmer ces observations : les couches plus tardives du réseau prédisent mieux les régions plus élevées de la voie ventrale.

Il est intéressant de noter que cette corrélation ne se limite pas aux tâches visuelles basiques, mais s'étend également à des aspects plus sophistiqués de la cognition. On arrive donc à la troisième caractéristique **(3)** qui cimente le rôle des CNN comme modèles de perception : leur aptitude à accomplir des tâches visuelles à des niveaux de performance similaires à ceux des humains. Les CNN ont fait preuve de compétences extraordinaires dans une variété de tâches, allant de la classification d'images à la reconnaissance de scènes complexes, même en présence de transformations qui préservent l'identité (position, taille, point de vue et contexte

visuel). Le fait que la succession des représentations neuronales créées par le système visuel puisse être reproduite par les CNN suggère qu'ils effectuent le même processus de « démêlement » (*disentanglement* en anglais). C'est-à-dire que les deux systèmes prennent des représentations de différentes catégories d'objets qui sont indissociables au niveau de l'image/rétine et créent des représentations qui permettent une séparabilité linéaire. Ils sont même capable de prédire des réponses psychologiques dans des tâches comme l'évaluation de la similarité d'objets (Peterson et al., 2016), la mémorabilité (R. Dubey et al., 2015) et l'encombrement -une dégradation de la reconnaissance des objets causée par des objets environnants- (Volkov et al., 2017).

Pour conclure, les réseaux neuronaux convolutifs (CNN) se sont récemment imposés comme des modèles puissants du traitement de l'information visuelle, allant de l'extraction de caractéristiques de bas niveau à l'interprétation sémantique de haut niveau.

#### Encadré 2 | CNN

Les « Convolutional Neural Networks » (CNN) ou en « Deep Convolutional neural Networks » (DCNN) sont des architectures de réseaux neuronaux spécialement adaptées à l'analyse d'images. Ils sont particulièrement efficaces pour classer des images dans diverses catégories. Prenons un exemple simple d'un CNN qui a pour tâche de classer des véhicules comme "voiture", "camionnette", "vélo", etc. (Figure 7). Si vous présentez à ce réseau une image inédite d'une voiture, il sera capable, grâce à son entraînement sur des milliers d'images étiquetées, de la catégoriser correctement comme une "voiture". Cette performance est due à sa capacité d'extraire et de reconnaître des caractéristiques uniques de chaque type de véhicule lors de sa phase d'apprentissage.

#### - Structure générale d'un CNN-

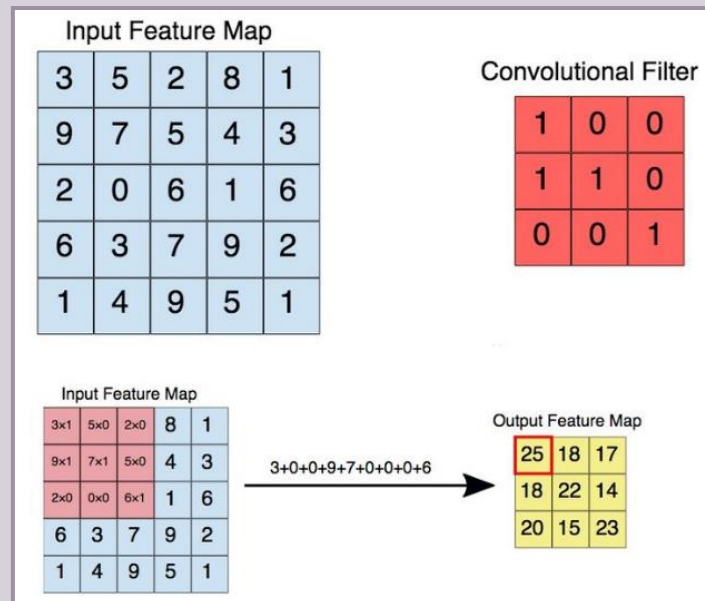
**Couche d'entrée :** Cette première couche reçoit les données brutes, souvent sous forme d'images.

**Couches Convolutives :** Le cœur d'un CNN est composé de couches convolutives. Ces couches utilisent des filtres 2D souvent de dimensions 3x3 ou 5x5, pour balayer l'image d'entrée. La convolution est une opération mathématique qui consiste à effectuer une multiplication élément par élément entre chaque filtre et la région de l'image sur laquelle il est



appliqué, suivie d'une sommation pour produire une valeur unique. Ce processus est répété sur toute l'image pour créer une carte des caractéristiques (ou carte d'activation).

$$\text{Valeur de l'activation} = \sum (\text{Poids du filtre} \times \text{Valeurs de la région de l'image couverte})$$



**Fonctions d'Activation :** Immédiatement après la convolution, une fonction d'activation est appliquée à chaque valeur dans la carte des caractéristiques. La fonction ReLU (Rectified Linear Unit) est la plus couramment utilisée, notamment parce qu'elle ajoute une non-linéarité en fixant toutes les valeurs négatives à zéro.

$$ReLU(x) = \max(0, x)$$

**Couches de Pooling :** Après une ou plusieurs couches de convolution et d'activation, viennent les couches de pooling. Ces couches réduisent la dimensionnalité des cartes de caractéristiques pour diminuer la complexité computationnelle du réseau et pour rendre le modèle plus généralisable. Le "max pooling" (prend la valeur maximale dans une fenêtre donnée) et l'"average pooling" (prend la valeur moyenne) sont les méthodes les plus couramment utilisées.

**Couches Complètement Connectées et couche de sortie :** Après plusieurs itérations de couches convolutives et de pooling, la dernière carte des caractéristiques est vectorisée (flattened). Ce vecteur est ensuite alimenté dans une ou plusieurs couches complètement connectées (Fully Connected). Ce sont ces couches qui effectuent la tâche finale, comme la classification. Des fonctions d'activation comme la softmax sont souvent utilisées ici, surtout pour les tâches de classification multi-classes.

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

**Profondeur du réseau :** La profondeur du réseau, c'est-à-dire le nombre de couches, peut varier considérablement selon la tâche et le modèle. Par exemple, VGG utilise des architectures avec 16 à 19 couches, tandis que ResNet peut aller jusqu'à 152 couches.

**En résumé :** Les CNN traitent des données en séquence, en passant par ces différentes couches. Ils transforment une entrée complexe, comme une image, en une représentation utile pour des tâches spécifiques comme la classification, la détection d'objets ou la segmentation. La combinaison de ces différentes couches et fonctions permet au CNN de capturer des caractéristiques de plus en plus abstraites et complexes à mesure que l'on avance dans les couches du réseau.

### **-Backpropagation et Apprentissage-**

**Fonction de Perte :** Dans l'apprentissage supervisé avec réseaux neuronaux, un concept clé est la "fonction de perte." Cette fonction quantifie à quel point les prédictions du modèle sont éloignées des valeurs réelles. Par exemple, si le réseau prédit que l'image d'un vélo est en réalité une voiture, la fonction de perte permettra de quantifier cette erreur.

Le choix de la fonction de perte dépend du problème à résoudre. Pour des problèmes de régression, l'Erreur Quadratique Moyenne (MSE) est souvent utilisée. Pour des problèmes de classification, c'est plutôt l'entropie croisée qui est privilégiée. L'objectif principal d'un réseau neuronal est de minimiser cette fonction de perte, ce qui équivaut à un problème d'optimisation où le réseau cherche le meilleur ensemble de poids.

**Règle d'Apprentissage :** Pour ajuster les paramètres (ou "poids") du réseau, on utilise généralement la règle suivante :

$$\text{Poids mis à jour} = \text{Poids initiaux} - \alpha \times \nabla J(\theta)$$

Avec  $\alpha$  le taux d'apprentissage, qui contrôle la taille du pas pendant l'optimisation et  $\nabla J(\theta)$  est le gradient de la fonction de perte par rapport aux poids  $\theta$ .

En utilisant cette règle, le réseau ajuste ses poids de manière à minimiser la fonction de perte.

**Descente de Gradient :** C'est ici que la rétropropagation entre en jeu. Cette technique calcule le gradient de la fonction de perte par rapport à chaque poids en partant de la sortie et en allant vers les couches d'entrée. En d'autres termes, elle "rétropropage" l'erreur à travers le réseau.

La descente de gradient, avec la rétropropagation, est la technique la plus couramment utilisée pour former les réseaux neuronaux. Cet algorithme itératif vise à trouver le minimum local d'une fonction, étant donné des paramètres spécifiques (c.-à-d., les poids initiaux). L'algorithme calcule la pente de la fonction de perte du réseau (c.-à-d. sa dérivée partielle) par rapport aux paramètres de poids à un point donné.

Vous pouvez imaginer la forme de la fonction de perte d'un réseau neuronal comme une montagne accidentée où votre objectif est de rejoindre la vallée. La descente de gradient estime la pente de la courbe entre le point de départ et le point cible pour indiquer la direction à suivre. Le taux d'apprentissage est un paramètre très important car il influence si l'optimisation réussit ou non.

**Autres Techniques :** Pour améliorer la performance et éviter le "surajustement" (où le modèle fonctionne bien sur les données d'apprentissage mais mal sur des données inconnues), on utilise souvent des techniques de régularisation comme le "dropout", qui éteint aléatoirement certaines connexions pendant l'apprentissage.

**Hyperparamètres :** Former un réseau neuronal est une tâche fastidieuse qui nécessite généralement l'ajustement fin de plusieurs hyperparamètres comme la taille du lot d'échantillons ("batch size"), le taux d'apprentissage, etc. Des modèles pré-entraînés sont souvent disponibles, ce qui peut grandement accélérer ce processus.

## 2. Métriques de la fluence : l'efficacité et l'efficacit 

L'inspiration premi re est tir e de la th orie de l'information de Shannon (Shannon, 1948). Cette th orie math matique distingue deux aspects cl s du traitement de l'information : l'**efficacit ** pour une transmission fiable et l'**efficacit ** pour une utilisation minimale des ressources (Hyvarinen, 2009). Ces concepts peuvent  tre utilis s pour mod liser la fluence (J. P. Renoult & Mendelson, 2019).

En neurosciences et en psychologie cognitive, l'efficacit  se rapporte   la pr cision et   la fiabilit  du cerveau dans le traitement de l'information. Il ne s'agit pas simplement d'une question de rapidit , mais  galement de la qualit  du r sultat obtenu. Par exemple, un syst me visuel efficace permet d'identifier rapidement et avec pr cision un visage dans une foule. Cette **efficacit ** peut  tre mesur e via divers indicateurs, tels que la vitesse de reconnaissance, le taux d'erreur ou l'activit  c r brale associ e   la t che (Owsley, 2013).

L'**efficacit ** concerne la capacit  du cerveau   accomplir des t ches de traitement de l'information tout en minimisant l'utilisation des ressources. Ce concept a  t  mis en lumi re par Barlow en 1961, qui a point  du doigt le co t m tabolique significatif de l'activit  neuronale, repr sentant entre 2,5 et 3,5% de la d pense  nerg tique d'un organisme au repos (Barlow, 1961). Selon Barlow, cette contrainte  nerg tique a orient  la s lection naturelle vers des syst mes neuronaux  conomiques en  nergie. Il a sugg r  que les neurones d di s   la perception et   la cognition sont s lectivement et parcimonieusement activ s. Appuyant cette notion, Olshausen a pr cis  en 2004 qu'un syst me est "parcimonieux" lorsque la majorit  des neurones restent au repos, laissant seuls quelques neurones hautement sp cialis s  tre intens ment activ s (Olshausen & Field, 2004). Cette s lectivit  neuronale illustre concr tement l'efficacit  dans le traitement de l'information par le cerveau.

L' quilibre entre efficacit  et efficacit  est un enjeu cl  dans le traitement de l'information par le cerveau. Ces deux dimensions sont souvent en concurrence : l'optimisation de l'une peut nuire   l'autre. Par exemple, augmenter la compression de donn es pour gagner en efficacit  peut introduire du bruit, compromettant ainsi l'efficacit  en termes de transmission d'information fiable. Il n'existe donc pas

d'optimum absolu : la meilleure stratégie dépend des objectifs spécifiques et du contexte dans lequel le traitement de l'information a lieu (J. P. Renoult & Mendelson, 2019).

L'efficacité et l'efficience jouent un rôle important dans nos préférences. Elles sont étroitement liées à la notion de **fluence**, qui désigne le degré de facilité ou de difficulté que nous ressentons lors de l'exécution de tâches mentales. Par exemple, notre penchant pour les formes symétriques s'explique par leur facilité de traitement cognitif (R. Reber & Schwarz, 2006). La symétrie, en nous permettant de déduire une moitié de l'objet à partir de l'autre, minimise l'effort mental requis, rendant ainsi le processus à la fois efficace et efficient. Notre attirance pour les prototypes est un exemple d'efficacité cognitive. Ces prototype « standards » nous aident à classer rapidement et précisément de nouveaux stimuli. Par exemple, une femme féminine sera vite associée au prototype « femme ». Ces prototypes agissent comme des "raccourcis" cognitifs, accélérant le traitement et la catégorisation de l'information tout en minimisant l'effort mental, ce qui les rend attractifs pour nous.

Pour évaluer à la fois l'efficacité et l'efficience dans le fonctionnement neuronal, des outils mathématiques tels que la sparsité peuvent être utilisés. La sparsité se réfère à la rareté de l'activation neuronale à un moment donné (*population sparseness*) et à la spécificité de la réponse d'un neurone sur sa durée de vie (*lifetime sparseness*). Plus précisément, la *population sparseness* mesure le nombre de neurones actifs à un instant précis, et un score faible indique une spécialisation neuronale, suggérant un traitement de l'information plus efficace. D'autres mesures, comme *l'activity sparseness* examinent la réactivité d'un neurone face à différents stimuli, tandis que la *temporal sparseness* étudie sa réactivité dans le temps. Ces mesures permettent de comprendre l'efficience avec laquelle le cerveau traite l'information, minimisant l'effort cognitif (pour plus de détails, voir : (Willmore & Tolhurst, 2001)).

Une étude a suggéré d'évaluer la fluence en analysant l'efficience de l'activation neuronale dans les réseaux neuronaux artificiels (Julien P. Renoult et al., 2016). Cette efficacité a été mesurée en utilisant le kurtosis pour évaluer la **sparsité** d'activation neuronale. Les auteurs ont utilisé un modèle du cortex visuel primaire humain (V1) et ont étudié la parcimonie d'encodage d'images de visages féminins par ce modèle. L'attractivité de ces visages a été évaluée par un panel masculin, révélant une

corrélation positive entre la sparsité neuronale et l'attractivité. Une autre étude a montré que la part de variance dans les évaluations d'attractivité expliquée par la sparsité dépassait celle d'autres facteurs comme l'indice de masse corporelle ou le dimorphisme sexuel (Holzleitner I.J., Lee A.J., Hahn A.C., Kandrik M., Bovet J., Renoult J., Simmons D., Garrod O., DeBruine L.M., Jones B.C., 2018). L'emploi de réseaux neuronaux profonds (CNN) pour évaluer la sparsité de l'activation neuronale à travers l'ensemble du circuit de traitement de l'information visuelle, depuis les récepteurs oculaires jusqu'aux niveaux supérieurs de perception, permettra une analyse plus complète de l'efficacité du traitement de l'information. Ceci, à son tour, nous aidera à comprendre plus précisément le rôle de la fluence dans la perception de l'attractivité faciale.

## V. Objectifs de la thèse

L'objectif centrale de cette thèse a été de modéliser la notion de beauté, en explorant sa manifestation au-delà de l'espèce humaine. Pour cela, je me penche sur l'impact de la prototypicalité, conceptualisée par le biais de la fluence, sur les comportements socio-sexuels des mandrills (*Mandrillus sphinx*), un primate non-humain vivant au Gabon. Mon travail s'ancre solidement dans la théorie de la fluence, en se concentrant spécifiquement sur une forme de fluence perceptive : la prototypicalité. Dans cette optique, la féminité est abordée comme une expression particulière de cette prototypicalité.

Dans le chapitre 1, je poserai les bases de mon étude en présentant le mandrill comme sujet d'observation privilégié ainsi que la base de données disponibles sur laquelle j'ai travaillé. Je présenterai les mandrills comme espèce animale, leur site d'étude et la population étudiée. Je fournirai un aperçu succinct des données comportementales recueillies et je détaillerai la base de données de photographies des visages, essentielle pour mesure sur la prototypicalité.

Dans le chapitre 2, je définirai une mesure de la féminité à travers le prisme de la fluence. La féminité des femelles mandrills sera définie en fonction de sa proximité avec le prototype central de la catégorie "femelle", en me basant sur la distance au centroïde. J'utiliserai notamment la dernière couche du réseau de neurones comme un espace d'encodage des visages. L'analyse débutera par une évaluation centrée sur

l'humain afin de valider cette approche de la féminité et de la généraliser à d'autres espèces. L'objectif est double : introduire la méthode de mesure de la féminité en utilisant un CNN et examiner comment cette représentation modélisée de la féminité des femelles mandrills influence leurs comportements socio-sexuels dans leur environnement naturel.

Dans le chapitre 3, j'explorerai la féminité chez les jeunes mâles mandrills, mais sous une perspective différente. Tout d'abord, je m'attacherai à comprendre "l'effet de féminisation" des jeunes mâles, avec la mesure de féminité établie dans le chapitre précédent. Tout d'abord, je me concentrerai sur les facteurs déterminants de la variation de la féminité chez les jeunes mâles, expliquée soit par l'effet de l'ordre de naissance fraternel, soit par l'hypothèse de la condition maternelle. Ensuite, je regarderai si la perception de la féminité chez les mâles provoque des divergences comportementales entre mâles et femelles juvéniles.

Dans le chapitre 4, l'objectif sera d'aller au-delà de la féminité comme simple illustration de la prototypicalité. L'approche méthodologique adoptée vise à concevoir une mesure généralisable de la prototypicalité en examinant chacune des couches d'un CNN, correspondant ainsi aux différentes étapes du traitement neuronal. Cette exploration visera à étudier la relation entre la prototypicalité et l'attractivité des visages humains, sans application concrète sur le comportement.

En somme, à travers cette thèse, je cherche à établir une passerelle entre les mécanismes de calcul neuronale et les comportements socio-sexuels observés en milieu naturel.

---

# CHAPITRE 1

---

---

Présentation de la population de mandrills étudiée et de la base de données disponible

---



© Berta Roura-Torres



# I. Présentation de l'espèce étudiée

## *Taxonomie et distribution spatiale*

Les mandrills appartiennent à la grande famille des Cercopithecidae (Figure 9). Cette famille comprend diverses espèces de babouins, macaques, mangabeys, drills et geladas (Kingdon et al., 2013) et est l'une des plus diverses parmi les primates, avec plus de 130 espèces réparties à travers l'Afrique et l'Asie. Historiquement, les mandrills étaient catégorisés comme des "babouins des forêts", principalement en raison de leur ressemblance apparente avec les babouins du genre *Papio*. Toutefois, les analyses génétiques contemporaines suggèrent une affiliation plus étroite avec les mangabeys du genre *Cercocebus* (Kingdon et al., 2013) (voir Figure 10).

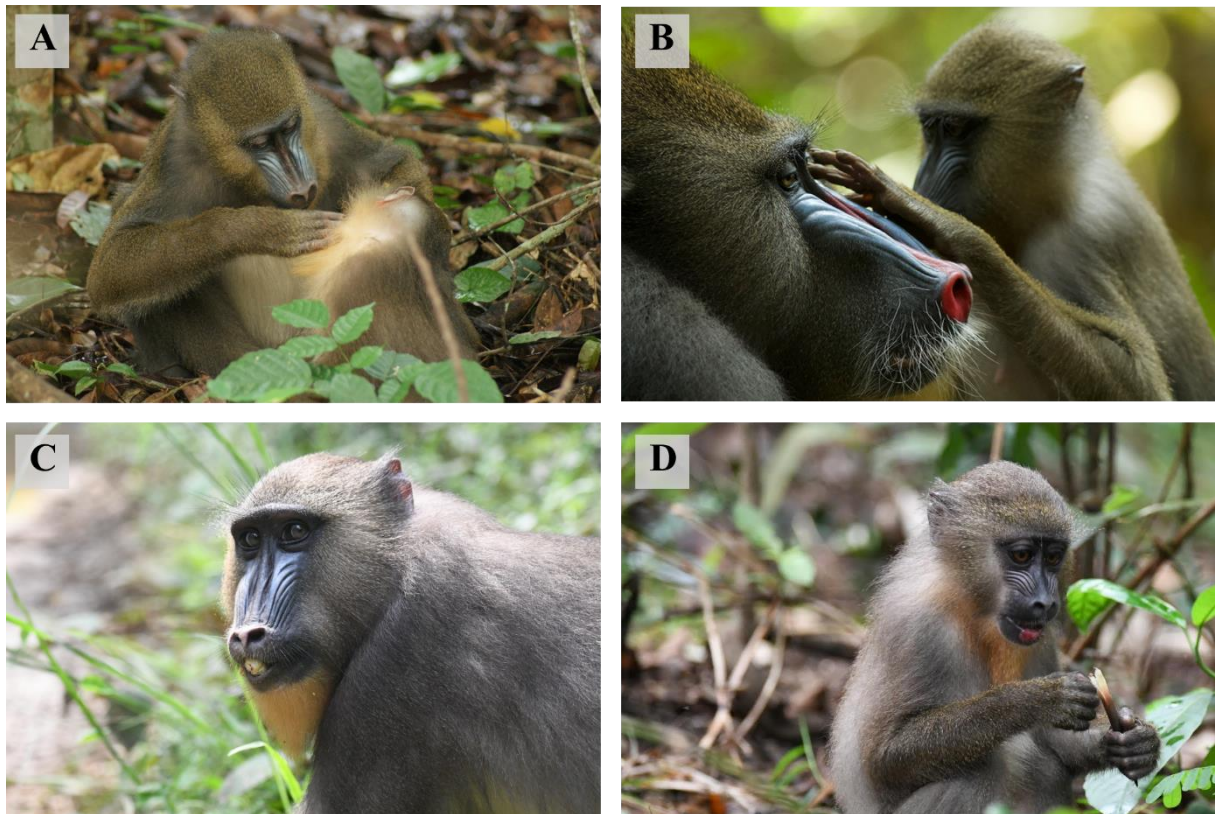


Figure 9 | Photographies de mandrills. (A) Un mâle adolescent toiletant un juvénile; (B) Une femelle adulte toiletant un mâle adulte ; (C) Mâle 178 -un des rares individus que j'arrive à reconnaître- (D) Juvénile. Photos © : Site Projet Mandrillus (A), Paul Amblard-Rambert (B) et les miennes lors de mes séances d'observations sur le terrain (C, D).

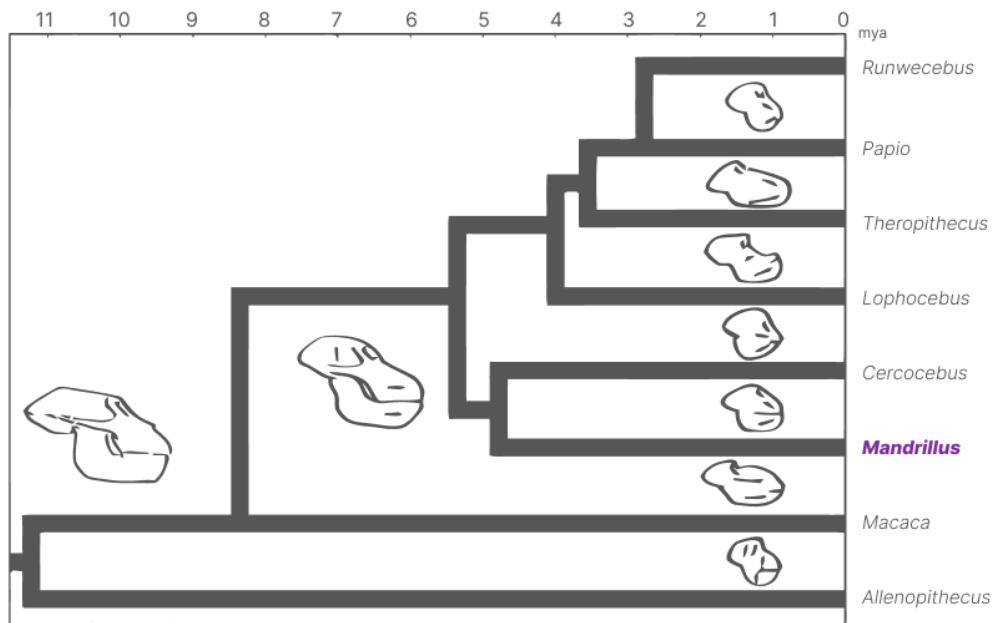


Figure 10 | Arbre phylogénétique de la famille des Cercopithecoidea (redessiné d'après (Kingdon et al., 2013) )

Le genre *Mandrillus* comprend seulement deux espèces (le drill : *M. leucophaeus* et le mandrill) et se retrouve uniquement en Afrique centrale, notamment comme au Cameroun et au Gabon. Ils habitent les forêts équatoriales humides et leur distribution spatiale reflète leur préférence pour ces habitats forestiers denses (Abernethy et al., 2002). Les mandrills sont principalement terrestres mais excellent également dans les arbres pour se reposer ou échapper aux prédateurs. Omnivores, ils consomment une grande variété d'aliments : fruits, fleurs, graines, feuilles, tubercules, insectes et petites proies (mammifères, oiseaux, reptiles). Cependant, ils ont une forte tendance frugivore (Nsi Akoue et al., 2017).



Figure 11. Crânes du mâle (a) et de la femelle (b) mandrills de la colonie CIRMF, illustrant le degré de dimorphisme sexuel en termes de taille. (Dirks et al., 2020)

### ***Caractéristiques biologiques, organisation sociale et système de reproduction***

Les mandrills forment de grands groupes multimâles-multifemelles comprenant plusieurs centaines d'individus (Abernethy et al., 2002). Au sein de ces groupes, ils adoptent un système d'accouplement polygynandrique, avec plusieurs mâles pouvant s'accoupler avec plusieurs femelles. Les femelles sont philopatriques et restent dans leur groupe natal toutes leurs vies, dans des familles matrilineaires (Brockmeyer et al., 2015). Les mâles, en revanche, dispersent souvent vers d'autres groupes lorsqu'ils atteignent la maturité sexuelle (Abernethy et al., 2002), bien que certains de ces mâles peuvent revenir dans leur groupe natal après quelques années (Charpentier In prep). La hiérarchie des mâles est déterminée par la compétition intra-sexuelle –et s'établit à l'issue de combats (parfois mortels), une situation qui conduit souvent à une instabilité de la hiérarchie (Marie Charpentier et al., 2005). En revanche, le rang de dominance des femelles est déterminée par leur appartenance à leur lignée maternelle : elles héritent de leur rang social de leur mère et forment ainsi des matrilineages relativement stables hiérarchiquement (J. Setchell et al., 2002).

La saison de reproduction des mandrills coïncide généralement avec la grande saison sèche (de juin à septembre). Pendant cette période, il est observé un afflux notable de mâles dans le groupe (Brockmeyer et al., 2015). La compétition entre les mâles devient alors féroce. Le mâle dominant (alpha) peut monopoliser 60 à 70 % des accouplements (Marie Charpentier et al., 2005; Mje Charpentier et al., 2020). Ces mâles dominants assurent la garde reproductive des femelles réceptives, reconnaissables à leurs gonflements péri-anal autour de la période d'ovulation (J. M. Setchell et al., 2005). Suite à cela, les naissances se produisent majoritairement entre décembre et janvier après une période de gestation environ 175 jours (Dezeure et al., 2022). La majorité des mâles quittent le groupe après la saison de reproduction alors que d'autres resteront plusieurs mois dans le groupe (Brockmeyer et al., 2015); Charpentier In Prep).

### ***Site et population étudiés***

En 2012, le Projet Mandrillus est créé par M. Charpentier. Il s'agit du seul suivi à long terme de mandrills sauvages habitués à l'Homme. La population étudiée évolue au Parc de la Lékédi et ses alentours, situé à 7 km au nord-ouest du village de Bakoumba dans la région du Haut-Ogooué au Sud du Gabon. Ce territoire où se mêlent forêts denses, prairies et savanes, épouse le rythme de quatre saisons écologiques distinctes: une longue saison de pluies de février à mai suivies d'un été sec jusqu'en septembre, une brève saison de pluies en Octobre et Novembre, puis une courte saison sèche en Décembre et Janvier.

L'histoire des mandrills étudiés débute au CIRMF (Centre International de Recherches Médicales de Franceville, Gabon), où 65 d'entre eux, alors semi-captifs, attendaient un destin différent. En 2002 et 2006, ils sont relâchés en deux vagues successives au parc de la Lékédi (Peignot et al., 2008). Les premiers mois de leur aventure sont accompagnés d'un rationnement en fruits locaux dispensés en faible quantité pour compléter leur régime alimentaire naturel. Cependant, cette supplémentation alimentaire a rapidement diminué entre 2008 et 2012, pour cesser complètement en avril 2012. Un an après le premier relâcher (en 2003), des mâles sauvages intègrent ce groupe et commencent à se reproduire avec les femelles issues des relâchés. À l'aube du Projet Mandrillus en 2012, ce groupe comptait environ 120 individus. Depuis leur nombre s'est accru, totalisant près de 350 individus en 2023,

parmi lesquels seulement 3-4 femelles adultes étaient issus de captivité, permettant ainsi l'habituation progressive des autres membres du groupe nés en milieu naturel (le mandrill est quasiment inhabituable en milieu naturel).

Grâce à l'expertise d'une équipe formée d'assistants de terrain, tous les individus, y compris les nouveau-nés et les juvéniles, sont désormais identifiables par des caractéristiques physiques distinctives. Ces assistants suivent quotidiennement le groupe tout au long de l'année, de l'aube au crépuscule. Toutefois, il arrive que les mandrills s'éclipsent dans la nature pendant quelques jours voire quelques semaines. Ma propre aventure s'est entremêlée à la leur durant la période tumultueuse du COVID (janvier 2022). J'ai prêté main-forte aux assistants (ou du moins, j'ai essayé, comme en témoigne la Figure 12) en essayant de les localiser après deux semaines de suivi interrompu (en raison de la quarantaine imposée à tous les assistants touchés par la COVID, pour la petite anecdote).

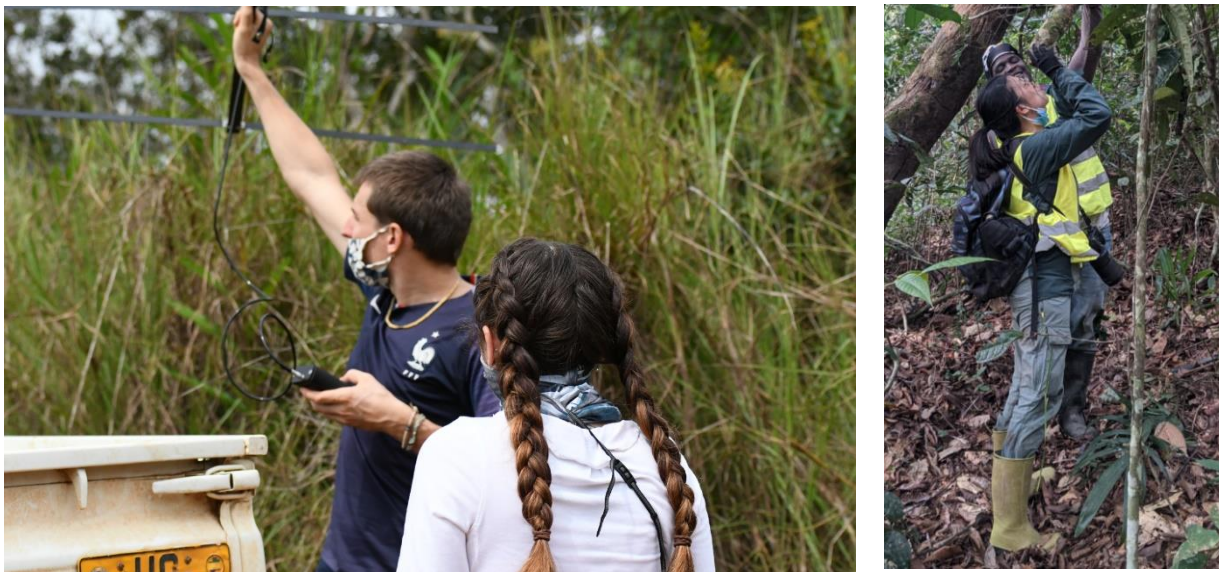


Figure 12 | (Gauche) Jules, le manager de l'équipe en 2022, avec son antenne réceptrice en main, cherche à localiser le groupe de mandrills égaré en traquant les fréquences des trois individus équipés de colliers émetteurs VHF. Marianne l'épaule dans cette mission. Et moi aussi en quelque sorte... (Droite) La boulette que je suis qui galère déjà avec l'eau, merci à Beбето pour ce système D (photo prise par Marianne)

## ***Collecte de données***

Des assistants de terrain formés aux différents protocoles du Projet se consacrent à l'observation détaillée des mandrills dans leur habitat naturel.

Données démographiques : Les observateurs réalisent des recensements quotidiens pour identifier chaque individu aperçu sur le terrain. Tout événement important, tel qu'une naissance, un décès ou une dispersion, est rigoureusement enregistré. L'état reproductif des femelles mandrills, qu'elles soient en phase de cycle, en gestation ou en allaitement, est également relevé quotidiennement. Les observations se basent en partie sur les gonflements sexuels, qui se manifestent lors de la réceptivité sexuelle des femelles. Ces gonflements augmentent en taille au cours de la première phase du cycle œstral, atteignent un pic autour de l'ovulation, puis diminuent jusqu'à disparaître lorsque les femelles ne sont plus réceptives. Une période prolongée sans gonflement sexuel peut indiquer une grossesse.

Données comportementales : En complément de ces données démographiques et reproductives, des données comportementales spécifiques sont collectées. Ces informations sont obtenues principalement par le biais d'observations ad libitum et de focales (Altmann, 1974). Lors de ces observations, tous les comportements des individus focals sont notés (comme le fourragement, les déplacements) incluant tous les comportements sociaux impliquant les membres du groupe, comme le toilettage, les comportements agressifs (comme les poursuites, morsures et frappes au sol), ainsi que divers types de jeux (notamment "course ludique", "jeu de bascule", "jeu de tête » etc...). Les copulations et autres comportements, comme les associations spatiales entre l'individu focal et les autres membres du groupe sont enregistrées (je ne liste pas exhaustivement tous ces comportements) sur des ordinateurs de poche. Ces données sont ensuite transférées et envoyées en France où elles nourrissent quotidiennement une base de données importante vérifiée, nettoyée et mise en forme par M. Charpentier. En 2023, environ 10,000h de focals collectés sur plus de 350 individus étaient disponibles.

En somme, cette méthodologie assure une vision complète et détaillée de la vie et des interactions des mandrills, permettant ainsi d'approfondir notre compréhension de cette espèce.

## ***Base de données des portraits de mandrills***

Au quotidien, les assistants sur le terrain capturent des images des mandrills en assurant une identification systématique des individus photographiés. Chaque photographie est annotée avec précision, contenant des informations telles que l'identité, le sexe, l'âge, la perspective du visage et la qualité de l'image. Suite à cela, un recadrage manuel est réalisé pour se concentrer sur le visage de l'individu photographié. Cet ensemble structuré de données forme la **Mandrillus Face Database**. Depuis le début de ce projet en 2012, la base de données a accumulé 29 495 portraits de 397 mandrills jusqu'en décembre 2023. Nous avons publié cette base de données pour un accès public et l'avons accompagnée du datapaper ci-dessous. J'ai moi-même participé aux prises photographiques, au recadrage et annotations de milliers de photos dans le cadre de cette thèse, et à la mise en place de script d'automatisation du processus.

# Manuscript 1

---

## *The Mandrillus Face Database: A portrait image database for individual and sex recognition, and age prediction in a non-human primate*

---

Sonia Tio<sup>a</sup>, Claudia Ximena Restrepo-Ortiz<sup>b</sup>, Berta Roura-Torres<sup>c,d</sup>, Loic Sauvade<sup>e</sup>, Mélanie Harté<sup>e</sup>, Marie J. E. Charpentier<sup>f,g</sup>, Julien P. Renoult<sup>a</sup>

<sup>a</sup>CEFE, University of Montpellier, CNRS, EPHE, IRD, Montpellier, France

<sup>b</sup>MARBEC, University of Montpellier, CNRS, Ifremer, IRD, Montpellier, France

<sup>c</sup>Behavioural Ecology & Sociobiology Unit, German Primate Center, Leibniz Institute of Primate Research, Kellnerweg 4, 37077 Gottingen, Germany.

<sup>d</sup>Department of Sociobiology/Anthropology, Johann-Friedrich-Blumenbach, Institute of Zoology and Anthropology, Georg-August University Gottingen, Kellnerweg 6, 37077 Gottingen, Germany.

<sup>e</sup>Projet Mandrillus, Fondation Lékédi Biodiversité, Bakoumba, Gabon

<sup>f</sup>Institut des Sciences de l'Evolution de Montpellier (ISEM), UMR5554 - University of Montpellier/CNRS/IRD/EPHE, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

<sup>g</sup>Department for the Ecology of Animal Societies, Max Planck Institute of Animal Behavior, Bücklestraße 5, Konstanz 78467, Germany

*Article accepted (Data in Brief)*

<https://doi.org/10.1016/j.dib.2023.108939>



## **Keywords**

Face recognition, artificial intelligence, deep learning, *Mandrillus sphinx*, primate

## **Abstract**

The Mandrillus Project is a long-term field research project in ecology and evolutionary biology, monitoring, since 2012, a natural population of mandrills (*Mandrillus sphinx*; primate) located in Southern Gabon. The Mandrillus Face Database was launched at the beginning of the project and now contains 29,495 photographic portraits collected on 397 individuals from this population, from birth to death for some of them. Portrait images have been obtained by manually processing images taken in the field with DSLR cameras: faces have been cropped to remove the ears and rotated to align the eyes horizontally. The database provides portrait images resized to 224X224 pixels associated with several manually annotated labels: individual identity, sex, age, face view, and image quality. Labels are stored within the image metadata and in a table accompanying the image database. This database will allow training and comparing methods on individual and sex recognition, and age prediction in a non-human animal.

## Specifications table

---

Subject	Ecology
Specific subject area	Ecology and animal behaviour.
Type of data	JPG Images and Table of Metadata (csv file)
How the data were acquired	The photographic portraits were acquired on wild mandrills daily monitored in their natural environment, using different models of digital SLR cameras equipped with a long-focal lens (varying from 70 to 500 mm depending on camera models and distance to the subjects).
Data format	Raw (JPG)
Description of data collection	The Mandrillus Face Database contains 29,495 portrait photos of wild mandrills taken between 2012 and 2021, stored in jpg format with a 224 × 224 pixel resolution. These individuals are daily monitored by field assistants, who take pictures and record individuals' identity directly in the field. Pictures are taken on both sexes of all ages (infants, juveniles, adolescents and adults). Data were normalized by manually cropping faces and rotating images to align the eyes along a horizontal line.
Data source location	· Institution: Centre d'Ecologie Fonctionnelle et Evolutive (CEFE), UMR5175, University of Montpellier/CNRS/IRD/EPHE, Montpellier, France Institut des Sciences de l'Evolution de Montpellier (ISEM), UMR5554, University of Montpellier/CNRS/IRD/EPHE, Montpellier, France · City/Town/Region: Lékédi Park, Bakoumba (GPS coordinates: - 1.7974667244585478, 13.020375049984194 - accuracy:10 km) Country: Gabon
Data accessibility	· Repository name: Mandrillus Face Database Data identification number: <a href="https://doi.org/10.5281/zenodo.7467318">10.5281/zenodo.7467318</a> · Direct URL to data: <a href="https://doi.org/10.5281/zenodo.7467318">https://doi.org/10.5281/zenodo.7467318</a>

---

## Value of the data

- This is the first public database that contains annotated face pictures of wild mandrills living in their natural environment. To our knowledge, this is also the largest photographic portrait database for non-human animals in the wild regarding the number of sampled individuals and the time frame (397 individuals totaling 29,495 pictures taken during 10 years)
- Pictures are labelled by experts in primatology for applications in Computer Vision, Machine Learning/Deep Learning, Data Science.
- The database is specifically designed to benchmark methods of individual recognition, face verification, sex recognition, and age prediction in a non-human primate

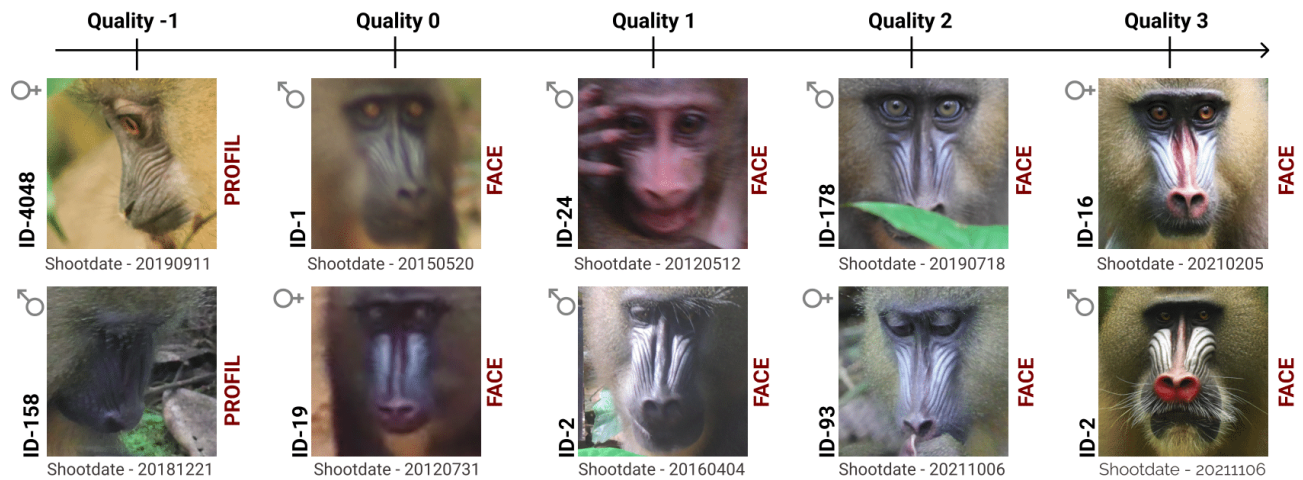
## **Objective**

This database is used to study the role of face attributes in visual communication in wild animals. Beyond applications in behavioural ecology, the database is also currently used to develop and compare Deep Learning methods of individual recognition, face verification, sex recognition, and age prediction in a non-human primate. The database also allows training Deep Learning models to automatically pre-process future data (automatic cropping, alignment, and labelling of face view and image quality).

## **Data description**

### **“Images” folder:**

The core dataset of the Mandrillus Face Database are photographic portraits in jpg format with a 224×224 pixel resolution, obtained following a processing step. These portraits are stored in a folder named “Images”. The folder contains subfolders, each subfolder including all images from one individual taken across the course of the study. Subfolders are named after individual identifiers. The image name follows the same syntax: date of shooting (YYYYMMDD) followed by ‘id’ and the individual identifier, and eventually the numbering of the pictures (whenever several pictures of the same individual were taken on the same day), with these three parts separated by an underscore. For instance, “20190313\_id170\_4.jpg” represents the 4<sup>th</sup> picture taken on an individual named 170 the 13<sup>th</sup> of March 2019. Each picture is annotated with labels stored in the XMP and IPTC metadata of the photo itself (see details of the labels below). Metadata includes information related to the individual (individual-specific metadata; e.g. identifier, sex, date of birth) or to the image (image-specific metadata; e.g. face view and image quality). See [Fig. 1](#) for a sample of images of different qualities.



**Fig1.** Sample of pictures collected on 10 individuals of both sexes and different ages, at different dates of shooting and of different quality and face view (see explanation of keywords below).

#### **MFD\_metadata.csv file:**

This metadata file is a .csv file with 29,496 rows (first row contains the name of each column; and one row per picture) and 9 columns (attributes of each picture).

Columns, arranged in that order in the file, contain the following information:

Photo Name (type 'string'): this column indicates the name of the picture (see above for the syntax). If the date of shooting is unknown, the date is "unknown" instead of "YYYYMMDD" format.

Id (type 'integer'): this column provides the identifier of the individual depicted in the picture.

Sex (type 'categorical'): this column indicates the sex of the individual on the picture ("f" for female, "m" for male and "unknown" if the sex was unknown).

dob (type 'date'): this column gives the date of birth of the individual on the picture (with "YYYYMMDD" format). If unknown, this cell returns "NaN".

dob\_estimated (type 'boolean'): this column indicates whether the date of birth is known with certainty ("False") or whether it has been estimated by the field assistants, based on observational data on the mother's ovulation cycle ("True"). If the date of birth is unknown, this cell returns "NaN".

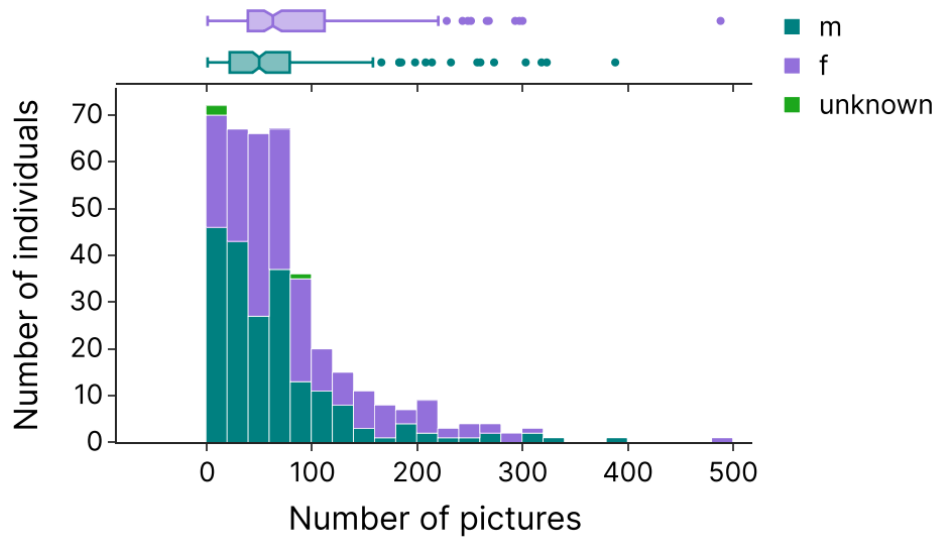
error\_dob (type 'integer'): if 'dob\_estimated = 'True', this column indicates the uncertainty (measured in days) around the date of birth. If 'dob\_estimated = 'False', this cell returns 0. If the dob is unknown, this cell returns "NaN".

FaceView (type 'integer'): this column indicates whether the mandrill's face depicted on the picture is in frontal (1) or in profile (0) view. The face is considered to be frontal when both eyes are visible and the face is fully frontal or on 3/4 (approximately  $<30^\circ$ ) and the occlusion covers less than 50% of the face, otherwise the face is considered to be in profile view. The database includes 26,846 frontal and 2,649 profile pictures.

FaceQual (type 'categorical'): this column indicates the quality of the picture, ranging from 0 to 3 (or -1: when the quality has not been evaluated because the individual is in profile view). 0: pictures of bad quality and for which experienced field assistants are unable to recognize the individual from the picture alone, without the contextual information. 1: pictures of average quality for which experienced field assistants are able to recognize the individual from the picture alone, with some difficulties but without any contextual information. 2: pictures of good quality for which individual recognition is straightforward but the portrait does not meet the criteria of quality 3. 3: pictures of high quality for which individual recognition is straightforward, and the face has a neutral expression and is in perfect frontal view, with no shadow, bright spot or partial occlusion ("id card-like" portraits). The majority of the portraits are of quality 2 and 3 (see [Fig. 4](#)).

Shootdate (type 'date'): date of shooting ("YYYYMMDD").

The database includes 191 females, 203 males and 3 individuals of unknown sex (infants aged less than a year, only). [Fig. 2](#) represents the number of males and females as a function of the number of pictures collected on them, and the number of pictures per sex. On average, females are represented by  $85 \pm 72$  pictures and males by  $65 \pm 64$  pictures (here average are distorted by extreme values: as the boxplot shows, there are several individuals represented by only one picture, and one individual with 488 pictures). [Fig. 2](#) also illustrates that the database contains numerous individuals with a few pictures, although 323 individuals are represented by more than 20 pictures.

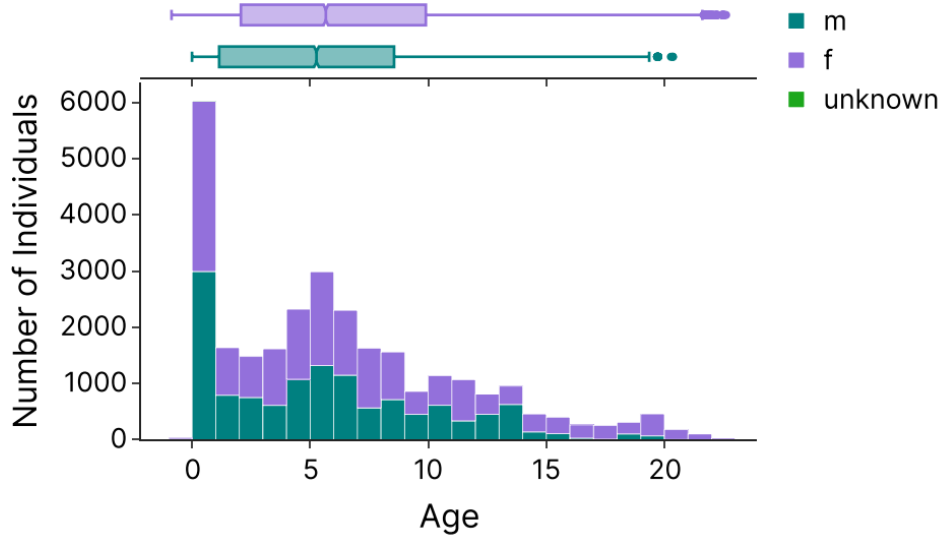


**Fig2.** Histogram of the number of individuals (coloured by sex) as a function of the number of pictures. The bin size represents 20 pictures. Marginal boxplots at the top display the minimum (e.g. 1 picture in females and males), first quartile (e.g. 39 pictures in females and 22 in males), median (e.g. 63 pictures in females and 50 in males), third quartile (e.g. 112 pictures in females and 79 pictures in males), and maximum (e.g. 488 in females and 388 in males) pictures for females (purple) and males (green).

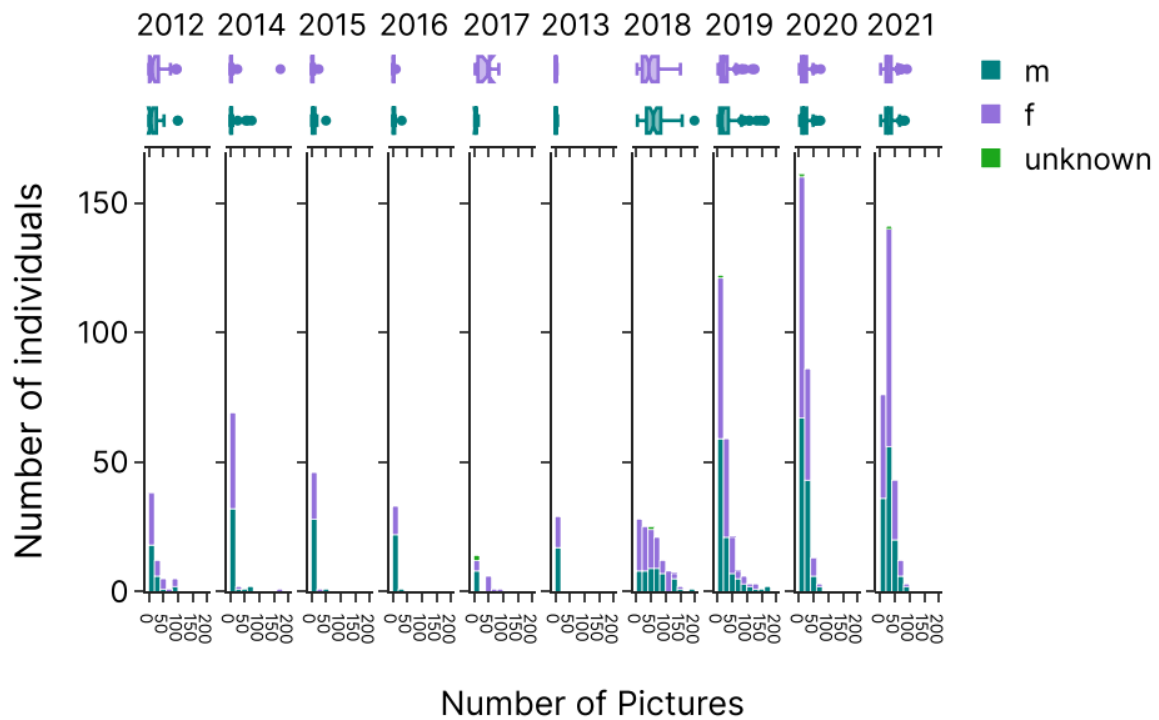
The database contains individuals from birth to 23 years old. The age is calculated as the difference between the shooting date ("Shootdate" column) and the date of birth ("dob" column). Fig. 3 provides the histogram of the age distribution of the portrayed individuals with infants (0-1 year) corresponding to 20% of the total number of pictures.

The Fig. 4 represents the number of pictures collected per year (extracted from the "ShootDate" column). Most pictures were taken from 2018 onward (more than 87% of the pictures).

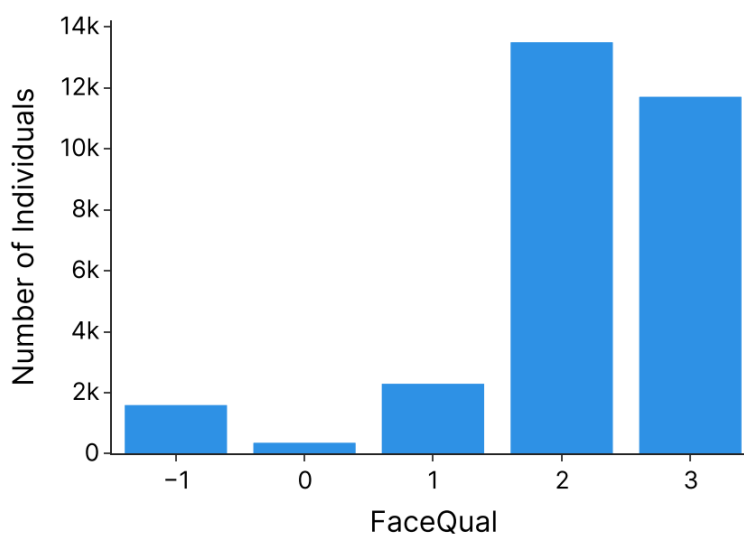
Finally, Fig. 5 represents the number of photos per quality score ("FaceQual" column). Most pictures (85%) are of quality 3 and 4.



**Fig3.** Histogram of the number of individuals (coloured by sex) as a function of their age. The bin size is 1 year old. Marginal boxplots of the ages at the top display the minimum (e.g. <1 year old in females and males), first quartile, median (e.g. 5.66 years old in females and 5.26 years old in males), third quartile, and maximum (e.g. 22.6 years old in females and 20 years old in males) ages for females (purple) and males (green).



**Fig4.** Yearly histograms of the number of individuals (coloured by sex) as a function of the number of pictures taken. The bin size represents 20 pictures. Marginal boxplots at the top display the minimum, first quartile, median, third quartile and maximum pictures for females (purple) and males (green), each year.



**Fig5.** Barplot of the number of pictures for each quality score.



## **Experimental design, materials and methods**

The database includes 29,495 photographic portraits collected on 397 individuals from the only wild social group of mandrills habituated to human presence. This group, which roams in the Lékédi Park and its surroundings, in southern Gabon (near to the village of Bakoumba), is daily monitored by the Mandrillus Project [1] (see for examples: [2,3]). The group was founded after the release of 65 semi-captive individuals (born and raised at CIRMF; Centre International de Recherches Médicales de Franceville, Gabon), in 2002 (36 individuals) and 2006 (29 individuals [4]). Starting as early as 2003, wild males joined the group to reproduce with released females. In 2021, most of the individuals of the group (>95%) were wild-born. Photos were taken directly in the forest by field assistants while following the study group. Since the beginning of the project, field assistants have used different models of DSLR cameras and long-focal lenses (varying from 70 to 500 mm, depending on camera models and distance to the subjects). Photos are uploaded on a computer regularly, and renamed by the assistants using the syntax presented above. Co-authors of this article (BRT, MH, MJEC, LS), who know the identity of the studied mandrills, monthly validated the names of the individuals depicted on all pictures from the database. Pictures were then processed using Adobe Photoshop Lightroom Software version 10.1.1. Images were first oriented to align the pupils of the eyes horizontally, and then centered and cropped to keep only the face (removing the neck and the ears). No further processing was applied.

## **Ethics statements**

The Mandrillus Face Database is based on non-invasive methods (pictures taken during the daily routine of the animals). Pictures were taken from a distance and without a flashlight. Photographers took pictures on the fly without any obvious perturbation of the study mandrills. These mandrills have been habituated to human presence since 2012. The Mandrillus Project and associated studies have been approved by an authorization from the CENAREST institute (permit number, AR017/22/MESRSTTCA/CENAREST/CG/CST/CSAR).

## **CRedit author statement**

Sonia Tieo: Data curation, Writing - Original draft.

Claudia Ximena Restrepo-Ortiz: Data curation.

Marie Charpentier: Data curation, Supervision, Reviewing and Editing.

Julien P. Renoult: Data curation, Supervision, Reviewing and Editing.

Berta Roura-Torres: Data curation, Photography.

Mélanie Harté: Data Curation.

Loic Sauvadet: Photography.

## **Acknowledgments**

We are grateful to past and present field assistants of the Mandrillus Project who took photographic portraits, and to the SODEPAL-COMILOG society (ERAMET group) for their long-term logistical support and contribution to the Mandrillus Project. This study was funded by several grants that allowed long-term collection of data: Deutsche Forschungsgemeinschaft (DFG, KA 1082-20-1), SEEG Lékédi (INEE-CNRS) to MJEC, and Agence Nationale de la Recherche to MJEC (ANR SLEEP 17-CE02-0002) and to JPR (ANR-20-CE02-0005-01). This is a Project Mandrillus publication number 31 and ISEM 2023-010-SUD.

## **Declaration of interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. Mandrillus Project. In: Mandrillus Project [Internet]. [cited 13 Dec 2022]. Available: <http://www.projetmandrillus.com/>.
2. Charpentier MJE, Harté M, Poirotte C, de Bellefon JM, Laubi B, Kappeler PM, et al. Same father, same face: Deep learning reveals selection for signaling kinship in a wild primate. *Sci Adv.* 2020;6: eaba3274.
3. Poirotte C, Massol F, Herbert A, Willaume E, Bomo PM, Kappeler PM, et al. Mandrills use olfaction to socially avoid parasitized conspecifics. *Sci Adv.* 2017;3: e1601721.
4. Peignot P, Charpentier MJE, Bout N, Bourry O, Massima U, Dosimont O, et al. Learning from the first release project of captive-bred mandrills *Mandrillus sphinx* in Gabon. *Oryx.* 2008;42: 122–131.

---

## CHAPITRE 2

---

---

### Conséquences sociales et sexuelles de la féminité faciale chez un primate non humain

---



© Berta Roura-Torres

# Manuscript 1

---

## *Social and sexual consequences of facial femininity in a non-human primate*

---

Sonia Tieo<sup>a,\*</sup>, Jules Dezeure<sup>b</sup>, Anna Cryer<sup>b</sup>, Pascal Lepou<sup>b</sup>, Marie J.E. Charpentier<sup>c,†</sup>, Julien P. Renoult<sup>a,†,\*</sup>

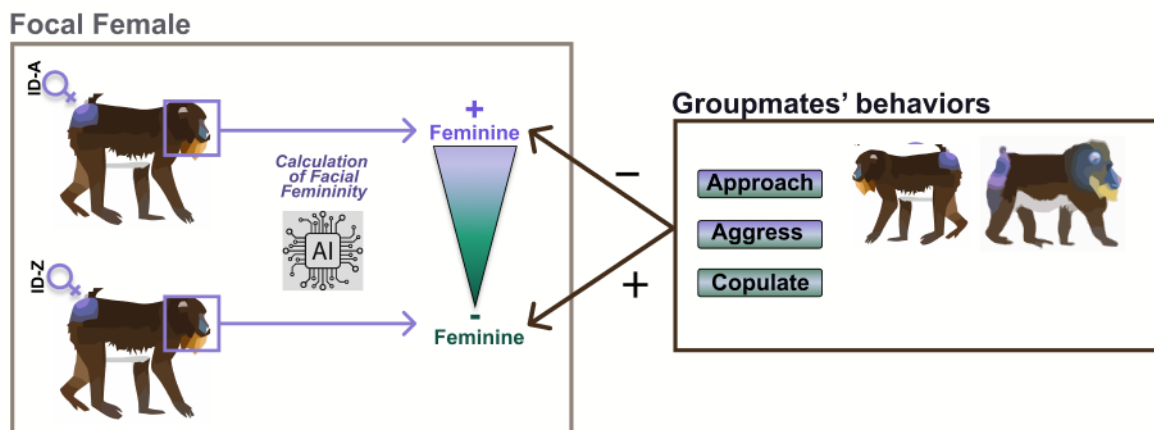
<sup>a</sup>CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

<sup>b</sup>Projet Mandrillus, Fondation Lékédi pour la Biodiversité, BP 52, Bakoumba, Gabon

<sup>c</sup>Institut des Sciences de l'Evolution de Montpellier (ISEM), UMR5554 - University of Montpellier/CNRS/IRD/EPHE, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

*Article accepted (iScience)*

<https://doi.org/10.1016/j.isci.2023.107901>



## **Subject areas**

Ethology, Biology of gender, Artificial intelligence

## **Summary**

In humans, femininity shapes women's interactions with both genders, but its influence on animals remains unknown. Using 10 years of data on a wild primate, we developed an artificial intelligence-based method to estimate facial femininity from naturalistic portraits. Our method explains up to 30% of the variance in perceived femininity in humans, competing with classical methods using standardized pictures taken under laboratory conditions. We then showed that femininity estimated on 95 female mandrills significantly correlated with various socio-sexual behaviors. Unexpectedly, less feminine female mandrills were approached and aggressed more frequently by both sexes and received more male copulations, suggesting a positive valuation of masculinity attributes rather than a perception bias. This study contributes to understand the role of femininity on animal's sociality and offers a framework for non-invasive research on visual communication in behavioral ecology.

## **Highlights**

- AI-based method to estimate facial femininity
- 10k female mandrill images of 95 adult female mandrills analyzed
- Less feminine mandrills attracted more social and sexual attention

## INTRODUCTION

In humans, an attractive face confers various benefits ranging from having more sexual partners<sup>1</sup> to being more successful at work,<sup>2</sup> earning more money,<sup>2,3</sup> and even receiving more lenient sentences in court.<sup>4</sup> In women, facial femininity,<sup>1-7</sup> the property of being representative of all female faces,<sup>5-7</sup> is an important feature influencing attractiveness. Feminine faces are generally preferred over masculine faces in mating contexts,<sup>5-7</sup> presumably because femininity is generally perceived as an indicator of fertility<sup>8</sup> or maternal competencies.<sup>9</sup> In various social contexts, both sexes also tend to prefer women with feminine traits, often associated with warmth, compared to masculine traits, generally associated with dominance and competitive tendencies.<sup>10-12</sup> The influence of facial femininity on conspecific behavior and socio-sexual interactions remains, however, largely unexplored in other animal species, making it unclear whether these effects are unique to humans.<sup>1</sup>

Previous studies on captive macaques have shown that individuals can assess the degree of sexual dimorphism, one proxy of femininity, in conspecifics' faces.<sup>13,14</sup> However, femininity is only one of the multiple components influencing conspecifics' behaviors. Indeed, both in humans<sup>15</sup> and animals,<sup>16</sup> behaviors emerge from multiple cognitive components, such as assessment (the quantification of a given value, e.g., attractiveness, associated with a stimulus), preference (the ranking of stimuli based on the assessment of different values), and choice (the selection of one stimulus based on multiple assessments, preferences, and other factors such as motor skills), which can be aligned or act in opposite directions.<sup>16</sup> For example, both homosexual and heterosexual men show opposite sexual preferences despite similarly assessing the attractiveness of women's faces,<sup>17</sup> and men's mate choice is modulated by contextual factors, such as the self-assessment of their own attractiveness (men who rate themselves as being more attractive generally choose more feminine women compared to men who rate themselves as being less attractive<sup>18,19</sup>). A large-scale cohort study found, moreover, that very attractive women (highest quartile) have more children than unattractive women, but fewer than moderately attractive women (second highest quartile),<sup>20</sup> illustrating the complexity of mate choice decisions. Last, femininity can

be evaluated differently in sexual contexts compared to social contexts. Although females' facial femininity seems consistently positively correlated with attractiveness in a sexual context,<sup>5,21</sup> in a social context, feminine women are often perceived as less competent, being possibly at a disadvantage compared to less feminine women.<sup>10,22,23</sup> The valence of femininity, corresponding to a positive or a negative assessment, depends thus on the nature of the interactions in humans. Whether or not such valence can also switch depending on the context in other animal species with a rich social life remains unknown.

These studies were, however, designed to in-lab assessment of human facial attractiveness and femininity, possibly explaining the discrepancies observed. They thus provide probably limited information about the impact of femininity on actual social relationships observed in the wild. It is now necessary to study animal behaviors expressed in their natural environments. The study of femininity in wild animals raises, however, significant methodological challenges. In humans, the classical approach for quantifying facial femininity is based on highly standardized photographic portraits on which landmarks are positioned to align faces.<sup>24,25</sup> Facial features are then extracted and used to build a face space, a multidimensional space that organizes variation among faces such that the distance between two faces describes their dissimilarity as perceived by humans.<sup>26,27</sup> In a face space, the axis connecting the centroid of female faces to the centroid of male faces represents the axis of variation in femininity.<sup>5,28</sup> In wild animals, obtaining standardized high-quality pictures is challenging. While captured and anesthetized individuals often show facial expressions that differ from their natural expressions, portrait images of freely moving animals are not standardized, showing high within-individual variation in pose, illumination, and facial expression that prevents the use of any landmark-based quantification methods. Fortunately, artificial intelligence (AI) now allows us to overcome these limitations. In particular, deep convolutional neural networks (DCNNs), the most widely used tool of AI in computer vision, automatically learn abstract features (e.g., colors, edges, simple and more complex textures) and concepts of interest (e.g., objects or parts of objects) when trained to optimally perform a specific task (e.g., individual recognition) from raw image data. Crucially, during training, DCNNs also learn to be insensitive to variations that are non-informative for that task (e.g., differences in illumination or distance to the camera in the case of individual



recognition).<sup>29,30</sup> Overall, DCNNs are powerful tools to build a face space in which individuals' position varies only with biologically relevant variables, such as identity, sex, or age.<sup>29,31</sup>

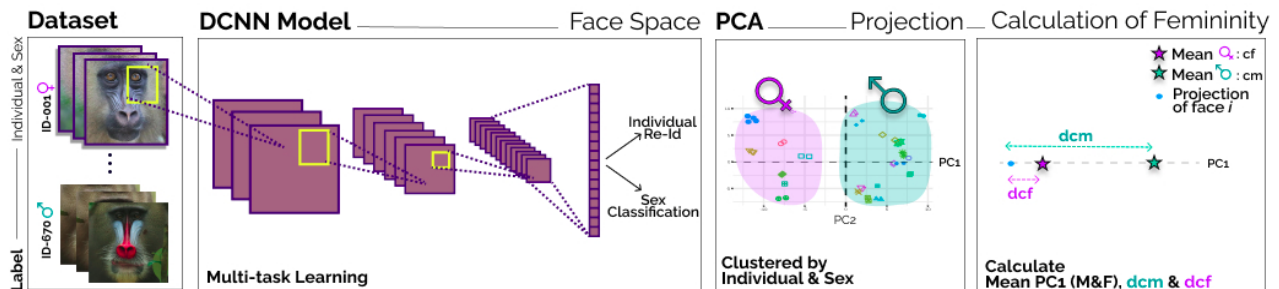


Figure 1. Pipeline to measure facial femininity in humans and mandrills

DCNN model (VGGFace) was first retrained with FaceScrub database<sup>37</sup> for humans and the Mandrillus Face Database<sup>40</sup> for mandrills to simultaneously perform identity verification and sex classification (i.e., multi-task learning). Portrait pictures of the studied individuals were then fed to the newly trained DCNN, and their features extracted. The 128-d penultimate layer of the DCNN was used as a face space. Images of CFD database (for analyses on humans) and Mandrillus Face Database restricted to adults (for analyses on mandrills) were projected into their respective face space. Retrieved features were reduced further to two dimensions using a PCA. The third box (from the left) represents the position of different pictures of different individuals in the PC1-PC2 plane (for clarity, only twenty individuals are presented; one shape of a given color corresponds to one individual). While PC1 and PC2 together allowed us to cluster portraits collected on the same individual, PC1 alone discriminated the two sexes. We thus defined the female centroid, cf, and the male centroid, cm, as the mean PC1 score of all female and male pictures, respectively. We then calculated two measures of femininity for a given portrait picture  $i$ : (1) dcf, the distance between  $PC1_i$  and cf, which describes the averageness of a face within the female faces' category, and (2) dcm, the distance between  $PC1_i$  and cm, which describes the dimorphism of females' faces compared to males' faces. In this depiction, a highly feminine face  $i$  is described by a high value of  $dcm_i$  and a small value of dcf

Here, we use DCNNs and long-term behavioral observations to investigate how facial femininity of female mandrills (*Mandrillus sphinx*), a primate from Central Africa, influences socio-sexual interactions in a large natural population of 250 wild individuals, monitored daily since 2012. Mandrills are a good study system for this topic because they are a highly dimorphic primate, with males being about 3.5 times heavier than females and exhibiting bright red and blue nasal and genital colorations mostly absent in females.<sup>32</sup> Mandrills live in large multimale-multifemale social groups where both sexes are sexually promiscuous and

where females form well-differentiated social interactions, with kinship and dominance ties having strong impacts on social partner choice. Mandrills are also seasonal breeders,<sup>33</sup> and females may overlap in their fertility period,<sup>34</sup> resulting in occasional male-mate choice. Indeed, during the mating season, males—usually the alpha male—mateguard cycling females, favoring high-ranking ones when several females are receptive at the same time.<sup>34</sup>

In this study, we first developed an AI-based measure of femininity applicable to field pictures. We benchmarked the method in humans using both standardized and non-standardized photographic portraits scored for femininity and attractiveness by human judges. We then applied this method to 13,000 photographic portraits of adult female mandrills collected in their natural environment during daily monitoring, to obtain an annual femininity score—quantified in two different ways—for each studied female. Finally, we assessed how females' femininity scores influence socio-sexual interactions with their groupmates (male copulations, intra- and inter-sexual spatial proximity, aggression, and grooming), using a large behavioral database collected over 10 years on 39 adult males and 95 adult females.

## **RESULTS**

### **AI model of facial femininity in women**

We first developed a predictive model of facial femininity in women. We used publicly available datasets of human faces for which an empirical measure of femininity was obtained through psychological questionnaires. We then identified the measure of femininity obtained through DCNNs that best correlates with perceived femininity, that is, the score of femininity assessed by judges ([Figure 1](#)).

### **Learning a face space**

Estimating femininity requires first to build a face space. DCNNs can yield biologically realistic face spaces when trained to recognize individuals. Here, we used VGGFace,<sup>35</sup> a DCNN<sup>36</sup> previously trained to individually recognize more than two thousands celebrities from 2.6 million face images, as a base model for all analyses, for both humans and mandrills. Because VGGFace needs to be retrained with a mandrill portrait dataset to learn a mandrill

face space (see in the following), for a comparable evaluation of model performance, we built a human face space by retraining VGGFace with another human face dataset (FaceScrub dataset<sup>37</sup>). Indeed, the quality of the retraining can be influenced by the size of the new dataset: we thus subsampled the FaceScrub dataset to keep the same number of individuals and mean number of images per individual as in our mandrill image dataset. Following previous studies,<sup>27,29</sup> we used the penultimate layer of the retrained model as the face space. Indeed, with a DCNN trained to recognize individuals from their face, this layer encodes information about identity but also other characteristics of individuals, such as their sex and ethnic origin,<sup>29,31</sup> while being relatively invariant to illumination, pose, and facial expression.<sup>29,30</sup> We first retrained VGGFace to recognize individuals of the FaceScrub dataset using a classification task. The model achieved good performance (mean average precision mAP@3 accuracy: 97%); however, when examining the distribution of the faces within the face space, we noticed that female and male faces were largely overlapping (see [Figure S1B](#) in Supplementary Information). This contrasts with results in psychology showing that both sexes are likely represented in separate regions within humans' perceptual face space.<sup>29,30</sup> When training to recognize sex, faces collected on the same individual did not cluster together ([Figure S1A](#)), again challenging empirical findings in psychology.<sup>30</sup> In order to build a face space that simultaneously separates sex and groups together portraits from the same individual (see [Figure S1C](#)), we thus retrained VGGFace to simultaneously recognize sex and verify individual identity. Face verification, the task of identifying whether two images depict the same individual or not, has been shown to better group together portraits collected on the same individual compared to a classification-based individual recognition.<sup>38</sup> Our final model was able to recognize sex and verify identity with an accuracy of 99% and 91%, respectively.

### **Measures of femininity in the face space**

Studies on human faces traditionally investigate two measures of femininity: the femaleness (sexual dimorphism between male and female faces<sup>39</sup>) and the averageness within the female faces' category.<sup>28</sup> We evaluated the ability of both measures to predict facial femininity using the publicly available Chicago Face Dataset (CFD hereafter). We first projected every image of CFD onto our human face space and then reduced the dimensionality of this face space using a principal-component analysis (PCA). The first principal component (PC1) explained

a relatively small fraction of the variance in features of the face space (15% for CFD); however, alone, it allowed us to separate the sexes with virtually no overlap (See [Figure S2](#)). Each face  $i$  was thus described by its  $PC_{1i}$  score. The female centroid,  $cf$ , and male centroid,  $cm$ , were computed as the mean of  $PC_{1i}$  score of all female and all male faces, respectively. We defined “averageness” within the female face category as  $dcf$ , the distance between  $PC_{1i}$  and  $cf$  and “femaleness” as  $dcm$ , the distance between  $PC_{1i}$  and  $cm$ , (see [Figure 1](#)). A high degree of averageness (low  $dcf$ ) describes faces that are typical of the female category. We thus expect a negative correlation between  $dcf$  and perceived femininity; that is, a greater divergence from the average female face (increased  $dcf$ ) is associated with a decreased perceived femininity. In contrast, faces exhibiting a higher degree of femaleness (high  $dcm$ ) are perceived as sexually more dimorphic; we thus predict a positive correlation between  $dcm$  and perceived femininity.

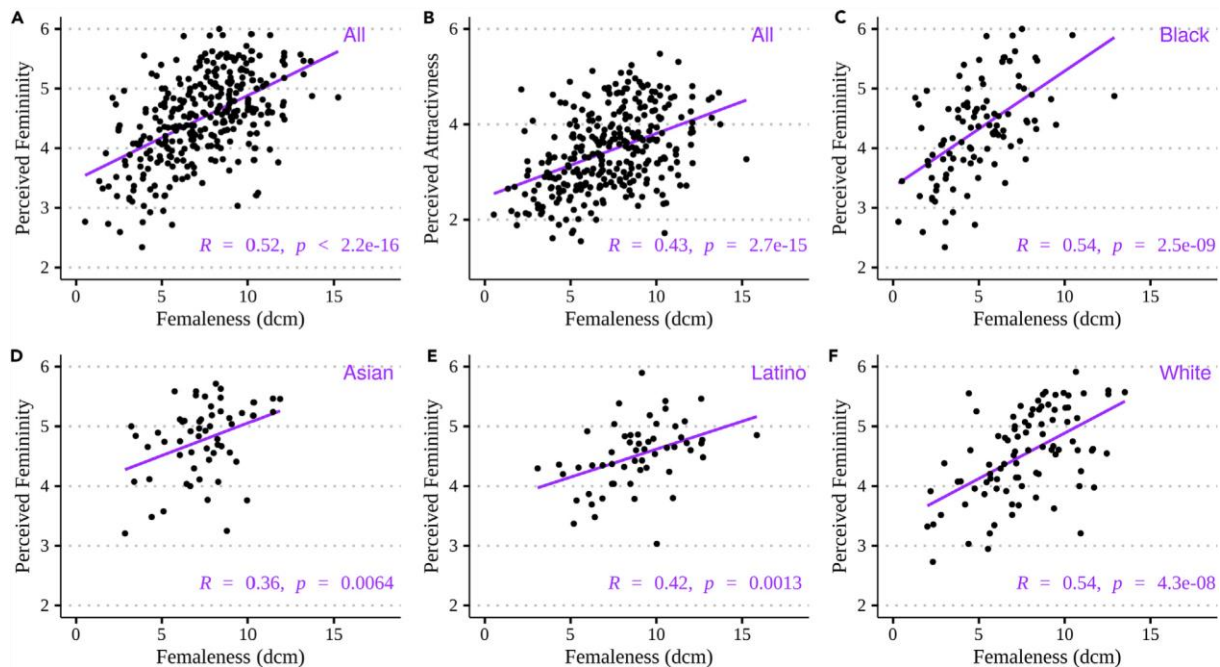
In accordance with our expectations, we found that  $dcm$  was positively ( $R = 0.52$ ,  $p = 2.2e-16$ ) and  $dcf$  was negatively ( $R = 0.16$ ,  $p = 0.005$ ) correlated with perceived femininity. Femaleness ( $dcm$ ) was noticeably more strongly correlated with the perceived femininity of CFD than was averageness ( $dcf$ ; [Figure 2A](#)). Because CFD includes faces of individuals from different ethnic backgrounds, we further calculated the correlation for each ethnic group. Across all ethnicities, we found, again, the highest correlations with  $dcm$  ([Table S1](#)) rather than  $dcf$ . The correlation was the highest for Black ( $R = 0.54$ ,  $p = 2.5e-9$ ; [Figure 2C](#)) and White ( $R = 0.54$ ,  $p = 4.3e-8$ ; [Figure 2F](#)) people, and the lowest for Asian people ( $R = 0.36$ ,  $p = 0.0064$ ; [Figure 2D](#); see details in [Table S1](#)).

Most studies on facial femininity investigate the correlation between predicted femininity and perceived attractiveness (i.e., attractiveness rated by people), rather than perceived femininity. For example, Ryali and collaborators<sup>41</sup> recently analyzed the CFD dataset using the widely used active appearance model<sup>42</sup> of face space. They predicted femininity as statistical typicality (log likelihood of being a woman), a measure close (but inversely related) to our measure of averageness ( $dcf$ ) when the probability density function is unimodal (as observed with CFD’s female face<sup>41</sup>), and they correlated their predicted femininity with the perceived attractiveness. They found a significant, positive correlation ( $R = 0.19$ ) between the two measures, close in magnitude to the correlation between  $dcf$  and

perceived femininity we found ( $R = 0.16$ ). For a more accurate comparison, we also calculated the correlation between our predictions of femininity (dcf and dcm) and perceived attractiveness. As expected, with dcf, the correlation is significant ( $R = 0.14$ ,  $p = 0.038$ ) but slightly weaker than that with perceived femininity. With dcm, the correlation with perceived attractiveness is markedly stronger ( $R = 0.43$ ,  $p = 2.7e-15$ ) compared to what was previously reported.<sup>41</sup>

Last, we evaluated the performance of our method with non-standardized portrait pictures, as found in the mandrill portrait dataset but not in the CFD. To our knowledge, there is no publicly available dataset of non-standardized portraits associated with empirical scores of perceived femininity. We thus used SCUT-FBP5500,<sup>43</sup> a dataset of non-standardized portraits retrieved from the Internet that display American and Asian people with faces of varying poses, expressions, and haircuts, with different illuminations and portrait backgrounds, and labeled for their perceived attractiveness. We found similar correlations between dcm and perceived attractiveness using either SCUTFBP5500 (Asian:  $R = 0.33$ ,  $p = 2.1e-51$ ; Caucasian:  $R = 0.49$ ,  $p = 1.1e-45$ ) or CFD (Asian:  $R = 0.34$ ,  $p = 0.0087$ ; Caucasian:  $R = 0.47$ ,  $p = 2.6e-6$ ; see [Tables S1](#) and [S2](#)).

Here, we thus showed that femaleness, i.e., the dimorphism of female faces compared to male faces, is a better predictor of femininity than the averageness within the category of female faces. We further showed that a face space automatically built from a DCNN trained to verify individual identity and discriminate sexes yields predictions of femininity with similar or higher performances compared to another, state-of-the-art face space. Crucially, however, with the DCNN-based face space, predictions are similarly good whether using in-lab, highly standardized portrait pictures, or non-standardized portraits of people in natural postures. Despite variations in performance, the method's ability to explain part of femininity in all ethnic groups encourages us to apply it to non-standardized portraits of mandrills.



**Figure 2. Correlation between predicted and perceived femininity/attractiveness in women**

(A–F) Results are shown for the Chicago Face Database, with predicted femininity measured as femaleness (dcm). Correlation with perceived femininity (A) and perceived attractiveness (B) for all 307 women, with perceived femininity for 104 Black (C), 57 Asian (D), 56 Latino, and 90 White females (F).  $R$  indicates the coefficient of correlation (Pearson) between predicted and perceived values, and  $p$  indicates the  $p$  value.

### Impact of femininity on mandrills' socio-sexual interactions

We estimated the femininity of adult females' faces in mandrills, using the same methodology as developed in humans. We first learned a mandrill face space by retraining VGGFace to simultaneously classify the sex (accuracy after training: 99%) and verify the identity (accuracy after training: 60%) with a dataset of ca. 13,000 photographic portraits collected on 169 wild individuals. We then projected the images of all adult males and females into the face space, reduced the dimensionality of their features using a PCA, and eventually generated dcm values (i.e., the distance of a female face to the mean score of males along PC1; [Figure 1](#)) for each female as a measure of their femininity.

Contrary to the human datasets (CFD and SCUT-FBP5500), most study mandrills are represented by several pictures taken at different ages and in different conditions (maximum time lapse between two pictures of the same individual: 10 years), with a different number of

pictures collected per individual (Figure S5). However, we found that the inter-individual variance in dcm (averaged over all pictures of a given individual) is 33.7, while the mean intra-individual variance is 1.4 (thus 23 times lower). In addition, the mean and variance in femininity scores are not correlated to the number of pictures considered (see STAR Methods and Figure S6). These results suggest that an estimate of femininity could be reliably obtained by simply averaging all images of a given individual. Nevertheless, in humans, the perception of femininity of a woman is influenced by age<sup>44</sup> and possibly the menstrual cycle (45,46 but see Marcinkowska et al.<sup>47</sup>). We thus first investigated whether females' age, dominance rank (high, medium, and low), sexual receptivity (whether she is cycling or not), and season (breeding vs. birthing) influence their facial femininity. To do so, we calculated one mean value of femininity per 6-month period for each of the 90 studied adult females (total: 5.5k pictures; mean  $\pm$  SD: 263  $\pm$  222 pictures per female and 140  $\pm$  127 pictures per female per 6-month period). We matched these 6-month periods to relevant ecological characteristics of the social and environmental conditions experienced by the mandrills: each 6-month period roughly corresponds to either the birthing season (that mainly occurs during the long rainy season) or the breeding season (that mainly occurs during the long dry season). We found that only females' age significantly correlates with dcm (estimate = 0.37, confidence interval [CI] = [-0.47 to -0.28], p value = <0.001; Table S3), with older females being less feminine than younger ones. This result is robust to various subsampling strategies accounting for potential spurious effects due to outliers (Figure S4). We thus used these femininity scores averaged over 6-month periods in the following behavioral analyses.

For each focal female, we analyzed the occurrence of different socio-sexual interactions with their male and female groupmates, each month of a given 6-month period: the monthly rate of copulations received from all adult males and the monthly rates of spatial association (time spent within 5 m), aggression, and grooming received from either all adult males or all other adult females. We performed one generalized linear mixed model per behavior, as the outcome variable, and considered the following explanatory variables: the female's femininity score (dcm), her age and dominance rank, whether or not she was sexually receptive that month, and the season of behavioral sampling.

With respect to behaviors involving adult male groupmates, we found that males copulate more and are more associated with less feminine females (respectively: estimate = 0.61, 95% CI = [0.38–0.98],  $p = 0.041$ ; estimate = 0.84, 95% CI = [0.74–0.95],  $p = 0.006$ ) (Figures 3A and 3B; Tables S4 and S5). Furthermore, males aggress less feminine females significantly more than more feminine females (estimate = 0.75, 95% CI = [0.59–0.94],  $p = 0.013$ ) (Figure 3D; Table S7). The effect of femininity on aggressions is partly explained by an increased spatial association of males with less feminine females, but not only. Indeed, the effect of femininity remains marginally significant (estimate = 0.82, 95% CI = [0.66–1.02],  $p = 0.070$ ) when the rate of spatial association was included as a covariate in the model based on aggressions (Table S9). Last, femininity has no influence on grooming received by males (estimate = 1.03, 95% CI = [0.62–1.71],  $p = 0.908$ ; Figure 3F; Table S11).

With respect to behaviors involving other adult female group mates, we found that less feminine females are significantly more associated to their female groupmates (estimate = 0.90, 95% CI = [0.83–0.98],  $p = 0.014$ ) (Figure 3C; Table S6) and also receive more aggression than more feminine females (estimate = 0.70, 95% CI = [0.54–0.89],  $p = 0.005$ ) (Figure 3E; Table S8). This last effect remains marginally significant when accounting for spatial associations (estimate = 0.78, 95% CI = [0.61–1.00],  $p = 0.053$ ; Table S10). Last, grooming relationships are not influenced by facial femininity of the study females (estimate = 0.89, 95% CI = [0.74–1.06],  $p = 0.197$ ) (Figure 3G; Table S12).

We then asked whether facial femininity could be associated with behavioral traits expressed by the females. Specifically, we investigated if females displaying lower femininity give higher rates of aggression compared to more feminine females. We used a generalized linear mixed model with the female's rate of emitted aggression as a response variable, and her femininity score (dcm), age, and dominance rank as explanatory variables. We found a significant negative association between femininity and the rate of emitted aggression. This association holds for both male (estimate = 0.75, 95% CI = [0.59–0.94],  $p = 0.013$ ; Figure 4 and Table S14) and female targets (estimate = 0.75, 95% CI = [0.59–0.94],  $p = 0.013$ ; Figure 4 and Table S15). Thus, females with lower facial femininity exhibit elevated rates of aggression.



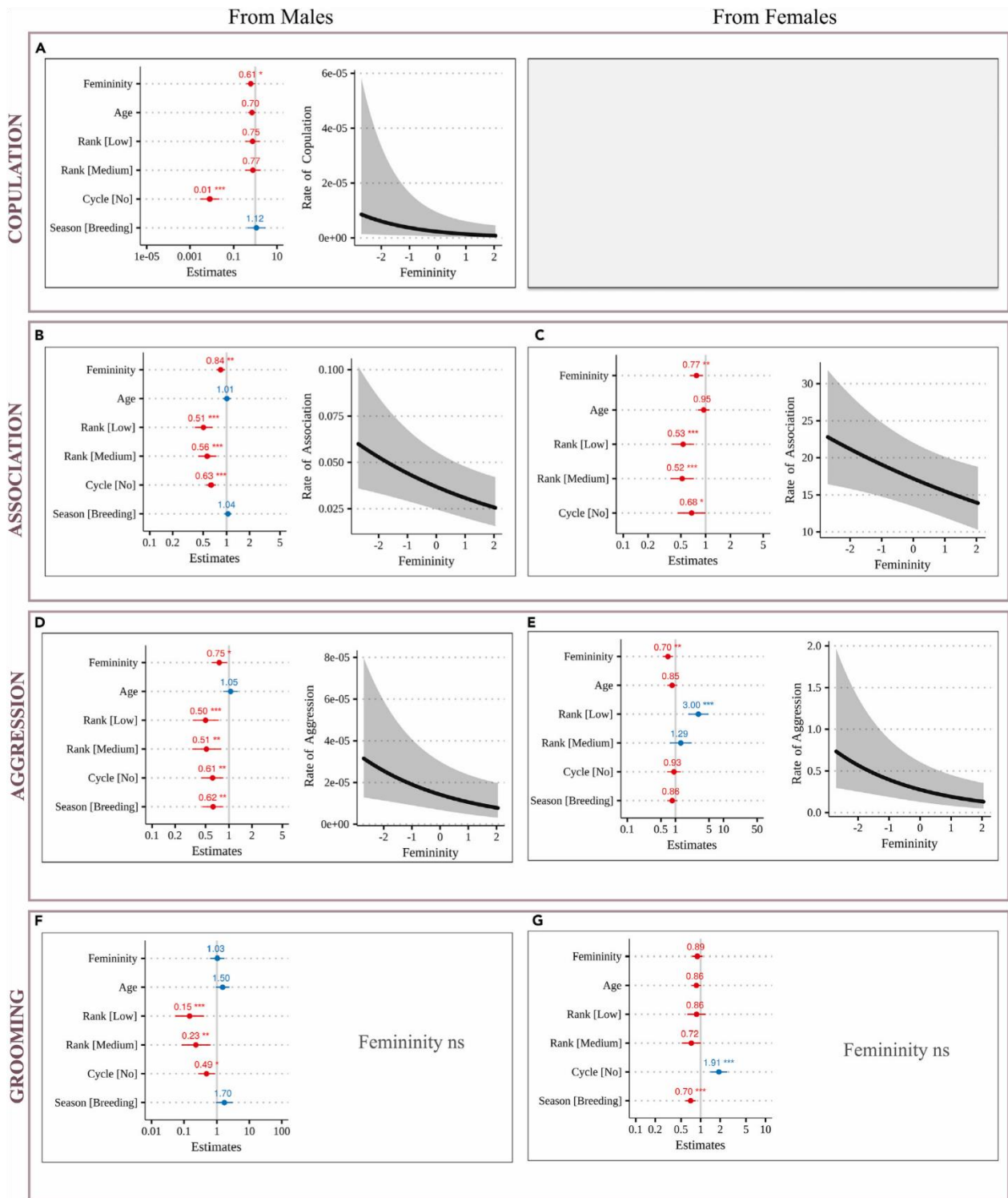


Figure 3. Predicted effect of femininity (dcm) on Mandrill Behavioral Rates

Rates of copulation from males (A); rates of aggression from males (B) and females (C); rates of spatial association with males (D) and females (E); rates of grooming received from males (F) and

females (G). Within each panel, the forest plot on the left gives the direction and size of the fixed effects on the outcome behavior. Significance is indicated (\*:  $p < 0.05$ , 0.01 and 0.001). The gray vertical line indicates no effect (here, 1 because it is a relative risk), and the horizontal bars indicate the confidence interval (95%). A dot on the right of the gray vertical line ( $>1$ , blue) indicates a positive correlation, whereas a dot to the left ( $<1$ , red) indicates a negative correlation. The predicted effect of dcm on the behavior is plotted on the right side of the panel, when significant. ns: not significant. Details in [Tables S4–S12](#).

## DISCUSSION

In this study, we investigated whether facial femininity influences different socio-sexual interactions in a non-human primate. Using datasets with empirical scores of perceived femininity in humans, we first developed an AI-based method to predict femininity scores applicable to non-standardized pictures. We found that femaleness, a measure of sexual dimorphism, is more strongly correlated with perceived femininity than averageness, a measure of similarity to a mean female face. In mandrills, femaleness significantly influences several behaviors received from both male and female conspecifics, and this effect is consistent across various socio-sexual contexts, in all groupmates.

### **Femininity influences socio-sexual behaviors in mandrills**

Our results provide compelling evidence that socio-sexual partners behave differently toward female mandrills depending on the facial femininity of the target female. However, the direction of these effects raises questions about the nature of the information conveyed by facial femininity. Contrary to our expectations, we found that less feminine females, compared to more feminine ones, receive more copulations from males, are spatially closer to both male and female groupmates, and are also more frequently aggressed by them, independent of the females' age, rank, sexual status, and the season of sampling. In this species, males are sexually coercive: they use long-term intimation through aggression toward females, which increases their future probability to mate with them once fertile.<sup>48</sup> Aggressive male behaviors toward female mandrills therefore likely indicate sexual interest, although in this study we observed higher aggression toward less feminine females throughout the year, not only during the breeding season. The similar effect found for female

aggressors, while controlling for female rank, could reflect intra-sexual competition. Female-female competition, compared to male-male competition, has received little scientific attention although reviews have highlighted that females may compete not only for food but also for access to mates or paternal care.<sup>49</sup> If female mandrills compete for access to mates, and if males prefer less feminine females, then increased female aggression toward these preferred (less feminine) females is expected. The evolutionary reasons why females compete are outside the scope of this article, but, interestingly, intra-sexual competition mediated by facial femininity has also been suggested in humans, where women preferred by men are often envied by other women who resort to direct or indirect competition, such as the use of makeup or plastic surgery<sup>50</sup> to rival them.<sup>51</sup>

Altogether, these results, in particular the increased rate of copulations from males toward less feminine females, may appear counterintuitive if femininity is a cue for females' fertility and other maternal skills.<sup>8,9,52</sup> However, the most feminine females are not always the most preferred sexual partners or the most feared sexual competitors, as has been observed in humans. For example, more feminine women are often judged less faithful and more promiscuous.<sup>18</sup> Choosing feminine women would therefore entail the risk of being cheated on, and raising a child that is not one's own.<sup>18</sup> Moreover, high femininity in women can be associated with lower success in competition for resources and lower dominance, leading men to prefer women with more masculine faces, especially in challenging environments.<sup>53,54</sup> These studies highlight the ubiquity of "flips" in the assignment of positive or negative valence to communicative traits.<sup>55</sup> Flips in valence are not limited to humans though. In *Xiphophorus swordtail* fishes, for example, the general preference for pigmented caudal elongations has been reverted in some lineages.<sup>56,57</sup> Observing a general preference for less feminine female faces in mandrills is thus not unexpected. A flip in the generally assumed valence of facial femininity would rather merely indicate that having masculine traits could be advantageous for female mandrills, e.g., to protect children against predators or to win resource conflicts. Interestingly, we found that less feminine females are more aggressive. This aggressiveness could drive the response of groupmates and thus partly explain that they are themselves more aggressive toward less feminine females. However, it is unlikely to explain all the differentiated behaviors expressed by groupmates, particularly the increase in copulations (in mandrills, females do not initiate copulations, pers. obs.). Since

aggressiveness is generally determined by testosterone,<sup>93,94</sup> this result rather reinforces the idea that a set of masculine behaviors could be selected behind the interest shown by males for females with less feminine faces. Future studies should explore this hypothesis, for example, by analyzing whether being a less feminine and more aggressive mother benefits their infants.

### **The origin of the link between femininity and preference**

A long-standing question in evolutionary psychology is whether the attractiveness of femininity results from specific adaptations to mate choice (i.e., the attribute indicator hypothesis), or whether it is rather a preference bias arising from general perceptual mechanisms.<sup>58,59</sup> Our previous interpretation based on a flip in valence relies on an adaptive hypothesis: facial femininity influences preferences because it indicates other physical, physiological, or behavioral attributes (or lack thereof).<sup>8,9,52</sup> Whether a preference bias is an alternative or a complementary explanation to our results, it is necessary to examine the relationship between femininity and femaleness first, and then those between femaleness and attractiveness.

Our result that femaleness in humans is a better predictor of femininity than averageness provides some support to the hypothesis that femininity is, at least partly, determined by perceptual mechanisms. This result probably reflects a contrast effect<sup>60</sup> in which the perception of an individual stimulus is influenced by the statistical distribution of all stimuli in the category of the stimulus and also the distribution of stimuli in other categories. To our knowledge, only one study<sup>61</sup> has previously compared how female facial averageness and sexual dimorphism predict perceived femininity. This study found that face shape dimorphism is correlated with femininity in two of the five ethnic groups studied (Colombian and Iranian), while averageness is correlated with femininity in two other groups (Czech and Turkish). The reason for the observed discrepancies across ethnicities in that study may be due to the fact their face space was based on morphometric features only, while our DCNN-based face space accounts for a greater diversity of features (see in the following). Beyond femininity, other studies have shown that dimorphism is a good predictor of the masculinity of male faces.<sup>39,62</sup> In mandrills as well, the effects of a female's femininity on the behaviors directed at her by both males and other females vanish when quantifying femininity

as averageness (dcf) rather than femaleness (dcm; see [Table S13](#)). This result suggests that femaleness is also a better predictor of femininity in mandrills, and that a contrast effect probably occurs in this primate species too.

In our analyses of human faces, we further showed that femaleness is a good predictor of attractiveness. Again, this relationship could be interpreted through the lens of either the attribute indicator hypothesis or an incidental preference bias driven by perceptual mechanisms. The idea that the link between femininity and attractiveness is driven by perceptual mechanisms is rooted in the fluency theory of attractiveness<sup>63</sup> and its vast empirical evidence.<sup>64,65</sup> The fluency theory proposes that stimuli triggering a subjective sensation of ease in information processing are evaluated as being more attractive.<sup>63,66</sup> Prototype-like stimuli are both easy to process (they are most quickly and precisely categorized and stored the longest in memory<sup>67,68</sup>) and are more attractive than non-prototype stimuli.<sup>66</sup> The attractiveness of prototypes has been shown in humans, using various biological, inanimate, or even abstract visual stimuli without any indicated attributes.<sup>66,67</sup> Regarding facial femininity, a study on chickens investigated preference for human faces: chickens trained to choose a typical female face from a range of male and female faces showed maximal response to female faces more distinctive than the mean during trials, specifically to the face that was also rated as the most attractive by human subjects.<sup>69</sup> This result has been interpreted as evidence that facial femininity is attractive because it is easy on the mind.<sup>70</sup> In mandrills, however, we found the opposite pattern: females with less feminine faces are spatially closer to groupmates, are more aggressed by groupmates, and receive more copulations from males than more feminine females. In this species, it therefore seems unlikely that the relationship between femininity and attractiveness is driven by fluency. This finding echoes classical critics of the fluency theory raised by researchers on artistic preferences, who stressed that attractive art can be perceptually challenging.<sup>71</sup> Further studies in empirical aesthetics have shown that fluency effects actually amplify attractiveness when the stimulus already has a positive valence (i.e., is evaluated as beneficial), but the effect vanishes, or may even reinforce aversion,<sup>72</sup> when the valence is flipped, as suspected in mandrills. To sum up, while we cannot rule out the possibility that femininity is a perceptual construct built from facial features only, our results highlight that in a non-human primate at

least, the influence of femininity on the attractiveness of females is most likely driven by an adaptive, evaluative judgment, here a positive judgment of having masculine attributes.

### **DCNN to study faces in the wild**

Our method allowed us to explain 27% of the variance in perceived femininity ratings when considering all the studied ethnic groups and thus outperformed other face space-based methods for predicting femininity. In comparison, a previous study using a face space to calculate shape dimorphism in women explained a maximum of 20% of the variance in femininity.<sup>61</sup> Another study<sup>41</sup> used the active appearance model<sup>42</sup> to construct a face space that takes into account both shape and texture information but not color. With this model, the authors explained 4% of the variance in attractiveness (these authors did not predict femininity) using the same dataset as the current study (CFD), compared to 18% of the variance in attractiveness with our method. As for studies on men, predictors of masculinity explain 15%<sup>62</sup> to 25%<sup>39</sup> of variance in perceived masculinity.

The transferability of the method to a non-human primate was encouraged by its ability to explain perceived femininity for all ethnic groups studied. Nevertheless, it is noteworthy that the performance varies substantially among ethnicities, ranging from 13% to 30% of variance. Unpublished investigations did not allow to identify the factors responsible for this variation in performance (particularly, the level of sexual dimorphism does not explain variation in performance, results not shown). Consequently, the actual ability of femaleness to predict perceived femininity by conspecifics outside the human species remains speculative.

Our high predictive performance in humans can probably be explained by the combination of our estimate of femininity (based on sexual dimorphism) and our use of a DCNN to define a face space rather than more classical methods. Indeed, a DCNN has several advantages for face studies. First, it can account for any features that are relevant to characterizing faces, including skin color and texture,<sup>73</sup> the global shape, and the relative position of facial elements (mouth, nose, eye<sup>74</sup>). Second, with a DCNN, it is not necessary to predetermine and manually design supposedly relevant features: training the model to perform a high-level predictive task with these faces, for example, face recognition, is

sufficient to learn and design features. Those features can be highly non-linear and more complex than handcrafted features.<sup>31,75</sup> Interestingly, a DCNN trained for face recognition generates a representation that retains structured information about the face (e.g., identity, demographic characteristics, appearance, social features, expression), indicating that the nature of the training task is not critical to learn generalist facial features relevant to many communicative processes beyond face recognition.<sup>29,31</sup> Nevertheless, in this study we showed that training a DCNN to perform both sex recognition and face verification, rather than face recognition, improves the prediction of femininity, indicating that it is worth adapting the learning task to a given research question (**Figure S3**). Third, DCNN-based face spaces are biologically realistic models of face representation. Many studies have highlighted similarities between DCNNs and the brain of primates in the way they process visual information and build internal representations of the world.<sup>76–79</sup> Notably, DCNNs explain physical similarity<sup>39,80</sup> and dissimilarity<sup>27</sup> judgments between faces remarkably well. Lastly, DCNNs can be trained to be insensitive to the general coloration, brightness, and quality of an image, the pose and the emotion of a face, and its relative size on the image, if all these variations are non-informative for a given question (the case when studying femininity<sup>29,31</sup>). The invariance property of DCNNs explains that predictions of femininity in humans were similarly good with in-lab, standardized portrait images and with non-standardized images retrieved from the Internet. Faces are important for the visual communication of many animal species, playing a role for example in individual recognition, mate choice, kin discrimination, or social status advertising.<sup>32,81,82</sup> Therefore, DCNNs provide new opportunities for non-invasive research in animal communication and behavioral ecology by enabling the investigation of the information conveyed through animals' faces from portraits captured in their natural environments.

### **Limitations of study**

First, our methodology employs a DCNN, an AI-based approach, for predicting femininity scores from non-standardized images. While the DCNN presents inherent benefits, the adaptation of specific training tasks to our research question implies that further tailoring may be needed for application in other research contexts or species. Second, our study focuses on facial femininity and its impact on socio-sexual behaviors. This correlative

approach does not allow, however, to explore the evolutionary mechanisms that drive the preference for less feminine facial features.

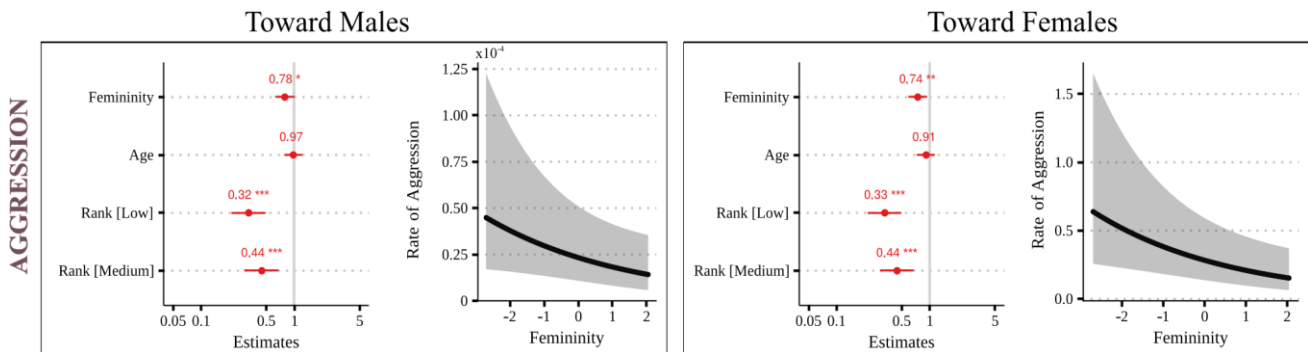


Figure 4. Predicted effect of femininity (dcm) on rates of aggression emitted toward males (left) and females (right)

Within each panel, the forest plot on the left gives the direction and size of the fixed effects on the outcome behavior. Significance is indicated (\*:  $p < 0.05$ , 0.01 and 0.001). The gray vertical line indicates no effect (here, 1 because it is a relative risk), and the horizontal bars indicate the confidence interval (95%). A dot on the right of the gray vertical line ( $>1$ , blue) indicates a positive correlation, whereas a dot to the left ( $<1$ , red) indicates a negative correlation. The predicted effect of dcm on the behavior is plotted on the right side of the panel, when significant. Details in [Tables S13](#) and [S14](#).



## STAR METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Deposited data</b>		
Mandrillus Face Dataset	Zenodo	<a href="https://doi.org/10.5281/zenodo.7467318">https://doi.org/10.5281/zenodo.7467318</a>
Cleaned Dataset	GitHub	<a href="https://github.com/soniamaitieo/Social_Sexual_Interactions_-Mandrills">https://github.com/soniamaitieo/Social_Sexual_Interactions_-Mandrills</a>
Code	GitHub	<a href="https://github.com/soniamaitieo/Social_Sexual_Interactions_-Mandrills">https://github.com/soniamaitieo/Social_Sexual_Interactions_-Mandrills</a>
<b>Experimental models: Organisms/strains</b>		
95 adult females ( $\geq 4$ years) and 39 males (subadults or adults, $\geq 9$ years) among free-ranging mandrills		
<b>Software and algorithms</b>		
R v3.6.3	Website	<a href="https://cran.r-project.org/">https://cran.r-project.org/</a>
R packages are cataloged in a csv	GitHub	<a href="https://github.com/soniamaitieo/Social_Sexual_Interactions_-Mandrills">https://github.com/soniamaitieo/Social_Sexual_Interactions_-Mandrills</a>
Python v3.8.5	Website	<a href="https://www.python.org/">https://www.python.org/</a>
Anaconda	Website	<a href="https://www.anaconda.com/">https://www.anaconda.com/</a>
Python packages for the Anaconda environment are listed in a yaml file	GitHub	<a href="https://github.com/soniamaitieo/Social_Sexual_Interactions_-Mandrills">https://github.com/soniamaitieo/Social_Sexual_Interactions_-Mandrills</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further information should be directed to and will be fulfilled by the lead contact, Sonia Tiew ([tiew.soniama@gmail.com](mailto:tiew.soniama@gmail.com)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

- Data are publicly available as of the date of publication. Link is listed in the [key resources table](#).
- All original code has been deposited on GitHub and is publicly available as of the date of publication. The link is listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### **Ethical guidelines**

The study received the necessary approval from the CENAREST institute (permit AR017/22/MESRSTTCA/CENAREST/CG/ CST/CSAR), adhering to all relevant international, national, and institutional guidelines for the proper care and humane treatment of animals.

### **Study subjects**

The study subjects are unfed and non-manipulated mandrills from a natural population living near the Lékédi Park in Southern Gabon. This group, which is fully habituated to humans, originated from captive mandrills that were released from the Center International de Recherches Médicales de Franceville (CIRMF) in 2002 and 2006.<sup>88</sup> The group, daily monitored by the Mandrillus Project since its inception in 2012, comprised approximately 220 mandrills by the end of 2021, which included only 7 females that were born in captivity. For this study we considered 95 females who were 4 years of age or older and 39 males, either subadults or adults, aged 9 years and above.

## METHOD DETAILS

### **Image datasets**

For retraining VGGFace to verify identity and classify sex, we used a subset of the FaceScrub dataset<sup>37</sup> that includes 37,193 portrait images representing 530 individuals. Femininity and attractiveness analyses were performed with Chicago Face Database (CFD<sup>83</sup>) and SCUTFBP5500 dataset.<sup>43</sup> CFD provides standardized (neutral expression, frontal view, same clothes) photographic portraits of 597 American people (one photo per individual) aged 17 to 65 years, of different ethnic backgrounds (number of Women-Men; East Asian: 57-52, Black: 104-93; Hispanic: 56-52; White: 90-91). The face photographs were scored for attractiveness (from 1 to 6) and femininity (from 1 to 7) by independent judges. SCUT-FBP5500 includes 5,500 non-standardized portrait images of Asian (women: 2,000; men: 2,000) and White

(women: 750; men: 750) people (one photo per individual) of different ages (15–60 years), and their associated scores of attractiveness (from 0 to 5) obtained by independent judges.

Mandrill portraits have been retrieved from the Mandrillus Face Database (MFD), which includes pictures collected between January 2012 and December 2021. The pictures were taken by field assistants directly in the forest, using different models of DSLR cameras and long-focal lenses, and different settings. Moreover, distance to the camera, illumination, pose and facial expression vary between pictures. Pictures of MFD are thus qualified “non-standardized”. All images have been cropped manually to keep the face only (without the ear), and rotated to align the eyes horizontally. The quality of each image has been manually scored. Quality 1 corresponds to pictures of average quality for which experienced field assistants are able to recognize the individual from the picture alone. Quality 2 describes pictures of good quality for which individual recognition is straightforward but the portrait does not meet the criteria of quality 3. Quality 3 describes portrait pictures of maximal quality (“id card-like” portraits). For further information on image acquisition and pre-processing, see.<sup>40</sup> For this study, we only considered adult individuals (females aged >4 years and males aged >9 years), corresponding to a total of 10,057 pictures collected on 119 adult females (83.8 G 74 pictures per female) and 2,928 pictures from 50 adult males (57.4 G 61.9 pictures per male). Femininity values (dcm and dcf) were obtained from these pictures on 119 females. Out of these 119 females, only 95 were used to determine their socio-sexual interactions with groupmates. Because the number of pictures per individual largely varies (Figure S5), we tested if this number influenced dcm and dcf when averaged over a 6-month period. We found no correlation between the number of pictures and both dcf and dcm values per individual, per 6-month period (Figure S6).

### Face space

The face space was defined as the latent space (penultimate layer) of a customized and retrained version of VGGFace.<sup>35,36</sup> After the second fully connected layer of the classical VGG16 architecture, we added a dropout layer (rate: 0.3), a dense layer with 128 neurons without activation (the Face Space), a lambda regularizer and a last dense layer with two neurons. The model was retrained to simultaneously verify the identity of individuals, using a Triplet Semi Hard loss function (from TensorFlow Addons<sup>84,85</sup> connected to the face space,

and to classify sex, using a sparse categorical cross entropy loss function connected to the 2-neuron dense layer. For training this model, we utilized 10,057 pictures with quality score ranging from 1 to 3. We used 60% of the images for training, 30% for validation and 10% for testing. Images were resized to 224x224 pixels. We used the Adam optimizer<sup>86</sup> (learning rate: 0.001) and a batch size of 128. Model performances were evaluated on the test dataset using the mean average precision mAP@3 accuracy for face verification, and categorical accuracy for sex classification. Once retrained (i.e., performances did not improve further, usually after approx. 100 epochs), we used the model to encode new portrait images. For estimating femininity, we used pictures of quality scores 2 and 3 (thus not considering low-quality portraits). DCNN analyses were performed with Python 3 and TensorFlow.v2<sup>87</sup> library on one NVIDIA GTX 1080 GPU.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### **Calculation of femininity on the PC1 axis**

To calculate dcf and dcm in mandrills, we used PC1 only. Indeed, we observed that the intra-individual variation of femininity is 23 times lower than the inter-individual variation when using PC1 only, while it is only 10 times or 14 times lower when using PC1-PC2 or the 128-day face space, respectively, because dimensions other than PC1 encode variations between faces that are not related to the sex. Female averageness (dcf), was calculated as the absolute distance between each individual face's PC1 score (PC1<sub>i</sub>) and the PC1 projection of the female centroid (cf.; see also [Figure 1](#)):

$$dcf = abs(PC1_i - cf)$$

Similarly, femaleness (dcm) was calculated as the absolute distance between PC1<sub>i</sub> and the PC1 projection of the male centroid (cm):

$$dcm = abs(PC1_i - cm)$$

Note that in our case (both in humans and mandrills) PC1<sub>i</sub> - cm was always positive, meaning that using the absolute or the relative value yields similar results.

## **Mandrills' behavioral data**

Behavioral data were collected in the context of a long-term field project (the Mandrillus Project) that studies the socio-ecology of a natural population of mandrills habituated to human presence. The study population roams in a private park (the Le' ke' di Park) and its vicinity, close to the village of Bakoumba, in Southern Gabon (for more information, see<sup>88</sup>). Trained observers, blind to the study questions, perform daily 5-min focal sampling and ad libitum observations on all studied individuals.<sup>89</sup> We used 13.019 h of focal data collected from 134 individuals: 95 adult females aged R4 years with 451 G 739h of focals per female and 39 subadult and adult males aged R9 years with 514 G 803h of focals per male (see details in [Figure S7](#)). During focals, all socio-sexual interactions are recorded between the focal individual and all individually-recognized social partners, including copulation (N = 132 events), grooming (N = 140 events), aggression (N = 656 events including chases, bites, ground-slapping; see<sup>48</sup> for more information). Spatial associations are further recorded between the focal individual (1–3 scans per focal observation) and all groupmates located within a circle of 5m of radius length, (36,517 scans). Following Altmann's observational research methods,<sup>89</sup> we used a 5-meter radius scan, a standard method that allows for capturing comprehensive behavioral data. For the purpose of this study, we pooled all the focals performed on each study female with those collected on all other females or all other males present in the study group each month and for whom focals were available. For each female, we retrieved the number of monthly copulations received by all males, and the number of monthly grooming and aggressions received from all males and all females. We further retrieved the number of times all males or all females were found associated with each focal female using all scans performed on that female and her male and female partners, each month.

Birth dates of the study mandrills are either exactly known or estimated within a few days or months from daily monitoring or from general body condition and tooth wear patterns.<sup>90</sup> Female dominance rank is estimated using outcomes of approach-avoidance collected during focal or ad libitum observations and calculated using standardized David's score.<sup>91</sup> Adult females are divided into three classes of rank of similar size over the study

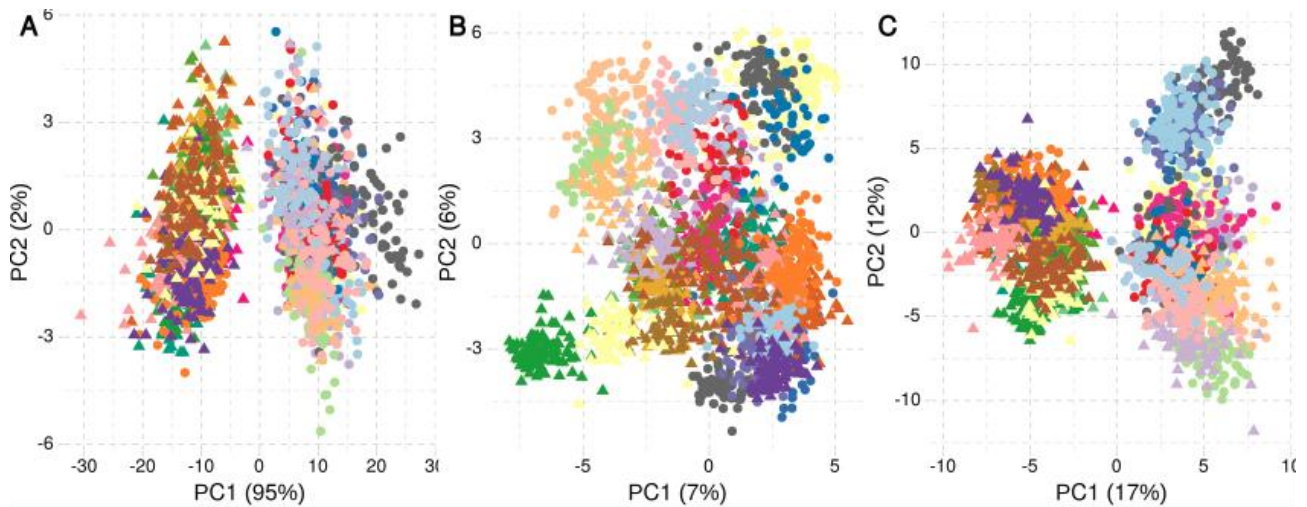
period (HR-high rank, MR-medium rank and LR-low rank) because female rank is stable in this matrilineal society.<sup>92</sup>

For each month of behavioral data collected on the focal females and their social partners, we determined whether these females were observed cycling that month (yes or no) thanks to a daily monitoring of female sexual cycles in the study population. Cycling females present conspicuous inflation of their genitalia, until they reach a maximum size corresponding to a fertile period that may last a few days.<sup>33</sup>

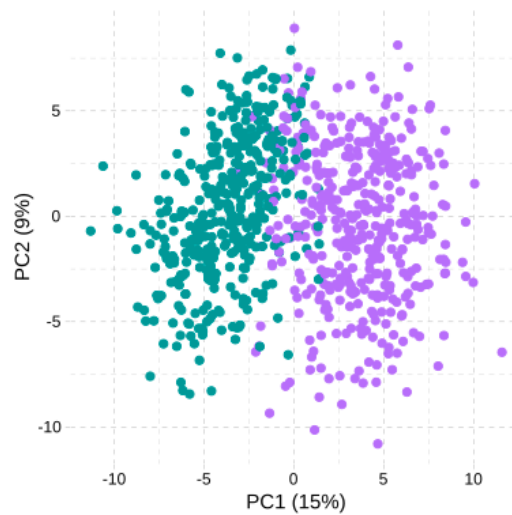
### **Statistical analyses**

To evaluate the effect of femininity (dcm) on response variables (study behaviors), we used Generalized Linear Mixed Models (GLMM) with a negative binomial distribution and a logit function for over-dispersed count data. We log-transformed the total duration of focals, and the total number of scans performed, collected each month on each study female and all male or female groupmates and considered them as an offset (Nb\_Scan\_IndAllmal) in their respective models to control for different monthly duration of focals or scans performed. In addition to dcm calculated over 6-month periods corresponding to the breeding vs. the birthing season, we considered the following predictors in our statistical models as fixed effects: a female's dominance rank (HR, MR or LR), age (years) and sexual status (cycling or not) at the month of behavioral data collection. Finally, we controlled for the season of sampling (breeding vs. birthing). We standardized all continuous predictors to compare estimates (dcm and age). The identity (id) of the focal female and the year of data collection were considered as two random effects to control for repeated monthly observations on focal females and yearly variation in sociality. For the rate of copulation, for example, the model is  $\text{RateOfCopulation} \sim (1 | \text{id}) + (1 | \text{year}) + \text{Femininity} + \text{Sexual status} + \text{Age} + \text{Dominance rank} + \text{offset}(\log(\text{Nb\_Scan\_IndAllmal}))$ . GLMMs were run using spAMM package with the fitme function (method = c("ML", "obs")). In preliminary analyses, we considered all possible first-order interactions involving dcm as well as the female age as a quadratic effect. None were significant, we thus removed these first-order interactions and the quadratic effect of age from final models. In order to represent behavioral rates, we first set the offset to 1 and then create a new data object; an argument to the pdep\_effects function to evaluate the effect of a variable and then plotted this with ggplot2. The estimates are back transformed.

## SUPPLEMENTARY INFORMATION

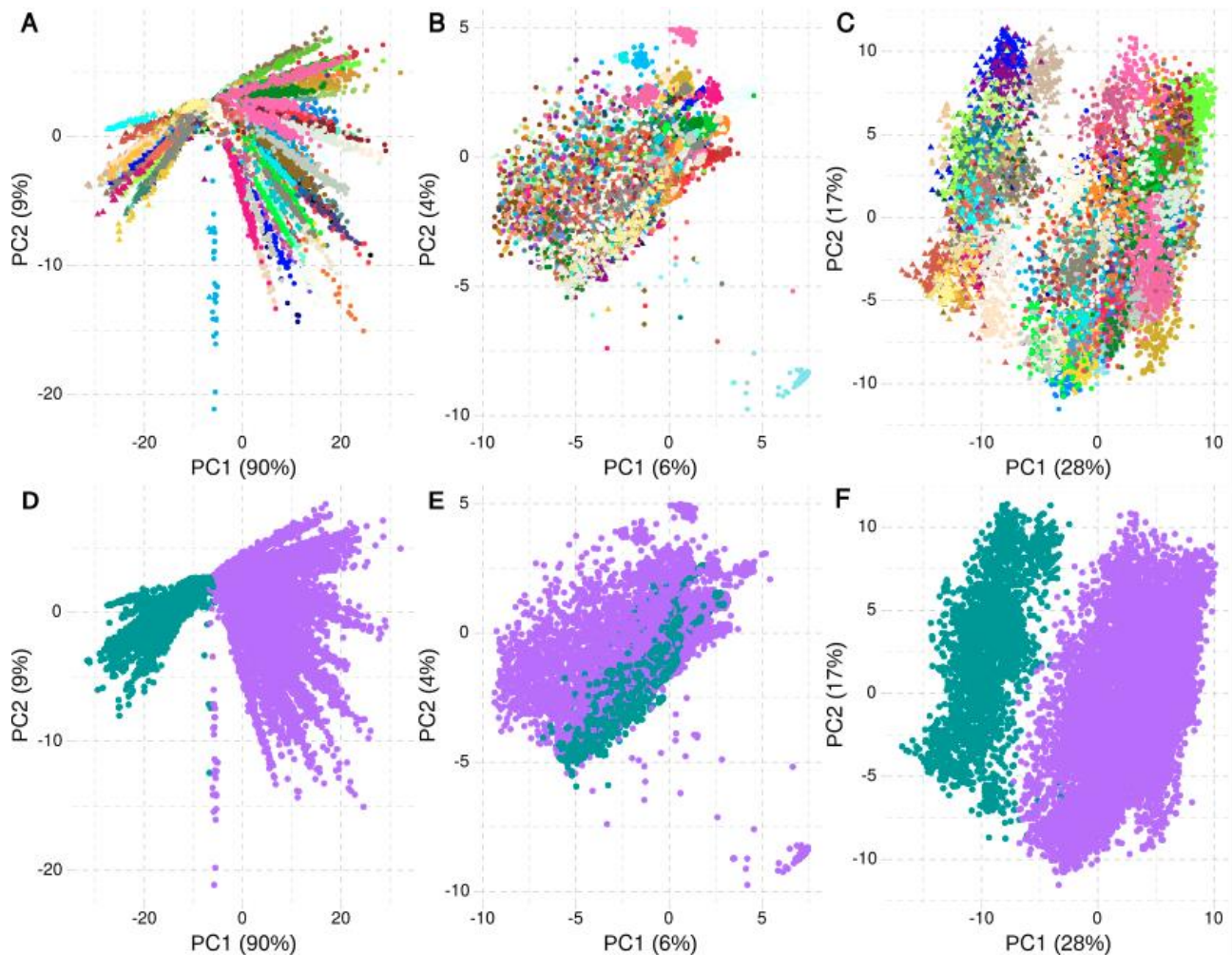


**Figure S1. Influence of the learning task on the human face space.** Human faces of FaceScrub dataset are projected into the face space, and their coordinates in the face space reduced by PCA. For visualization, coordinates along the first two principal components are shown for portraits of only 30 individuals colored by their identity. The face space corresponds to the 128-dimensional penultimate fully connected layer of VGGFace retrained on the FaceScrub dataset to perform different tasks: (A) sex classification, (B) individual verification, (C) individual verification and sex classification. This last task gives a face space that both separates sex and clusters same-individual portrait images. The percentages in the axis labels indicate the fraction of variance explained by PC1 and PC2. Triangles correspond to males, and circles to females.

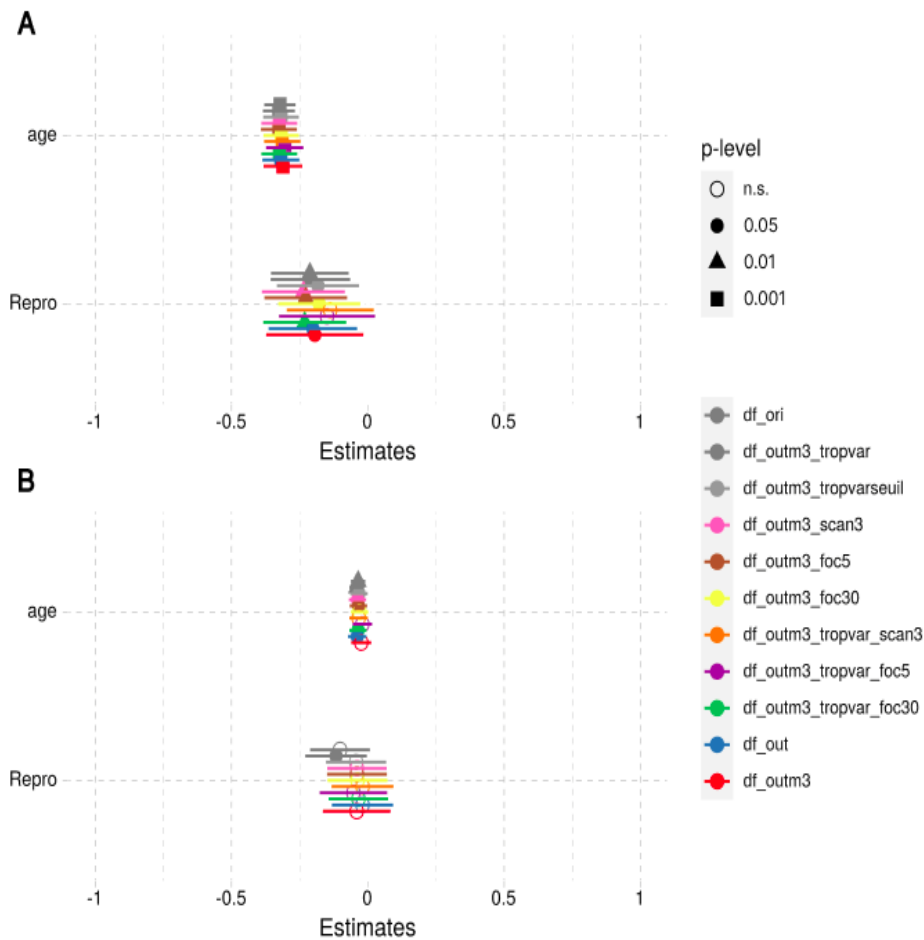


**Figure S2. Visualization of the faces of the Chicago Face Dataset (CFD) into the PC1-PC2 plan of the face space.** The portraits of the 597 individuals of CFD are colored by their sex (green for male and violet for female). The face space corresponds to the 128-dimensional penultimate fully connected layer of VGGFace retrained on the FaceScrub dataset to simultaneously perform identity verification and sex classification. Coordinates of CFD portraits in the face space were then reduced to two dimensions corresponding to the first two components (PC1 and PC2) of a PCA performed on the original 128 dimensions. The percentages in the axis labels indicate the fraction of variance explained by PC1 and PC2.



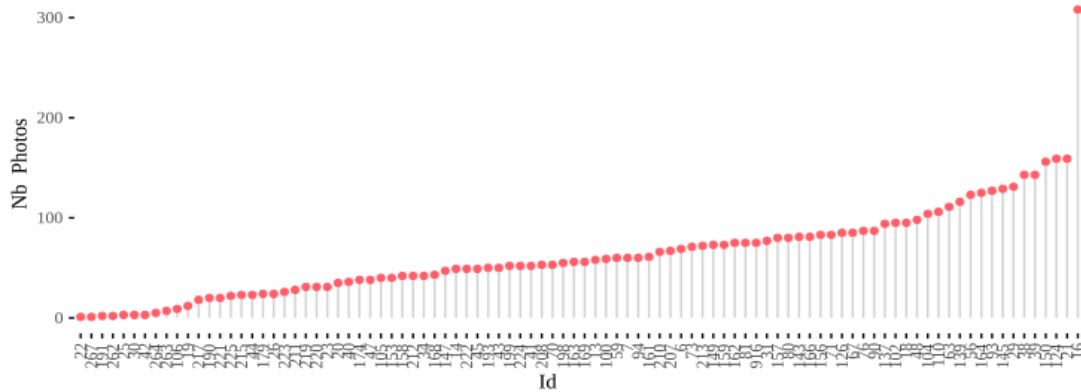


**Figure S3. Influence of the learning task on the mandrill face space.** The mandrill face space displays 12,985 pictures of 169 adult individuals from Mandrillus Face Dataset colored by their identity (A, B, C; triangles: males and circles: females) and their sex (D, E, F ; green: males and violet: females). We used the penultimate 128-dimensional fully connected layer of VGGFace retrained on the Mandrillus Face Dataset dataset as a face space. Portrait images of Mandrillus Face Dataset were then projected onto this face space, and their coordinates reduced by PCA. The percentages in the axis labels indicate the fraction of variance explained by PC1 and PC2. The projection on PC1 and PC2 gives different distributions of the faces depending on the learning task: (A,D) sex classification, (B,E) individual verification, and (D,F) individual verification plus sex classification. This last task gives a face space that both separates sex and clusters same-individual portrait images.

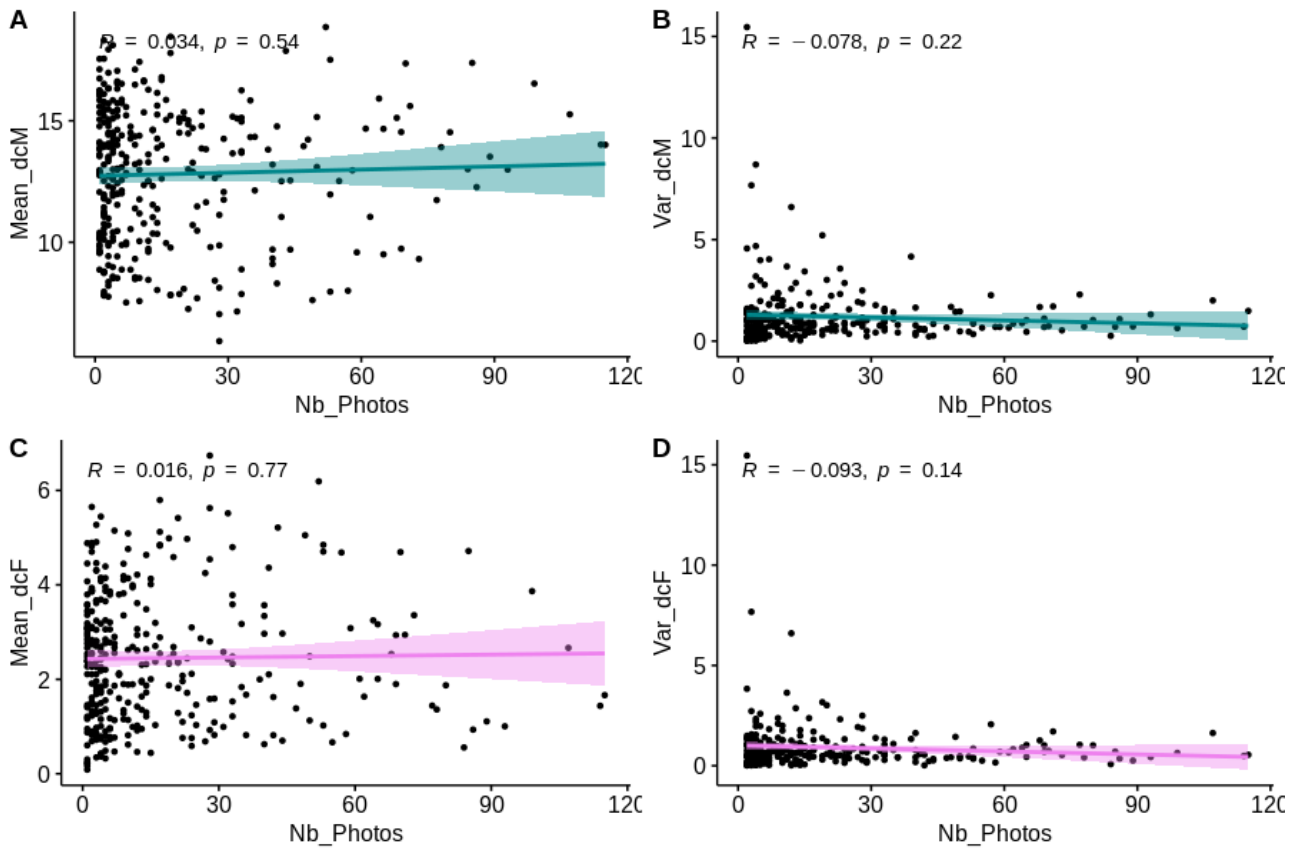


**Figure S4. Effect of age and reproductive season on dcm (A) and dcf (B) on different subdataset.** We fitted a linear mixed model (estimated using REML and nloptwrap optimizer) to predict dcm or dcf with age (continuous variable) and season (two-level categorical variable; formula:  $dcm \sim age + season$ ). The model included the individual identity (id) and year of shooting (year) as random effects (formula:  $list(\sim 1 | id, \sim 1 | year)$ ). The forest plot of odds ratios gives the direction and size of the effects of the variables of interest (age and season). The symbols indicate the significance level and the error bar the standard error of the effect (see legend within the figure). Estimates equal to 0 indicate no effect. Dots on the left, negative side of zero shows a negative effect. These plots highlight the robustness of the model, that is, effect size and direction are similar across different strategies of data sampling: (1) df\_ori: all observations included, (2) df\_ot: photo outliers removed (based on IQR criteria), (3) df\_outm3: photo outliers and observations with less than 3 pictures removed, (4) df\_outm3\_tropvar: photo outliers, observations with less than 3 pictures, and observations for which the intra-individual variance of dcm or dcf within a season is larger than the inter-individual variance removed, (5) df\_outm3\_tropvar: photo outliers, observations with less than 3 pictures, and observations for which the intra-individual variance of dcm or dcf within a season - with 10% tolerance- is larger than the inter-individual variance removed, (6) df\_outm3\_tropvar\_scans3: photo outliers, observations with less than 3 pictures, observations for

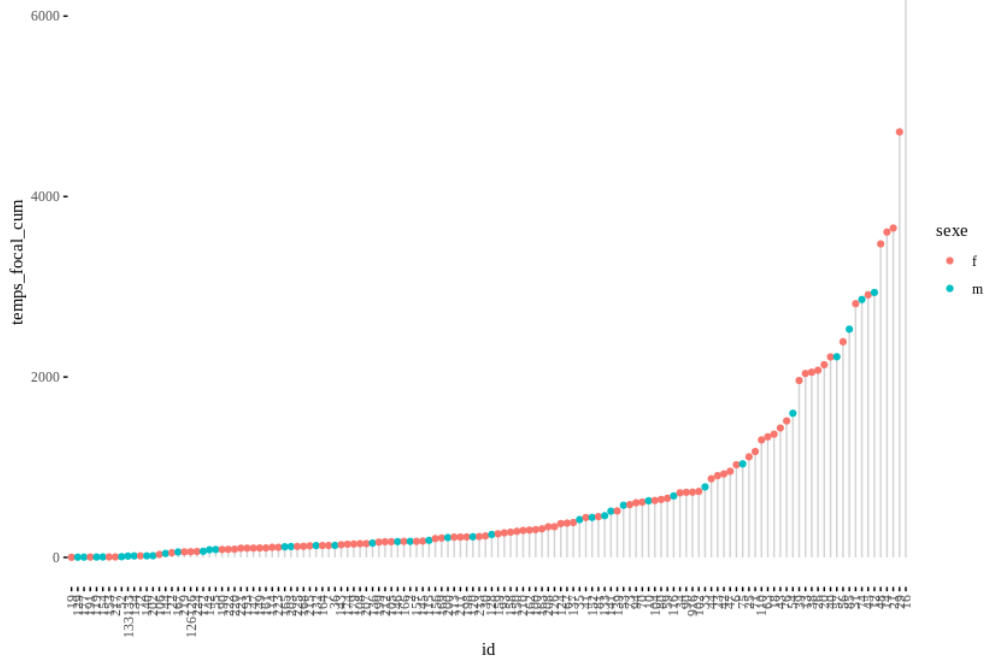
which the intra-individual variance of dcm or dcf within a season -with 10% tolerance- is larger than the inter-individual variance, and observations with less than 3 scans removed, (7) df\_outm3\_tropvar\_foc5: photo outliers, observations with less than 3 pictures, observations for which the intra-individual variance of dcm or dcf within a season -with 10% tolerance- is larger than the inter-individual variance, and observations with less than 3 scans and with focal lengths of less than 5 minutes removed, (8) df\_outm3\_tropvar\_foc30: photo outliers, observations with less than 3 pictures, observations for which the intra-individual variance of dcm or dcf within a season -with 10% tolerance- is larger than the inter-individual variance, and observations with less than 3 scans and with focal lengths of less than 5 minutes removed.



**Figure S5. Number of photos per female of the studied mandrills population.** Id along the x-axis indicates individual identifiers.



**Figure S6. Correlation plot between mean or variance in femininity and the number of pictures.** Observations are pooled by Y-S periods. Mean femininity (Mean\_dcm or Mean\_dcf) corresponds to the mean of dcm (or dcf) of all pictures of a female per Year-Season period, and Var\_dcm (or Var\_dcf) to the variance of dcm (or dcf) of all pictures of a female per Year-Season period. R corresponds to the Pearson coefficient of correlation and p to the p-value.



**Figure S7. Cumulative focal times per individual.** Id along the x-axis indicates individual identifiers.

**Table S1. Correlation between the perceived femininity or attractiveness and the predicted femaleness or averageness for CFD dataset.** Femaleness is calculated on PC1 (dcm) or in the 128d-FaceSpace (dcm\_FaceSpace). Averageness is calculated on PC1 (dcf) or in the 128d-FaceSpace (dcf\_FaceSpace). Separate analyses were conducted for all 307 females (All), 104 Black females, 57 Asian females, 56 Latino females, and 90 White females. R corresponds to the Pearson coefficient of correlation and p to the p-value.

CFD	All		Asian		White		Black		Latino	
	R	p	R	p	R	p	R	p	R	p
dcf_FaceSpace ~ Feminine	0.16	0.005	0.13	0.35	0.1	0.35	0.18	0.061	0.27	0.046
dcm_FaceSpace ~ Feminine	0.42	3e-14	0.24	0.07	0.39	0.00017	0.45	1.3e-06	0.4	0.0021
dcf ~Feminine	-0.16	0.005	-0.12	0.36	-0.2	0.058	-0.017	0.86	-0.11	0.42
dcm ~Feminine	0.52	1e-22	0.36	0.0064	0.54	4.3e-08	0.54	2.5e-09	0.42	0.0013
dcf_FaceSpace ~ Attractive	0.12	0.034	0.19	0.15	0.12	0.26	0.12	0.22	0.11	0.44
dcm_FaceSpace ~Attractive	0.33	2.8e-09	0.26	0.052	0.35	0.00061	0.36	0,0002	0.25	0.064
dcf ~Attractive	-0.12	0.04	-0.013	0.92	-0.14	0.17	0.047	0.63	-0.22	0.11
dcm ~ Attractive	0.43	2.7e-15	0.34	0.0087	0.47	2.6e-06	0.48	3.4e-07	0.3	0.025

**Table S2. Correlation between the perceived attractiveness and the predicted femaleness or averageness for SCUT-FRB5000 dataset.** Femaleness is calculated on PC1 (dcm) or in the 128d-FaceSpace (dcm\_FaceSpace). Averageness is calculated on PC1 (dcf) or in the 128d-FaceSpace (dcf\_FaceSpace). Separate analyses were conducted for all 2,750 females (All), 2,000 Asian females (Asian) and 750 White females. R corresponds to the Pearson coefficient of correlation and p to the p-value. The very low R score for dcm with All is due to that PC1 does not separate the sexes when all ethnicities are grouped together, since variation in ethnicities explains most of the variation between faces.

SCUT	All		Asian		White	
	R	p	R	p	R	p
dcf_FaceSpace ~ Attractive	0.098	2.6e-07	0.04	0.072	0.21	4.4e-09
dcm_FaceSpace ~ Attractive	0.23	3.4e-34	0.19	2,00E-18	0.37	1.1e-25
dcf ~ Attractive	-0.043	0.024	-0.013	0.57	0.061	0.092
dcm ~ Attractive	0.067	0.00042	0.33	2.1e-51	0.49	1.1e-45

**Table S3. Effect of the season and female’s age, dominance rank and sexual status on predicted femininity (dcm).** We fitted a generalized linear mixed model (estimated using REML and nloptwrap optimizer) to predict dcm from the different predictors. Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using a Wald t-distribution approximation. Estimates are also provided. Significant effects are highlighted in bold. The mean value of dcm is 13 and the variance is 7.

<i>Predictors</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.02	-0.28 – 0.31	0.918
age	-0.37	-0.47 – -0.28	<b>&lt;0.001</b>
saison repro [Repro]	0.01	-0.02 – 0.03	0.483
Cycle [NO]	0.01	-0.03 – 0.04	0.772
rank [LR] (ref: HR)	-0.08	-0.40 – 0.24	0.627
rank[MR] (ref: HR)	-0.01	-0.33 – 0.30	0.935
<b>Random Effects</b>			
$\sigma^2$	0.07		
ICC	0.88		
N id	88		
N year	10		
Observations	2294		
Marginal R2 / Conditional R2	0.196 / 0.903		



**Table S4. Model predicting the rate of copulation between a focal female and all males.**

We fitted a generalized linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of copulation (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfCopulation} \sim (1 \mid \text{id}) + (1 \mid \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{Season} + \text{offset}(\log(\text{Nb\_Scan\_IndAllmal}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 \mid \text{id}, \sim 1 \mid \text{year})$ ). The effect of femininity (dcm) is statistically significant and positive: males copulate more with less feminine females. Model performed with spaMM and results reported with sjPlot (R packages). Significant effects are highlighted in bold. **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.00	0.00 – 0.00	<b>&lt;0.001</b>
Femininity	0.61	0.38 – 0.98	<b>0.040</b>
Age	0.70	0.45 – 1.08	0.107
Rank [LR] (ref: HR)	0.75	0.34 – 1.65	0.471
Rank [MR] (ref: HR)	0.77	0.34 – 1.77	0.538
Cycle [No]	0.01	0.00 – 0.02	<b>&lt;0.001</b>
Season [Breeding]	1.12	0.42 – 2.97	0.821
N id	88		
N year	10		
Observations	2239		

**Table S5. Model predicting the rate of spatial association between a focal female and all males.** We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of spatial association (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfSA} \sim (1 \mid \text{id}) + (1 \mid \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{Season} + \text{offset}(\log(\text{Nb\_Scan\_IndAllmal}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 \mid \text{id}, \sim 1 \mid \text{year})$ ). The effect of femininity (dcm) is statistically significant and positive: males associate more with less feminine females. Model performed with spaMM and results reported with sjPlot (R packages). **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.09	0.06 – 0.13	<0.001
Femininity	0.84	0.74 – 0.95	0.006
Age	01.01	0.89 – 1.14	0.905
Rank [LR] (ref: HR)	0.51	0.39 – 0.66	<0.001
Rank [MR] (ref: HR)	0.56	0.43 – 0.73	<0.001
Cycle [No]	0.63	0.55 – 0.72	<0.001
Season [Breeding]	01.04	0.94 – 1.16	0.436
N id	88		
N year	10		
Observations	2239		

**Table S6. Model predicting the rate of spatial association between a focal female and all females.** We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Spatial Association (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfSA} \sim (1 \mid \text{id}) + (1 \mid \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{Season} + \text{offset}(\log(\text{Nb\_Scan\_IndAllfem}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 \mid \text{id}, \sim 1 \mid \text{year})$ ). The effect of femininity is statistically significant and positive: females associate more with less feminine females. Model performed with spaMM and results reported with sjPlot (R packages). **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.04	0.03 – 0.06	<0.001
Femininity	0.90	0.83 – 0.98	0.014
Age	0.94	0.86 – 1.03	0.172
Rank [LR] (ref: HR)	0.87	0.71 – 1.05	0.146
Rank [MR] (ref: HR)	0.76	0.62 – 0.93	0.006
Cycle [No]	1.09	0.99 – 1.19	0.073
Season [Breeding]	0.92	0.86 – 0.98	0.006
N id	88		
N year	10		
Observations	2239		

**Table S7. Model predicting the rate of Aggression between a focal female and all males.**

We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Aggression (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfAggression} \sim (1 \mid \text{id}) + (1 \mid \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{Season} + \text{offset}(\log(\text{Nb\_Scan\_IndAllfem}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 \mid \text{id}, \sim 1 \mid \text{year})$ ). The effect of femininity (dcm) is statistically significant and positive: males aggress more the less feminine females. Model performed with spaMM and results reported with sjPlot (R packages). **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.00	0.00 – 0.00	<0.001
Femininity	0.75	0.59 – 0.94	0.013
Age	01.05	0.84 – 1.31	0.671
Rank [LR] (ref: HR)	0.50	0.34 – 0.74	<0.001
Rank [MR] (ref: HR)	0.51	0.33 – 0.79	0.002
Cycle [No]	0.61	0.44 – 0.84	0.003
Season [Breeding]	0.62	0.46 – 0.84	0.002
N id	88		
N year	10		
Observations	2239		

**Table S8. Model predicting the rate of Aggression between a focal female and all females.** We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Aggression (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfAggression} \sim (1 | \text{id}) + (1 | \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{Season} + \text{offset}(\log(\text{Nb\_Scan\_IndAllfem}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 | \text{id}, \sim 1 | \text{year})$ ). The effect of femininity (dcm) is statistically significant and positive: females aggress more the less feminine females. Model performed with spaMM and results reported with sjPlot (R packages). **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.00	0.00 – 0.00	<0.001
Femininity	0.70	0.54 – 0.89	0.005
Age	0.85	0.67 – 1.08	0.181
Rank [LR] (ref: HR)	3.00	1.84 – 4.88	<0.001
Rank [MR] (ref: HR)	1.29	0.76 – 2.18	0.341
Cycle [No]	0.93	0.67 – 1.29	0.671
Season [Breeding]	0.86	0.67 – 1.09	0.211
N id	88		
N year	10		
Observations	2239		

**Table S9. Model predicting the rate of Aggression between a focal female and all males with Spatial Association as an explanatory variable.** We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Aggression (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfAggression} \sim (1 | \text{id}) + (1 | \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{RateOfSA} + \text{Season} + \text{offset}(\log(\text{Tps\_Foc\_IndAllmal}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 | \text{id}, \sim 1 | \text{year})$ ). The effect of femininity (dcm) is not statistically significant but positive. Model performed with spaMM and results reported with sjPlot (R packages). The effect of femininity on aggressions is partly explained by an increased spatial association of males with less feminine females, but not only. Indeed, the effect of femininity remains marginally significant (estimates = 0.82, 95% CI = [0.66 – 1.02], p = 0.070) when the rate of spatial associations was included as a covariate in the model predicting aggressions from femininity. **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.00	0.00 – 0.00	<0.001
Femininity	0.82	0.66 – 1.02	0.070
Age	45078	0.87 – 1.30	0.542
Rank [LR] (ref: HR)	0.67	0.48 – 0.94	0.022
Rank [MR] (ref: HR)	0.64	0.43 – 0.97	0.036
Cycle [No]	0.70	0.50 – 0.98	0.039
Season [Breeding]	0.63	0.46 – 0.85	0.003
Rate of S.A.	857.37	85.24 – 8623.30	<0.001
N id	88		
N year	10		
Observations	2122		

**Table S10. Model predicting the rate of Aggression between a focal female and all females with Spatial Association as an explanatory variable.** We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Aggression (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfAggression} \sim (1 | \text{id}) + (1 | \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{RateOfSA} + \text{Season} + \text{offset}(\log(\text{Tps\_Foc\_IndAllfem}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 | \text{id}, \sim 1 | \text{year})$ ). The effect of femininity (dcm) is almost statistically significant and positive: females attack less feminine females. Model performed with spaMM and results reported with sjPlot (R packages). This effect remains marginally significant when accounting for spatial associations in the statistical model (estimates = 0.78, 95% CI = [0.61 – 1.00], p = 0.053; Table S10). **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.00	0.00 – 0.00	<0.001
Femininity	0.78	0.61 – 1.00	0.053
Age	0.95	0.75 – 1.19	0.647
Rank [LR] (ref: HR)	3.25	2.03 – 5.19	<0.001
Rank [MR] (ref: HR)	1.47	0.88 – 2.45	0.140
Cycle [No]	0.92	0.65 – 1.30	0.646
Season [Breeding]	0.61	0.47 – 0.79	<0.001
Rate of S.A.	5.27	0.57 – 49.03	0.144
N id	88		
N year	10		
Observations	2122		

**Table S11. Model predicting the rate of Grooming between a focal female and all males.**

We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Grooming (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfGrooming} \sim (1 \mid \text{id}) + (1 \mid \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{Season} + \text{offset}(\log(\text{Tps\_Foc\_IndAllmal}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 \mid \text{id}, \sim 1 \mid \text{year})$ ). The effect of femininity (dcm) is not statistically significant. Model performed with spaMM and results reported with sjPlot (R packages). **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.00	0.00 – 0.00	<0.001
Femininity	1.03	0.62 – 1.71	0.908
Age	1.50	0.93 – 2.44	0.100
Rank [LR] (ref: HR)	0.15	0.05 – 0.41	<0.001
Rank [MR] (ref: HR)	0.23	0.08 – 0.64	0.005
Cycle [No]	0.49	0.28 – 0.89	0.018
Season [Breeding]	1.70	0.93 – 3.09	0.085
N id	88		
N year	10		
Observations	2122		



**Table S12. Model predicting the rate of Grooming between a focal female and all females.** We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Grooming (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfGrooming} \sim (1 | \text{id})++ (1 | \text{year}) + \text{Femininity} + \text{Cycle} + \text{Age} + \text{Rank} + \text{Season} + \text{offset}(\log(\text{Tps\_Foc\_IndAllfem}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 | \text{id}, \sim 1 | \text{year})$ ). The effect of femininity (dcm) is not statistically significant. Model performed with spaMM and results reported with sjPlot (R packages). **Related to Figure 3.**

<i>Predictors</i>	<i>Incidence Rate Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.00	0.00 – 0.00	<0.001
Femininity	0.89	0.74 – 1.06	0.197
Age	0.86	0.73 – 1.02	0.075
Rank [LR] (ref: HR)	0.86	0.62 – 1.20	0.381
Rank [MR] (ref: HR)	0.72	0.51 – 1.01	0.054
Cycle [No]	1.91	1.42 – 2.56	<0.001
Season [Breeding]	0.70	0.58 – 0.84	<0.001
N id	88		
N year	10		
Observations	2122		

**Table S13. Summary of dcm and dcf estimates with p-values for each behavior by males and females: Bold text indicates significant effects Related to Figure 3.**

		From Males	From Females
<b>dcm</b>	<b>MO</b>	<b>0.61 (p=0.040)</b>	
	<b>AG</b>	<b>0.75 (p=0.013)</b>	<b>0.70 (p=0.005)</b>
	<b>SCAN</b>	<b>0.84 (p=0.006)</b>	<b>0.90 (p=0.014)</b>
	<b>AG avec SCAN</b>	<b>0.82 (p=0.070)</b>	<b>0,78 (p=0.053)</b>
	<b>TO</b>	1.03 (p=0.908)	0.89 (p=0.197)
<b>dcf</b>	<b>MO</b>	1.21 (p=0.245)	
	<b>AG</b>	0.99 (p=0.896)	1.01 (p=0.890)
	<b>SCAN</b>	1.00 (p=0.934)	0.96 (p=0.134)
	<b>AG avec SCAN</b>	0.95 (p=0.467)	1.02 (p=0.797)
	<b>TO</b>	0.96 (p=0.823)	0.98 (p=0.779)

**Table S14.** Model predicting the rate of Aggression given by females with femininity score toward males. We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Aggression (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfAggression} \sim (1 | \text{id}) + (1 | \text{year}) + \text{Femininity} + \text{Age} + \text{Rank} + \text{offset}(\log(\text{Nb\_Scan\_IndAllmal}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 | \text{id}, \sim 1 | \text{year})$ ). The effect of femininity (dcm) is statistically significant and positive: females aggress more females when they are less feminine. Model performed with spaMM and results reported with sjPlot (R packages). **Related to Figure 4.**

Predictors	Incidence Rate Ratios	CI	p
(Intercept)	0.00	0.00 – 0.00	<0.001
Mean dcM	0.78	0.62 – 0.99	0.041
age	0.97	0.78 – 1.21	0.772
rang [LR] (ref: HR)	0.32	0.21 – 0.49	<0.001
rang [MR] (ref: HR)	0.44	0.29 – 0.67	<0.001
N id	88		
N year	10		
Observations	2239		

**Table S15.** Model predicting the rate of Aggression given by females with femininity score toward females. We fitted a linear mixed model (estimation of fixed effects by Laplace ML approximation (p\_v.)) to predict the rate of Aggression (number of events with the number of the observations as an offset; for predictions, the offset is set to 1). Formula:  $\text{NbOfAggression} \sim (1 \mid \text{id}) + (1 \mid \text{year}) + \text{Femininity} + \text{Age} + \text{Rank} + \text{offset}(\log(\text{Nb\_Scan\_IndAllfem}))$ . The model included id and year as random effects (formula:  $\text{list}(\sim 1 \mid \text{id}, \sim 1 \mid \text{year})$ ). The effect of femininity (dcm) is statistically significant and positive: females aggress more males when they are less feminine. Model performed with spaMM and results reported with sjPlot (R packages). **Related to Figure 4.**

Predictors	Incidence Rate Ratios	CI	p
(Intercept)	0.00	0.00 – 0.00	<0.001
Mean dcm	0.74	0.59 – 0.93	0.010
age	0.91	0.73 – 1.13	0.392
rang [LR] (ref: HR)	0.33	0.22 – 0.49	<0.001
rang [MR] (ref: HR)	0.44	0.29 – 0.67	<0.001
N id	88		
N year	10		
Observations	2239		

## REFERENCES

- Rhodes, G., Simmons, L.W., and Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evol. Hum. Behav.* 26, 186–201.
- Hamermesh, D.S. (2013). *Beauty Pays: Why Attractive People Are More Successful* (Economics Books).
- Biddle, J.E., and Hamermesh, D.S. (1995). Beauty, Productivity and Discrimination: Lawyers' Looks and Lucre (NBER Working Papers).
- Ahola, A.S., Christianson, S.Å., and Hellström, Å. (2009). Justice needs a blindfold: Effects of gender and attractiveness on prison sentences and attributions of personal characteristics in a judicial process. *Psychiatr. Psychol. Law* 16, S90–S100.
- Perrett, D.I., Lee, K.J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D.M., Henzi, S.P., Castles, D.L., and Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887.
- Rhodes, G., Chan, J., Zebrowitz, L.A., and Simmons, L.W. (2003). Does sexual dimorphism in human faces signal health? *Proc. Biol. Sci.* 270 (Suppl 1), S93–S95.
- Fraccaro, P.J., Feinberg, D.R., DeBruine, L.M., Little, A.C., Watkins, C.D., and Jones, B.C. (2010). Correlated male preferences for femininity in female faces and voices. *Evol. Psychol.* 8, 447–461.
- Smith, M.J.L., Perrett, D.I., Jones, B.C., Cornwell, R.E., Moore, F.R., Feinberg, D.R., Boothroyd, L.G., Durrani, S.J., Stirrat, M.R., Whiten, S., et al. (2006). Facial appearance is a cue to oestrogen levels in women. *Proc. Biol. Sci.* 273, 135–140.
- Law Smith, M.J., Deady, D.K., Moore, F.R., Jones, B.C., Cornwell, R.E., Stirrat, M., Lawson, J.F., Feinberg, D.R., and Perrett, D.I. (2012). Maternal tendencies in women are associated with estrogen levels and facial femininity. *Horm. Behav.* 61, 12–16.
- Wen, F., Zuo, B., Ma, S., Xu, Y., Coley, J.D., and Wang, Y. (2020). Do We See Masculine Faces as Competent and Feminine Faces as Warm? Effects of Sexual Dimorphism on Facial Perception. *Evol. Psychol.* 18, 1474704920980642.
- Owen, H.E., Halberstadt, J., Carr, E.W., and Winkielman, P. (2016). Johnny Depp, Reconsidered: How Category-Relative Processing Fluency Determines the Appeal of Gender Ambiguity. *PLoS One* 11, e0146328.
- Neave, N., and Shields, K. (2008). The effects of facial hair manipulation on female perceptions of attractiveness, masculinity, and dominance in male faces. *Pers. Individ. Differ.* 45, 373–377. <https://doi.org/10.1016/j.paid.2008.05.007>.
- Rosenfield, K.A., Semple, S., Georgiev, A.V., Maestriperi, D., Higham, J.P., and Dubuc, C. (2019). Experimental evidence that female rhesus macaques perceive variation in male facial masculinity. *R. Soc. Open Sci.* 6, 181415.
- Damon, F., Meary, D., Quinn, P.C., Lee, K., Simpson, E.A., Paukner, A., Suomi, S.J., and Pascalis, O. (2017). Preference for facial averageness: Evidence for a common mechanism in human and macaque infants. *Sci. Rep.* 7, 46303.
- Evolution and the Mechanisms of Decision Making (2012) (Oxford University Press).
- Mendelson, T.C., Fitzpatrick, C.L., Hauber, M.E., Pence, C.H., Rodriguez, R.L., Safran, R.J., Stern, C.A., and Stevens, J.R. (2016). Cognitive Phenotypes and the Evolution of Animal Decisions. *Trends Ecol. Evol.* 31, 850–859.
- Kranz, F., and Ishai, A. (2006). Face perception is modulated by sexual preference. *Curr. Biol.* 16, 63–68.
- Little, A.C., Jones, B.C., Feinberg, D.R., and Perrett, D.I. (2014). Men's strategic preferences for femininity in female faces. *Br. J. Psychol.* 105, 364–381. <https://doi.org/10.1111/bjop.12043>.
- Burriss, R.P., Welling, L.L., and Puts, D.A. (2011). Men's attractiveness predicts their preference for female facial femininity when judging for short-term, but not long-term, partners. *Pers. Individ. Differ.* 50, 542–546. <https://doi.org/10.1016/J.PAID.2010.11.022>.
- Jokela, M. (2009). Physical attractiveness and reproductive success in humans: Evidence from the late 20 century United States. *Evol. Hum. Behav.* 30, 342–350.
- O'Toole, A.J., Deffenbacher, K.A., Valentin, D., McKee, K., Huff, D., and Abdi, H. (1998). The perception of face gender: the role of stimulus structure in recognition and classification. *Mem. Cognit.* 26, 146–160.
- Sczesny, S., Spreemann, S., and Stahlberg, D. (2006). Masculine = Competent? Physical Appearance and Sex as Sources of Gender Stereotypic Attributions. *Swiss J. Psychol. Schweiz. Z. Psychol. Rev. Suisse Psychol.* 65, 15–23. <https://doi.org/10.1024/1421-0185.65.1.15>.
- Banchefsky, S., Westfall, J., Park, B., and Judd, C.M. (2016). But You Don't Look Like A Scientist!: Women Scientists with Feminine Appearance are Deemed Less Likely to be Scientists. *Sex. Roles* 75, 95–109.
- Scheib, J.E., Gangestad, S.W., and Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proc. Biol. Sci.* 266, 1913–1917.
- Jones, B., Little, A., Penton-Voak, I., Tiddeman, B., Burt, D., and Perrett, D. (2001). Facial symmetry and judgements of apparent health: Support for a "good genes" explanation of the attractiveness-symmetry relationship. *Evol. Hum. Behav.* 22, 417–429.
- Tredoux, C. (2002). A direct measure of facial similarity and its relation to human similarity perceptions. *J. Exp. Psychol. Appl.* 8, 180–193.
- Jozwik, K.M., O'Keefe, J., Storrs, K.R., Guo, W., Golan, T., and Kriegeskorte, N. (2022). Face dissimilarity judgments are predicted by representational distance in morphable and image-computable models. *Proc. Natl. Acad. Sci. USA* 119, e2115047119.
- Jones, A., and Jaeger, B. (2019). Biological Bases of Beauty Revisited: The Effect of Symmetry, Averageness, and Sexual Dimorphism on Female Facial Attractiveness. *Symmetry* 11, 279.
- Hill, M.Q., Parde, C.J., Castillo, C.D., Colon, Y.I., Ranjan, R., Chen, J.-C., Blanz, V., and O'Toole, A.J. (2019). Deep convolutional neural networks in the face of caricature. *Nat. Mach. Intell.* 1, 522–529.
- Hopper, W.J., Finklea, K.M., Winkielman, P., and Huber, D.E. (2014). Measuring sexual dimorphism with a race-gender face space. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1779–1788.
- O'Toole, A.J., Castillo, C.D., Parde, C.J., Hill, M.Q., and Chellappa, R. (2018). Face Space Representations in Deep Convolutional Neural Networks. *Trends Cognit. Sci.* 22, 794–809.
- Setchell, J., and Wickings, E.J. (2005). Dominance, Status Signals and Coloration in Male Mandrills (Mandrillus sphinx). <https://doi.org/10.1111/J.1439-0310.2004.01054.X>.
- Dezeure, J., Charpentier, M.J., and Huchard, E. (2022). Fitness effects of seasonal birth timing in a long-lived social primate living in the equatorial forest. *Anim. Behav.* 185, 113–126. <https://doi.org/10.1016/j.anbehav.2022.01.002>.
- Setchell, J.M., and Jean Wickings, E. (2006). Mate Choice in Male Mandrills (Mandrillus sphinx). *Ethology* 112, 91–99.

35. Parkhi, O., Vedaldi, A., and Zisserman, A. (2015). Deep face recognition. In *BMVC 2015 - Proceedings of the British Machine Vision Conference 2015*.
36. Simonyan, K., and Zisserman, A. (2015). Very Deep Convolutional Networks for LargeScale Image Recognition.
37. Ng, H., and Winkler, S. (2014). A Data-Driven Approach to Cleaning Large Face Datasets.
38. Gu, J., Jiang, W., Luo, H., and Yu, H. (2021). An efficient global representation constrained by Angular Triplet loss for vehicle re-identification. *Pattern Anal. Appl.* 24, 367–379.
39. Mitteroecker, P., Windhager, S., Müller, G.B., and Schaefer, K. (2015). The morphometrics of “masculinity” in human faces. *PLoS One* 10, e0118374.
40. Tio, S., Restrepo-Ortiz, C.X., Roura-Torres, B., Sauvadet, L., Harte, M., Charpentier, M.J.E., and Renoult, J.P. (2023). The Mandrillus Face Database: A portrait image database for individual and sex recognition, and age prediction in a non-human primate. *Data Brief* 47, 108939.
41. Ryali, C.K., Goffin, S., Winkelman, P., and Yu, A.J. (2020). From likely to likable: The role of statistical typicality in human social assessment of faces. *Proc. Natl. Acad. Sci. USA* 117, 29371–29380.
42. Cootes, T.F., Edwards, G.J., and Taylor, C.J. (1998). Active appearance models. *Computer Vision – ECCV’98*, 484–498.
43. SCUT-FBP5500: a diverse benchmark dataset for multi-paradigm facial beauty prediction. <https://ieeexplore.ieee.org/abstract/document/8546038>.
44. Kloth, N., Damm, M., Schweinberger, S.R., and Wiese, H. (2015). Aging affects sex categorization of male and female faces in opposite ways. *Acta Psychol.* 158, 78–86.
45. Puts, D.A., Bailey, D.H., Cardenas, R.A., Burriss, R.P., Welling, L.L.M., Wheatley, J.R., and Dawood, K. (2013). Women’s attractiveness changes with estradiol and progesterone across the ovulatory cycle. *Horm. Behav.* 63, 13–19.
46. Roberts, S.C., Havlicek, J., Flegr, J., Hruskova, M., Little, A.C., Jones, B.C., Perrett, D.I., and Petrie, M. (2004). Female facial attractiveness increases during the fertile phase of the menstrual cycle. *Proc. Biol. Sci.* 271 (Suppl 5), S270–S272.
47. Marcinkowska, U.M., and Holzleitner, I.J. (2020). Stability of women’s facial shape throughout the menstrual cycle. *Proc. Biol. Sci.* 287, 20192910.
48. Smit, N., Baniel, A., Roura-Torres, B., Amblard-Rambert, P., Charpentier, M.J.E., and Huchard, E. (2022). Sexual coercion in a natural mandrill population. *Peer Community J.* 2, e36. <https://doi.org/10.24072/pcjournal.134>.
49. Stockley, P., and Bro-Jørgensen, J. (2011). Female competition and its evolutionary consequences in mammals. *Biol. Rev.* 86, 341–366. <https://doi.org/10.1111/j.1469185x.2010.00149.x>.
50. Nash, R., Fieldman, G., Hussey, T., Lévêque, J.L., and Pineau, P. (2006). Cosmetics: They Influence More Than Caucasian Female Facial Attractiveness. *J. Appl. Soc. Psychol.* 36, 493–504. <https://doi.org/10.1111/j.00219029.2006.00016.x>.
51. Fink, B., Klappauf, D., Brewer, G., and Shackelford, T.K. (2014). Female physical characteristics and intrasexual competition in women. *Pers. Individ. Differ.* 58, 138–141. <https://doi.org/10.1016/j.paid.2013.10.015>.
52. Little, A.C., Jones, B.C., and DeBruine, L.M. (2011). Facial attractiveness: evolutionary based research. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1638–1659. <https://doi.org/10.1098/rstb.2010.0404>.
53. Marcinkowska, U.M., Kozlov, M.V., Cai, H., Contreras-Gardun˜o, J., Dixon, B.J., Oana, G.A., Kaminski, G., Li, N.P., Lyons, M.T., Onyishi, I.E., et al. (2014). Cross-cultural variation in men’s preference for sexual dimorphism in women’s faces. *Biol. Lett.* 10, 20130850. <https://doi.org/10.1098/rsbl.2013.0850>.
54. Little, A.C., Cohen, D.L., Jones, B.C., and Belsky, J. (2006). Human preferences for facial masculinity change with relationship type and environmental harshness. *Behav. Ecol. Sociobiol.* 61, 967–973.
55. Rosenthal, G.G. (2018). Evaluation and hedonic value in mate choice. *Handling editor: Becky Fuller. Curr. Zool.* 64, 485–492.
56. Basolo, A.L. (1998). Evolutionary change in a receiver bias: a comparison of female preference functions. *Proc. Biol. Sci.* 265, 2223–2228.
57. Wong, B.B.M., and Rosenthal, G.G. (2006). Female Disdain for Swords in a Swordtail Fish. *Am. Nat.* 167, 136–140. <https://doi.org/10.1086/498278>.
58. Halberstadt, J. (2006). The generality and ultimate origins of the attractiveness of prototypes. *Pers. Soc. Psychol. Rev.* 10, 166–183.
59. Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* 57, 199–226.
60. Rosch, E., and Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognit. Psychol.* 7, 573–605. [https://doi.org/10.1016/00100285\(75\)90024-9](https://doi.org/10.1016/00100285(75)90024-9).
61. Fiala, V., Trebický, V., Pazhoohi, F., Leongomez, J.D., Turecek, P., Saribay, S.A., Akoko, R.M., and Kleisner, K. (2021). Facial attractiveness and preference of sexual dimorphism: A comparison across five populations. *Evolut. Hum. Sci.* 3, 1–37.
62. Sanchez-Pages, S., Rodriguez-Ruiz, C., and Turiegano, E. (2014). Facial masculinity: how the choice of measurement method enables to detect its influence on behaviour. *PLoS One* 9, e112157.
63. Reber, R., Wurtz, P., and Zimmermann, T.D. (2004). Exploring “fringe” consciousness: The subjective experience of perceptual fluency and its objective bases. *Conscious. Cognit.* 13, 47–60. [https://doi.org/10.1016/s10538100\(03\)00049-7](https://doi.org/10.1016/s10538100(03)00049-7).
64. Halberstadt, J., and Rhodes, G. (2000). The attractiveness of nonface averages: implications for an evolutionary explanation of the attractiveness of average faces. *Psychol. Sci.* 11, 285–289.
65. Rubenstein, A.J., Kalakanis, L., and Langlois, J.H. (1999). Infant preferences for attractive faces: a cognitive explanation. *Dev. Psychol.* 35, 848–855.
66. Halberstadt, J. (2003). The Hedonic Marking of Processing Fluency: Implications for Evaluative Judgment. *The Psychology of Evaluation*, 205–234. <https://doi.org/10.4324/9781410606853-14>.
67. Winkelman, P., Halberstadt, J., Fazendeiro, T., and Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychol. Sci.* 17, 799–806.
68. Whitfield, T.W.A., and Slatter, P.E. (1979). The effects of categorization and prototypicality on aesthetic choice in a furniture selection task. *Br. J. Psychol.* 70, 65–75.
69. Ghirlanda, S., Jansson, L., and Enquist, M. (2002). Chickens prefer beautiful humans. *Hum. Nat.* 13, 383–389.
70. Renoult, J.P., and Mendelson, T.C. (2019). Processing bias: extending sensory drive to include efficacy and efficiency in information processing. *Proc. Biol. Sci.* 286, 20190165. <https://doi.org/10.1098/rspb.2019.0165>.

71. Muth, C., Hesslinger, V.M., and Carbon, C.(2015). The appeal of challenge in the perception of art: how ambiguity, solvability of ambiguity, and the opportunity for insight affect appreciation. *Psychol. Aesthet. Creat. Arts* 9, 206–216. <https://doi.org/10.1037/A0038814>.
72. Albrecht, S., and Carbon, C.-C. (2014). The Fluency Amplification Model: fluent stimuli show more intense but not evidently more positive evaluations. *Acta Psychol.* 148, 195–203.
73. Kim, Y., Hwang, I., and Cho, N.I. (2017). Convolutional neural networks and training strategies for skin detection. In 2017 IEEE International Conference on Image Processing (ICIP) (IEEE). <https://doi.org/10.1109/icip.2017.8297017>.
74. Khan, K., Attique, M., Khan, R.U., Syed, I., and Chung, T.-S. (2020). A Multi-Task Framework for Facial Attributes Classification through End-to-End Face Parsing and Deep Convolutional Neural Networks. *Sensors* 20, 328.
75. Baek, S., Song, M., Jang, J., Kim, G., and Paik, S.-B. (2019). Spontaneous generation of face recognition in untrained deep neural networks. Preprint at bioRxiv. <https://doi.org/10.1101/857466>.
76. Yamins, D.L.K., Hong, H., Cadieu, C.F., Solomon, E.A., Seibert, D., and DiCarlo, J.J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl. Acad. Sci. USA* 111, 8619–8624.
77. Kriegeskorte, N. (2015). Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. *Annu. Rev. Vis. Sci.* 1, 417–446.
78. Raman, R., and Hosoya, H. (2020). Convolutional neural networks explain tuning properties of anterior, but not middle, faceprocessing areas in macaque inferotemporal cortex. *Commun. Biol.* 3, 221–314.
79. Lindsay, G.W. (2021). Convolutional Neural Networks as a Model of the Visual System: Past, Present, and Future. *J. Cognit. Neurosci.* 33, 2017–2031.
80. Jozwik, K.M., Kriegeskorte, N., Storrs, K.R., and Mur, M. (2017). Deep Convolutional Neural Networks Outperform Feature-Based But Not Categorical Models in Explaining Object Similarity Judgments. *Front. Psychol.* 8, 1726.
81. Sheehan, M.J., and Tibbetts, E.A. (2011). Specialized face learning is associated with individual recognition in paper wasps. *Science* 334, 1272–1275.
82. Charpentier, M.J.E., Harte, M., Poirotte, C., de Bellefon, J.M., Laubi, B., Kappeler, P.M., and Renoult, J.P. (2020). Same father, same face: Deep learning reveals selection for signaling kinship in a wild primate. *Sci. Adv.* 6, eaba3274. <https://doi.org/10.1126/sciadv.aba3274>.
83. Ma, D.S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behav. Res. Methods* 47, 1122–1135.
84. Schroff, F., Kalenichenko, D., and Philbin, J.(2015). FaceNet: A unified embedding for face recognition and clustering. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (IEEE). <https://doi.org/10.1109/cvpr.2015.7298682>.
85. `tfa.losses.TripletSemiHardLoss` TensorFlow. [https://www.tensorflow.org/addons/api\\_docs/python/tfa/losses/TripletSemiHardLoss](https://www.tensorflow.org/addons/api_docs/python/tfa/losses/TripletSemiHardLoss).
86. Kingma, D.P., and Ba, J. (2014). Adam: A Method for Stochastic Optimization. <https://doi.org/10.48550/arXiv.1412.6980>.
87. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). TensorFlow: A System for Large-Scale Machine Learning. <https://doi.org/10.48550/arXiv.1605.08695>.
88. Peignot, P., Charpentier, M.J., Bout, N., Bourry, O., Massima, U., Dosimont, O., Terramorsi, R., and Wickings, E.J. (2008). Learning from the first release project of captive-bred mandrills *Mandrillus sphinx* in Gabon. *Oryx* 42, 122–131.
89. Altmann, J. (1974). *Observational Study of Behavior: Sampling Methods*.
90. Galbany, J., Romero, A., Mayo-Aleson, M., Itsoma, F., Gamarra, B., Pe´rez-Pe´rez, A., Willaume, E., Kappeler, P.M., and Charpentier, M.J.E. (2014). Age-related tooth wear differs between forest and savanna primates. *PLoS One* 9, e94938.
91. Poirotte, C., Massol, F., Herbert, A., Willaume, E., Bomo, P.M., Kappeler, P.M., and Charpentier, M.J.E. (2017). Mandrills use olfaction to socially avoid parasitized conspecifics. *Sci. Adv.* 3, e1601721.
92. Setchell, J.M., and Jean Wickings, E. (2005). Dominance, Status Signals and Coloration in Male Mandrills (*Mandrillus sphinx*). *Ethology* 111, 25–50. <https://doi.org/10.1111/j.14390310.2004.01054.x>.
93. Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neurosci. Biobehav. Rev.* 30, 319–345.
94. Batrinos, M.L. (2012). Testosterone and aggressive behavior in man. *Int. J. Endocrinol. Metabol.* 10, 563–568. <https://doi.org/10.5812/ijem.3661>.

---

## CHAPITRE 3

---

---

### Causes et conséquences de la variation de féminité faciale chez les jeunes mâles mandrills

---



© Berta Roura-Torres



## **Manuscript 3**

---

*Causes and Consequences of variations in  
Facial Feminization in Juvenile Male  
Mandrills*

---

Sonia Tio, Michel Raymond, George Havill, Guillaume Lespagnol, Marianne Paisley,  
Julien P. Renoult<sup>†</sup>, Marie J.E. Charpentier<sup>†</sup>

<sup>†</sup>Equal contribution

**(Article in preparation)**

## **Abstract**

Facial feminization, though extensively studied in humans, remains less understood in other species. In this study, we investigated both the causes and consequences of facial feminization in juvenile male mandrills (*Mandrillus sphinx*). We primarily examined two theories: the Trivers-Willard hypothesis, which posits that maternal physical conditions can influence the sex and subsequent feminization of offspring, and the Fraternal Birth Order (FBO) effect, which associates heightened male feminization with an increasing number of older brothers due to maternal immunity. Our findings indicated that younger maternal age and a reduced number of older sisters were significant predictors of enhanced facial femininity in these male juveniles. Our results support the Trivers-Willard theory because younger mothers, who have not yet reached full physical maturity, may be energetically constrained, leading to the production of female offspring or of more feminized males. These feminized males did not display, however, expected female-like social behaviors. This study highlights the need to consider human-centric findings within a broader evolutionary framework to broaden our understanding of mother-offspring early interactions.

## **Introduction**

Individualized faces and elaborated facial expressions in animal species, especially in the primate order, result from a combination of genetic, hormonal and environmental factors. Facial traits, and their extended range of variation observed across individuals, may also provide significant information about emotions but also physiological or behavioral characteristics of their carriers. Within this wide panel of facial diversities, feminization of features typically representative of males has been the focus of numerous studies and has long attracted the interest of researchers. Indeed, in human males, facial feminization has often been explored in relation to sexual orientation. Homosexuality in men, for example, is associated to a set of physical, behavioral, and cognitive traits often typically associated to women. For instance, homosexual men show some degree of feminization in the proportion of long bones [1] and a "feminized" facial structure with rounder jaws, larger distances between mouth and eyes, and shorter noses [2] than generally observed in heterosexual men. Additional studies have further highlighted feminization of voices [3] and cognitive skills, like spatial cognition and verbal fluency, in homosexual men [4,5]. Recently, the

use of Deep Neural Networks (DNN) allowed to accurately determine sexual orientation from men faces, outperforming human judges [6].

Prenatal androgen exposure vs. the maternal immunity against male fetuses have been both proposed as candidate physiological mechanisms to explain feminization of male-typical traits. The first mechanism posits that male feminization may be triggered by a deficiency in androgen during the sexual differentiation phase in fetal brain development, often as a repercussion of maternal stress during pregnancy. It is well-documented that testosterone exposure, beginning around the seventh gestational week, critically influences the development of genitalia and secondary characteristics, including facial morphology [7] and brain structure development [8]. An increased exposure could, in turn, causes masculinization in females [9]. In male rats, the relationship between androgens and sexual preferences demonstrates an inverted-U pattern [10]. Specifically, both low and high androgen exposures during development lead to an attraction to males. However, low androgen exposure results in female-typical behaviors (e.g. lordosis), whereas high androgen exposure leads to male-typical behaviors (e.g. mounting and thrusting). In humans, measuring exposure to prenatal androgen is challenging. Hence, indirect measures like the 2D:4D finger length ratio are often used as proxies of physiological feminization. The 2D:4D ratio's connection to homosexual preferences remains, however, debated, with some studies showing a higher (feminized) ratio in homosexual men than heterosexual men [11–13], while others indicate the opposite [14,15].

The Trivers-Willard hypothesis is one evolutionary pathway supporting the prenatal androgen exposure mechanism. This theory, as early proposed, suggests that maternal physical conditions can influence the sex and subsequent traits of the offspring [16]. It posits that optimally conditioned females might exhibit a bias towards producing male offspring, whereas those in suboptimal conditions might be biased towards female offspring because testosterone is energetically costly to produce [17,18]. For example, women bearing male fetuses are known to intake roughly 10% more energy compared to women pregnant with female fetuses [19]. Additionally, historical data from Bangladesh recorded a skewed sex ratio favoring females during periods of famine [20]. Meanwhile, studies in Ethiopia revealed that nutritionally healthy women (assessed via weight) predominantly gave birth to male offspring [21]. In non-human species, dairy cows fed with a nutritionally rich diet give birth to a

significantly higher proportion of bulls [22] and studies on rodents echoed similar findings [23]. Such data accentuates the overarching role that environmental resources play in determining offspring traits including their sex. Dominance rank has been also suggested as an indirect indicator of maternal condition with high-ranking females, in a wide range of species, often enjoying privileged access to food [24] and experiencing fewer injuries [25]. As such, female red deer of higher ranks tend to produce more male offspring, aligning with the predictions of Trivers and Willard [26]. In primates, results are more contradictory: in some species, high-ranking females tend to produce more female offspring [27], while in others, the reverse is observed [28,29]. These conflicting results can be explained by the females' tendency to establish stable, linear, matrilineal dominance hierarchies in some primates [30,31]. In such species, female offspring often inherit their mother's rank. High-ranking females should thus produce more high-ranking, philopatric daughters. Conversely, male offspring, who typically disperse from their natal groups, may enjoy less benefits of being born from high-ranking mothers. Other maternal traits may also impact the production of males: in humans, the probability to give birth to daughters rather than sons is higher in younger women and those close to menopause. This is potentially linked to variations in metabolic energy and social support [32].

The second physiological mechanism proposed to explain an elevated feminization observed in some human males is the "maternal immune hypothesis". This mechanism posits that, with each successive male pregnancy, a mother accumulates antibodies against male-specific proteins such as NLGN4Y (presumed to influence synaptic functioning in the male fetal brain; [33,34]). These accumulated antibodies may potentially disrupt the usual brain masculinization processes, likely leading to a more feminized neural architecture which, in turn, might influence sexual orientation and gendered behaviors. This mechanism was originally proposed to explain the "fraternal birth order effect", where the likelihood of a man being homosexual rises with the number of older brothers he has [35]. Notably, a fraternal birth order effect is not observed among adopted brothers, suggesting that environmental, familial aspects, such as education, do not determine sexual orientation [36]. In addition, in one study, the fraternal birth order effect has been related to enhanced facial femininity in homosexual men by showing first that manipulated faces resembling homosexual averages are perceived as more feminized [37]. They further found that homosexual men exhibit more older brothers and, to

some extent, older sisters, implying the presence of an older sibling effect, even after controlling for sibship size. While this study offers valuable insights for humans, it remains unclear whether these findings can be generalized to other species.

Here, we first examine the variation in facial feminization in juvenile male mandrills (*Mandrillus sphinx*) and test the two concurrent hypotheses to explain such variation (namely the Trivers-Willard hypothesis vs. fraternal birth order (FBO) hypothesis). Additionally, we investigated the potential association between facial feminization and behavioral patterns, hypothesizing that males exhibiting more feminized features would demonstrate behaviors typically observed in females, in contrast to their more masculinized peers. Nonhuman primates share many developmental and behavioral traits with humans, including an extended developmental lifespan, complex social relationships, and often pronounced sexual dimorphism and sexually-differentiated behaviors. Mandrills are characterized by an extreme sexual dimorphism, with adult males exhibiting conspicuous facial coloration and being 3.5 times heavier than females. Mandrills' complex behavioral repertoire highlights elaborated social structures. Our study population, subjected to longitudinal studies since 2012, provides an extensive database of over 20,000 facial images [38]. This database permits rigorous assessment of facial femininity metrics in immature male mandrills, prior to the marked growth acceleration they undergo around puberty and associated sexual differentiation characteristic of the species. Within this population, long-term demographical records enable accurate birth order determination in these males. Additionally, our dataset allowed to obtain proxies for maternal conditions, such as age and dominance rank. This article specifically targets juvenile mandrills.

Our first objective is to elucidate the factors driving variations in facial femininity in male juvenile mandrills. In the context of the Trivers-Willard hypothesis, we predict that firstborn sons of energetically constrained mothers—characterized by youth or lower social rank—will present more feminine facial attributes. By contrast, the Fraternal Birth Order (FBO) hypothesis and its associated physiological mechanism, the maternal immune hypothesis, predicts that males with a greater number of elder siblings, especially older brothers, should exhibit more pronounced facial feminization. Our second objective investigates if males exhibiting increased

facial femininity also demonstrate behaviors typically attributed to females, including elevated grooming and reduced play and aggression.

## **Materials**

### **Study System and Behavioral Data Collection**

This study was conducted on a natural population of mandrills, originally established in the private Lékédi Park, Gabon, through the release of 36 captive individuals in 2002, and further supplemented by 29 individuals released in 2006 [39]. Since 2003, wild immigrating males joined the group to reproduce and in Dec 2022, the group was composed of c.a. 300 individuals, including only 3 captive-born females. The Mandrillus Project was initiated in 2012 to study the socio-ecology of this habituated mandrill population [40,41]

We used behavioral and life-history information recorded between January 2012 and December 2022 on 133 female and 161 male mandrills aged between 2 and 5 years. Behavioral observations were made by trained field assistants blind to the focus of this study: during 5-min focal observations [42], social interactions between focal individuals and their groupmates are recorded. We used a total of 22,183 focal observations (1,457 hours). We specifically focused on mandrills' behaviors related to playing (e.g. 'playing chase', 'invitation to play', 'rocking game', 'mutual play head', 'solitary play' and 'play head'), aggression ('chases', 'bites' and 'ground-slapping'), and grooming. Birth dates were known for 68 individuals (from 0 to a few days of approximation). Birth dates were estimated for 226 individuals thanks to body condition and tooth eruption and wear patterns [43] We focused on 37 male mandrills for whom we had all the following data available: facial images, precise pedigree, and detailed behavioral observations, allowing us to ascertain both birth order and femininity metrics within this group. Of the 37 males analyzed, maternal data, encompassing age and dominance rank, is available for 35 individuals. Female dominance ranks were deduced from the outcomes of approach-avoidance interactions across all adult females, observed either during focal or ad libitum observations and were calculated using the standardized David's score [44]. Adult females were categorized into three similarly sized rank classes (HR-high rank, MR-medium rank, and LR-low rank) as female rank is stable in this matrilineal society [31].

## Mandrillus Face Database (MFD):

The Mandrillus Face Database (MFD) serves as our main source for facial images, comprising photographs we collected from January 2012 to December 2022 (see for details: [38]). We used a dataset comprising 12,706 images (5,802 for males and 6,904 for females, across 242 individuals aged 2 to 6) for the assessment of facial femininity. We evaluated the distance to the female centroid, termed as dcf, as an indicator of facial femininity or sexual dimorphism in males. A shorter distance to the female centroid -or a lower dcf value- indicates a more feminine appearance in male faces [45].

For our analyses, we focused on images collected on the 37 males aged 2-5 years, leading to a dataset of 1,371 images (averaging  $97 \pm 31$  images per individual). Before proceeding with our main analysis, we initially investigated the relationship between the quantity of images available for each male and the respective dcf scores garnered per age-class. Notably, between the ages of 2 and 5, these males exhibit significant morphological alterations. The sample for this initial correlation analysis consists of 95 male-years observations. Only a slight correlation between averaged dcf and the number of pictures was observed, suggesting that the number of images has a minimal influence on the calculated femininity score for each study individual (for further details, see: Figure S1).

In subsequent analyses, we calculated the ‘mean dcf’ (mean femininity) for each male, categorized by age class. Intra-individual variability in femininity scores across different ages in males was assessed using pairwise Wilcoxon signed-rank tests.

## Statistics

### Fraternal Birth Order (FBO) Hypothesis Testing

To investigate the relationship between facial femininity and the FBO, we performed Generalized Linear Mixed Models (GLMMs) with a Gaussian distribution. In our models, we treated the ‘mean dcf’ as the response variable. The three factors—‘birth order’, ‘number of old brothers’ (i.e FBO), and ‘number of old sisters’ —were assessed separately as predictor variables due to their substantial correlation (see Table S2). We also incorporated other explanatory variables, including the mother's

rank and the age-class of the male, as explanatory variables. Because study males contributed to several datapoints (different age-class), we included the identity of each male as a random effect. Furthermore, we ensured that the residuals of each model fit to a normal distribution using the Shapiro test. Notably, two individuals had to be excluded from the analysis due to the unavailability of maternal identity. Consequently, the dataset consisted of 35 individuals from the original 37, resulting in a total of 88 observations.

### Fraternal Birth Order (FBO) Hypothesis Testing

To explore the association between facial femininity and mother's age, we used a GLMM with a Gaussian distribution. We treated the 'mean dcf' within each age class as the response variable, while considering the mother's age and rank, as well as the male's age-class, as three predictor variables. The identity of each male mandrill was introduced as a random effect. As previously mentioned, the dataset comprised 35 individuals out of the original 37.

### Behavioral Sex Differences and Age-related Variations in Mandrills

We first identified sex-typical behaviors among mandrills. In each age class for both male and female subjects, we determined behavior rates for play, grooming, and aggression by dividing the total event count by the corresponding total focal time. We evaluated disparities across sex and age classes using t-tests or Welch's tests. The choice between the Student's t-test (assuming equal variances) and the Welch's t-test (addressing unequal variances) depended on the variance outcomes.

### Influence of facial femininity on Behavioral Sex Differences

We finally explored the relationships between facial femininity in the 37 study males and those behaviors that were found to vary across sexes. We used Generalized Linear Mixed Models (GLMMs) with a negative binomial error distribution, augmented by a zero-inflated model, to address the abundance of zeros in the counts of the studied social behaviors. In our models, we included the femininity (mean distance to the female centroid dcf) as a predictor and further considered categorical age variables (expressed in years). We included male's identity as a random effect to account for repeated measurements of the same individuals across years. Total observation times were considered as an offset in our models.



All statistical analyses are conducted in R Studio. We use “t.test” for Student and Welch test and ‘lme4’ and ‘glmmTMB’ ( with *nbinom2* parameters) packages for the Gaussian and Negative Binomial GLMMs, respectively. To validate the Gaussian models, the distribution of residuals is examined and we checked their normality with the Shapiro test. The ‘ggplot2’ and ‘sjPlot’ packages were used for data visualization and to assess the models' effect sizes and confidence intervals, respectively.

## **Results**

### **Determinants of facial feminization of males juveniles**

#### Intra-Individual Stability in Facial Femininity Scores Across Age Categories in Male Mandrills.

Analyses demonstrated a consistent intra-individual stability in femininity scores across the different studied age-classes. Specifically, there was no significant difference in femininity scores when comparing 2 years old to 3 years old , 2 years old to 4 years old, 2 years old to 5 years , 3 years old to 4 years old, 3 years old to 5 years old and 4 years old to 5 years old (see Table S1) . This consistent intra-male stability suggests that the femininity measurement for a given male remains robust across age categories.

#### Results of Fraternal Birth Order (FBO) Hypothesis Analysis

Contrary to our expectations, we showed no effect of the birth order and the number of older brothers on male femininity scores, although males with higher birth order tend to be less feminine (estimate = 0.21, 95% CI [-0.03, 0.45], p-value = 0.084; Model 1 in Table 1). Similarly, males having an increased count of older brothers did not manifest heightened femininity, with the observed effect being statistically insignificant (estimate = 0.31, 95% CI [-0.18, 0.79], p-value = 0.211; Model 2 in Table 1). By contrast, having more older sisters was associated with male mandrills being less feminine (estimate = 0.46, 95% CI [0.02, 0.91], p-value = 0.041; see Model 3 in Table 1).

### Results of Trivers-Willard Hypothesis Analysis

We found a significant positive association between maternal age and mean dcf, suggesting that males born to older mothers showed less feminine facial features (estimate = 0.18, 95% CI [0, 0.37], p-value = 0.048; see Model 4 in Table 2).

**Table 1. Analysis of facial femininity in male mandrills based on birth order, number of older brothers, and number of older sisters.** Results of mixed models analyzing facial femininity ('mean dcf') as influenced by birth order (Model 1), number of older brothers (Model 2), and number of older sisters (Model 3) are displayed in the table. The models' estimates, 95% confidence intervals (CI), and p-values are provided for each predictor. The reference category for categorical predictors is mentioned within parentheses. Significant effects are shown in bold. Random effects (identity of males) are broken down into residual variance ( $\sigma^2$ ), variance of random intercepts ( $\tau_{00 \text{ id}}$ ), the Intraclass Correlation Coefficient (ICC) which signifies between-group differences, unique identifier count (N id), total data entries (Observations), and variance elucidated by fixed effects and combined effects (Marginal R<sup>2</sup> / Conditional R<sup>2</sup>).

**Response variable : mean dcf**

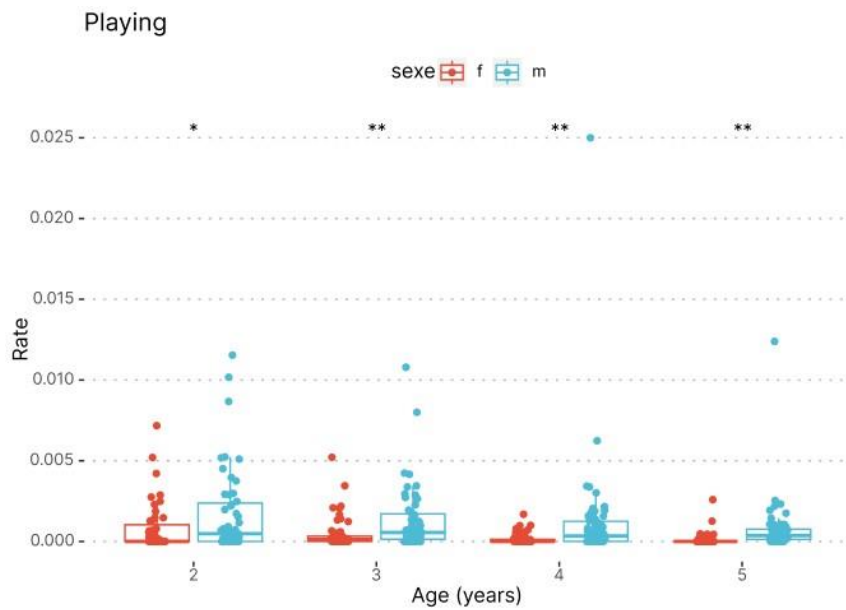
	<i>Predictors</i>	<i>Estimates</i>	<i>CI</i>	<i>p</i>
<b>Model 1 – Effect of birth order</b>	(Intercept)	8.69	7.32 – 10.06	<0.001
	birth order	0.21	-0.03 – 0.45	0.084
	age year categ [3] (ref: age year categ [2])	-0.47	-1.49 – 0.55	0.363
	age year categ [4] (ref: age year categ [2])	-0.55	-1.54 – 0.44	0.272
	age year categ [5] (ref: age year categ [2])	0.19	-0.88 – 1.26	0.728
	mother rank [LR] (ref: HR)	0.18	-1.12 – 1.48	0.784
	<b>Random Effect</b>			
	σ <sup>2</sup>	2.05		
	τ <sub>00</sub> id	2.28		
	ICC	0.53		
	N id	35		
	Observations	88		
	Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.083 / 0.535		
<b>Model 2 – Effect of number of old brothers</b>	(Intercept)	9	7.71 – 10.28	<0.001
	number of older brothers	0.31	-0.18 – 0.79	0.211
	age year categ [3] (ref: age year categ [2])	-0.47	-1.50 – 0.56	0.365
	age year categ [4] (ref: age year categ [2])	-0.55	-1.54 – 0.45	0.276
	age year categ [5] (ref: age year categ [2])	0.17	-0.90 – 1.25	0.748
	mother rank [LR] (ref: HR)	0.16	-1.18 – 1.49	0.817
	<b>Random Effect</b>			
	σ <sup>2</sup>	2.05		
	τ <sub>00</sub> id	2.43		
	ICC	0.54		
	N id	35		
	Observations	88		
	Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.058 / 0.569		
<b>Model 3 – Effect of number of old sisters</b>	(Intercept)	8.93	7.76 – 10.10	<0.001
	<b>number of older sisters</b>	0.46	0.02 – 0.91	<b>0.041</b>
	age year categ [3] (ref: age year categ [2])	-0.49	-1.51 – 0.54	0.347
	age year categ [4] (ref: age year categ [2])	-0.58	-1.56 – 0.41	0.247
	age year categ [5] (ref: age year categ [2])	0.18	-0.89 – 1.24	0.745
	mother rank [LR] (ref: HR)	0.17	-1.11 – 1.44	0.798
	<b>Random Effect</b>			
	σ <sup>2</sup>	2.05		
	τ <sub>00</sub> id	2.18		
	ICC	0.52		
	N id	35		
	Observations	88		
	Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.108 / 0.569		

**Table 2. Analysis of facial femininity in male mandrills based on mother's age.** Results of mixed models analyzing facial femininity ('mean dcf') as influenced by mother's age (Model 4). The models' estimates, 95% confidence intervals (CI), and p-values are provided for each predictor. The reference category for categorical predictors is mentioned within parentheses. Significant effects are shown in bold. Random effects (identity of males) are broken down into residual variance ( $\sigma^2$ ), variance of random intercepts ( $\tau_{00}$  id), the Intraclass Correlation Coefficient (ICC) which signifies between-group differences, unique identifier count (N id), total data entries (Observations), and variance elucidated by fixed effects and combined effects (Marginal R<sup>2</sup> / Conditional R<sup>2</sup>).

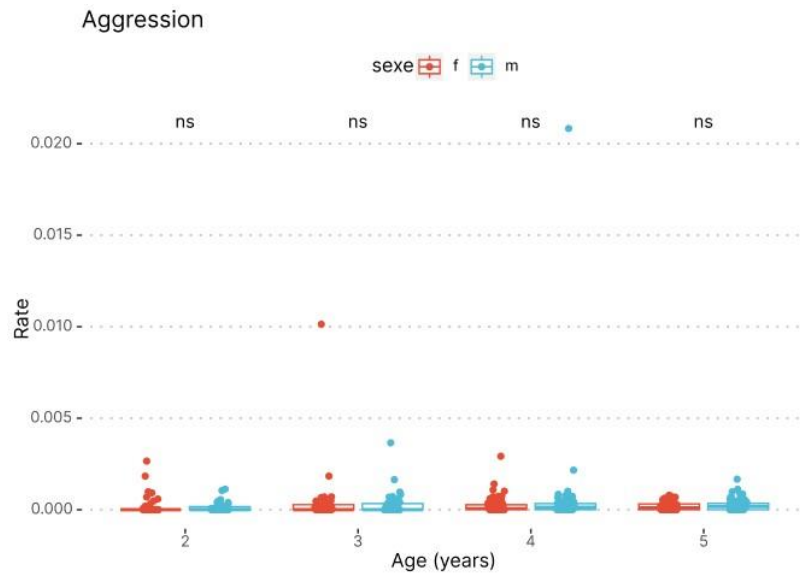
<b>Response variable : mean dcf</b>					
<b>Predictors</b>		<b>Estimates</b>	<b>CI</b>	<b>p</b>	
<b>Model 4 - Effect mothers' age</b>	(Intercept)	7.53	5.37 – 9.69	<b>&lt;0.001</b>	
	<b>mother age</b>	0.18	0 – 0.37	<b>0.048</b>	
	age year categ [3] (ref: age year categ [2])	-0.64	-1.68 – 0.29	0.219	
	age year categ [4] (ref: age year categ [2])	-0.90	-1.94 – 0.13	0.087	
	age year categ [5] (ref: age year categ [2])	0.34	-1.51 – 0.83	0.567	
	mother rank [LR] (ref: HR)	0.14	-1.13-1.41	0.826	
	<b>Random Effect</b>				
	$\sigma^2$	2.05			
	$\tau_{00}$ id	2.18			
	ICC	0.52			
	N id	35			
	Observations	88			
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.108 / 0.569				

### **Behavioral Sex Differences and Age-related Variations in Mandrills**

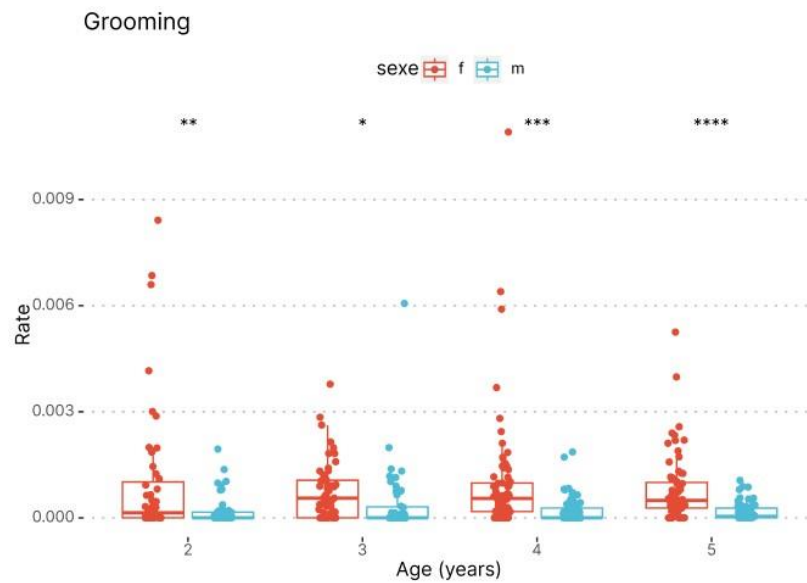
We found that males consistently showed higher playing frequencies than females from ages 2 to 5 (Figure 1). Significant sex disparities were noted at 2 years ( $p=0.04$ ), 3 years ( $p=0.01$ ), 4 years ( $p=0.009$ ), and 5 years ( $p=0.002$ ). Contrary to our prediction, both females and males displayed comparable aggression emission frequencies across all age-classes (Figure 2). Females consistently displayed higher grooming rates relative to males at every age-class examined. Sex disparities in grooming rates were significant at 2 years old ( $p = 3.86e-3$ ), 3 years old ( $p = 2.03e-2$ ), and 4 years old ( $p = 2.76e-4$ ).



**Figure 1. Playing rates across age categories (2, 3, 4, 5) expressed in years and sex.** Within each boxplot, a color distinction was employed, with blue denoting males (m) and red representing females (f). The boxplots were enhanced with overlaid significant difference indicators resulting from t-tests, visually highlighting the statistical comparisons of playing rates between males and females. Statistically significant differences are represented by asterisks (\*) on the boxplots, with the common levels of significance shown as \*( $p < 0.05$ ), \*\*( $p < 0.01$ ), \*\*\*( $p < 0.001$ ) and \*\*\*\*( $p < 0.0001$ ).



**Figure 2. Aggression rates across age categories (2, 3, 4, 5) expressed in years and sex.** Within each boxplot, a color distinction was employed, with blue denoting males (m) and red representing females (f). The boxplots were enhanced with overlaid significant difference indicators resulting from t-tests, visually highlighting the statistical comparisons of aggression rates between males and females. Statistically significant differences are represented by asterisks (\*) on the boxplots, with the common levels of significance shown as \*( $p < 0.05$ ), \*\*( $p < 0.01$ ), \*\*\*( $p < 0.001$ ) and \*\*\*\*( $p < 0.0001$ ).



**Figure 3. Grooming rates by across age categories (2, 3, 4, 5) expressed in years and sex:** Within each boxplot, a color distinction was employed, with blue denoting males (m) and red representing females (f). The boxplots were enhanced with overlaid significant difference indicators resulting from t-tests, visually highlighting the statistical comparisons of grooming rates between males and females. Statistically significant differences are represented by asterisks (\*) on the boxplots, with the common levels of significance shown as \*( $p < 0.05$ ), \*\*( $p < 0.01$ ), \*\*\*( $p < 0.001$ ) and \*\*\*\*( $p < 0.0001$ ).

### Relationships between facial feminization in male mandrills and behavior typically associated with either masculinity or femininity

In our analysis of 'mean dcf', we found no statistically significant effects on the examined behaviors, yet observable trends emerged. Specifically, data suggests that less feminine males might play less (estimate = 1.08, 95% CI [0.92 – 1.26], p-value = 0.354). Similarly, they might also display more aggression (estimate = 1.06, 95% CI [0.96 – 1.18], p-value = 0.258) and engage more in grooming (estimate = 1.06, 95% CI [0.92 – 1.22], p-value = 0.449). However, these trends, though observable, did not reach statistical significance (see Table3).

**Table 3. Analysis of the effect of Facial Femininity on the emission of play, aggression, and grooming.** Results for predictions 4 to 6 regarding number of play, aggression, and grooming in mandrills. The "Predictors" column includes femininity (mean dcf), and age categories (age year categories [2-5] and 2 as age reference). "Estimates" column provides the coefficients indicating the relationship between predictors and dependent variables, with the "CI" (Confidence Interval) column showing the range within which the true value lies at 95% confidence, and "p" representing the p-value. Random effects (identity of males id) are broken down into residual variance ( $\sigma^2$ ), variance of random intercepts ( $\tau_{00}$  id), the Intraclass Correlation Coefficient (ICC) which signifies between-group differences, unique

identifier count (N id), total data entries (Observations), and variance elucidated by fixed effects and combined effects (Marginal R2 / Conditional R2). Boldface is employed to denote significant effects.

<b>Response variable : mean dcf</b>					
	<b>Predictors</b>	<b>Estimates</b>	<b>CI</b>	<b>p</b>	
<b>Model 8- Count of playing</b>	(Intercept)	0.00	0.00 – 0.01	<b>&lt;0.001</b>	
	mean dcf	01.08	0.92 – 1.26	0.354	
	age year categ [3] (ref: age year categ [2])	0.67	0.26 – 1.72	0.409	
	age year categ [4] (ref: age year categ [2])	0.54	0.22 – 1.29	0.162	
	age year categ [5] (ref: age year categ [2])	0.32	0.12 – 0.82	<b>0.018</b>	
	<b>Random Effect</b>				
	$\sigma^2$	18.99			
	$\tau_{00}$ id	0.34			
	ICC	0.02			
	N id	37			
	Observations	83			
Marginal R2 / Conditional R2	0.008 / 0.025				
<b>Model 9- Count of aggression</b>	(Intercept)	0.00	0.00 – 0.00	<b>&lt;0.001</b>	
	mean dcf	1,06	0.96 – 1.18	0.258	
	age year categ [3] (ref: age year categ [2])	2.75	0.77– 9.88	0.366	
	age year categ [4] (ref: age year categ [2])	3.23	0.97– 10.74	0.056	
	age year categ [5] (ref: age year categ [2])	4.15	1.23–14.03	<b>0.022</b>	
	<b>Random Effect</b>				
	$\sigma^2$	17.40			
	$\tau_{00}$ id	0.50			
	ICC	0.03			
	N id	37			
	Observations	83			
Marginal R2 / Conditional R2	0.014 / 0.041				
<b>Model 10- Count of grooming</b>	(Intercept)	0.00	0.00 – 0.00	<b>&lt;0.001</b>	
	mean dcf	01.06	0.92 – 1.22	0.449	
	age year categ [3] (ref: age year categ [2])	1.62	0.48 – 5.45	0.438	
	age year categ [4] (ref: age year categ [2])	1.81	0.60 – 5.41	0.289	
	age year categ [5] (ref: age year categ [2])	1.57	0.50 – 4.93	0.442	
	<b>Random Effect</b>				
	$\sigma^2$	18.44			
	$\tau_{00}$ id	0.37			
	ICC	0.02			
	N id	37			
	Observations	83			
Marginal R2 / Conditional R2	0.003 / 0.022				



## **Discussion**

Our study sought to uncover the causes and consequences of facial femininity in juvenile male mandrills. We identified maternal age and the number of older sisters as predominant factors. Specifically, younger mothers and males with few or no older sisters were more likely to exhibit pronounced facial femininity. Feminized male mandrills did not display, however, social behaviors typically attributed to females.

### **Causes of Facial Feminization Variations**

In this study, we first tested the maternal immune hypothesis and its extension, the Fraternal Birth Order (FBO) hypothesis, as a plausible biological mechanism explaining feminized faces in male non-human primates, a question only explored in humans so far. Prior human studies have shown a positive correlation between the number of older brothers and both homosexual orientation [35] and, to some extent, facial feminization [37]. Further research suggests that mothers accumulate antibodies against male-specific antigens (such as NLGN4Y), following consecutive male pregnancies, which could contribute to the "demasculinization" of subsequent male offspring [46]. Because homosexual behaviors were only scarcely documented in the study mandrill population, we considered male's facial feminization as the principal metric in our analyses, limiting the generalization of our findings. Yet, we found a unexpected, but significant, pattern in mandrills: young males with few older sisters are more feminine on average than those with numerous older sisters. In addition, males born to young females tend to be more feminine than males born to older mothers. While the number of daughters produced is positively correlated to maternal age (as expected), the two effects found on facial feminization are likely to be relatively independent. Indeed, if only maternal age was responsible for a feminization effect (i.e. masculinized offspring in old mothers), then we should have further observed that males with a large number of older brothers being more masculinized, which was not the case.

Taken together, these results are in direct contrast with the FBO hypothesis and rather supports the Trivers-Willard theory on the influence of maternal condition on offspring sex ratio. This hypothesis stipulates that when conditions are favorable and maternal resources are abundant, mothers will produce offspring of the sex that will

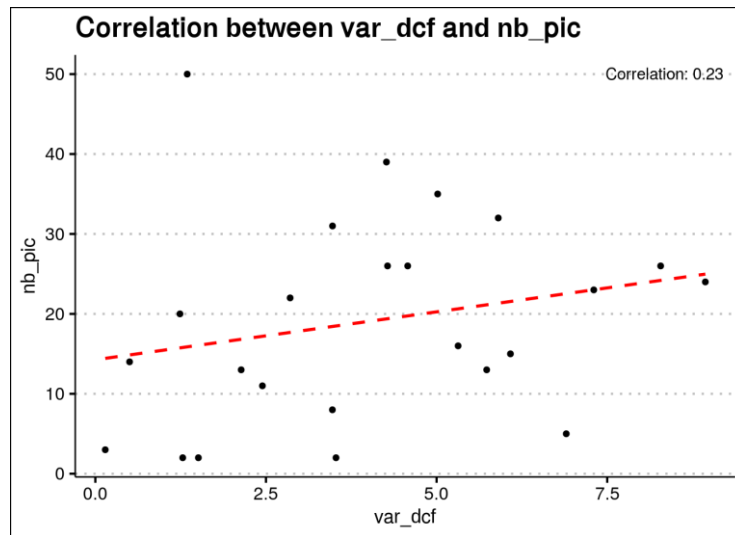
benefit the most from these favorable conditions in terms of reproductive success. In mandrills, if bearing male fetuses is more costly than bearing female fetuses, as observed in several other primate species [19,47,48], then females in good condition only should produce sons. Primiparous female mandrills have not reached yet their full adult body size when giving birth to their first offspring. In our study population, age at first birth in females is on average 4.7 yrs ( $\pm$  0.66) and the youngest female gave birth at 3.2 yrs, almost 4 yrs before reaching a full adult size (at 7; [49]). Consequently, primiparous and young mothers probably did not meet the energy requirements needed to give birth to males, possibly explaining the fact that sons born to young mothers are more feminine than sons born to older mothers. As well, females that gave birth to numerous daughters may have also stored the energy needed to give birth to masculinized sons, as observed. Future research should assess whether older mothers nearing menopause (so with supposedly less energy to invest in their pregnancies) tend to also produce also more feminine male offspring or a greater number of female offspring, and whether female's age at first birth influences the likelihood of having a female offspring, as expected under the Trivers-Willard hypothesis. In addition, other maternal characteristics, such as maternal weight, parasitic infection status, and stress levels are potential interesting routes to explore the relationship between maternal condition and offspring sex ratio and by extension, the feminization of their sons. Intriguingly, in this study, maternal rank did not influence males' facial feminization by contrast to our predictions drawn from the Trivers-Willard hypothesis. Yet, high-ranking females in mandrills are generally in better physical conditions than low-ranking females [50]. As such, they give birth earlier to their first offspring and show shorter inter-birth intervals [51]. While the Trivers-Willard hypothesis focused on maternal condition driving offspring sex ratio, our results indicate a slightly different mechanism. Rather than influencing offspring sex directly, maternal condition seems to impact facial feminization of their male fetuses. Mothers in suboptimal conditions might limit testosterone investment, which is highly costly to females, in male fetuses, leading to pronounced feminization. Further research is thus now needed to fully grasp the complex interplay between maternal rank, condition, and offspring sex ratio in mandrills.

## **Consequences of Facial Feminization Variations**

In this study, we then examined the potential impact of males' facial feminization on their social relationships. We first determined sex-typical behaviors in juvenile mandrills. In primates, pronounced sex-specific behaviors occur at early developmental stages, notably evidenced in play and grooming patterns. Female primates consistently engage in more social grooming than their male counterparts, reflecting the importance of grooming in social bonding and kinship ties (e.g., in Japanese macaques:[52,53]; in rhesus macaques [54,55]). By contrast, male primates exhibit a pronounced preference for playing behavior, especially those of rough-and-tumble nature [55–58]; In male mandrills, an extended adolescent phase reveals a transition in behavior: younger males predominantly engage in playful interactions with their peers. However, as they mature, their behavioral dynamics shift towards heightened aggression and interactions with adults [59].

Contrary to our initial expectations, feminized male mandrills did not exhibit female-typical behaviors: they play and groom their groupmates as often as more masculinized males. The concept of behavioral feminization versus masculinization has its theoretical underpinnings in studies on rodents. Prenatally androgenized guinea pigs show lasting changes in sexually dimorphic behaviours, because in-utero androgen exposure can reshape neural structures [60] In male rat models, prenatal stressors that altered the maternal androgenic environment resulted in modified sexual behaviors, characterized by reduced male copulatory responses and increased female lordotic behaviors [61]. In rhesus macaques, prenatal androgen exposure influences neural pathways affecting sexual and social behaviors too [62]. Elevated levels of these hormones result in masculinized behaviors. For instance, behaviors like playing and mounting show masculinization in juveniles macaques. In adulthood, behaviors such as approaches, sex contacts, and mounts appear masculinized, while sexual solicitations exhibit de-feminization [62]. Considering these findings, future studies should incorporate hormonal assessments to estimate prenatal androgen levels in mothers. Such an in-depth analysis might determine if the identified behavioral patterns in mandrills correlate with these hormonal factors, although non-apparent in facial feminized traits.

## Supplementary Data



**Figure S1. Correlation between the variance in dcf and the number of averaged photos (Variance dcf (var\_dcf)~ number of pictures (nb\_pic)).** Initial observations suggest a minor correlation between these variables, informing potential biases introduced by the number of photos used to compute dcf and dcm.

**Table S1. Pairwise Wilcoxon Signed-Rank Tests for Comparing Facial Femininity Scores (mean dcf) Across Different Age Categories in Male Mandrills: "Age A" and "AgeB"** indicate the first and second age categories being compared. "V-statistic." represents the V-statistic, a measure of the test statistic. "p-value" indicates the statistical significance of the comparison. These comparisons assess the consistency of facial femininity scores across different age categories, with associated statistical measures.

Age A	Age B	p-value	V-statistics
2	3	0.9697266	38
2	4	0.2438965	63
2	5	0.3125000	26
3	4	0.6327286	142
3	5	0.3393555	26
4	5	0.6112213	65

**Table S2. Correlations Coefficient between 'birth order', 'number of old brothers' (i.e FBO), and 'number of old sisters'**

<b>Correlations</b>	<b>R</b>	<b>p</b>
<b>birth order ~ number of old brothers</b>	0.96	2.2 x 10 <sup>16</sup>
<b>birth order ~ number of old sisters</b>	0.96	2.2 x 10 <sup>16</sup>
<b>number of old sisters ~ number of old brothers</b>	0.84	2.2 x 10 <sup>16</sup>

## REFERENCES

1. Martin JT, Nguyen DH. Anthropometric analysis of homosexuals and heterosexuals: implications for early hormone exposure. *Horm Behav.* 2004;45: 31–39.
2. Valentova JV, Kleisner K, Havlíček J, Neustupa J. Shape differences between the faces of homosexual and heterosexual men. *Arch Sex Behav.* 2014;43: 353–361.
3. Suire A, Tognetti A, Durand V, Raymond M, Barkat-Defradas M. Speech acoustic features: A comparison of gay men, heterosexual men, and heterosexual women. *Arch Sex Behav.* 2020;49: 2575–2583.
4. McCormick CM, Witelson SF. A cognitive profile of homosexual men compared to heterosexual men and women. *Psychoneuroendocrinology.* 1991;16: 459–473.
5. Xu Y, Norton S, Rahman Q. Sexual orientation and neurocognitive ability: A meta-analysis in men and women. *Neurosci Biobehav Rev.* 2017;83: 691–696.
6. Wang Y, Kosinski M. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *J Pers Soc Psychol.* 2018;114: 246–257.
7. Banks A, Gartrell NK. Hormones and sexual orientation: a questionable link. *J Homosex.* 1995;28: 247–268.
8. Klinga K, Bek E, Runnebaum B. Maternal peripheral testosterone levels during the first half of pregnancy. *Am J Obstet Gynecol.* 1978;131: 60–62.
9. Tasos E. To what extent are prenatal androgens involved in the development of male homosexuality in humans? *J Homosex.* 2022;69: 1928–1963.
10. Swift-Gallant A. Individual differences in the biological basis of androphilia in mice and men. *Horm Behav.* 2019;111: 23–30.
11. Lippa RA. Are 2D:4D finger-length ratios related to sexual orientation? Yes for men, no for women. *J Pers Soc Psychol.* 2003;85: 179–188.
12. Manning JT, Churchill AJG, Peters M. The effects of sex, ethnicity, and sexual orientation on self-measured digit ratio (2D:4D). *Arch Sex Behav.* 2007;36: 223–233.
13. McFadden D, Shubel E. Relative lengths of fingers and toes in human males and females. *Horm Behav.* 2002;42: 492–500.
14. Rahman Q, Wilson GD. Born gay? The psychobiology of human sexual orientation. *Pers Individ Dif.* 2003;34: 1337–1382.
15. Robinson SJ, Manning JT. The ratio of 2nd to 4th digit length and male homosexuality. *Evol Hum Behav.* 2000;21: 333–345.
16. Trivers RL, Willard DE. Natural selection of parental ability to vary the sex ratio of offspring. *Science.* 1973;179: 90–92.
17. Jasienska G, Bribiescas RG, Furberg A-S, Helle S, Núñez-de la Mora A. Human reproduction and health: an evolutionary perspective. *Lancet.* 2017;390: 510–520.
18. Muehlenbein MP, Bribiescas RG. Testosterone-mediated immune functions and male life histories. *Am J Hum Biol.* 2005;17: 527–558.
19. Tamimi RM, Lagiou P, Mucci LA, Hsieh C-C, Adami H-O, Trichopoulos D. Average energy intake among pregnant women carrying a boy compared with a girl. *BMJ.* 2003;326: 1245–1246.
20. Hernández-Julián R, Mansour H, Peters C. The effects of intrauterine malnutrition on birth and fertility outcomes: evidence from the 1974 Bangladesh famine. *Demography.* 2014;51: 1775–1796.
21. Cagnacci A, Renzi A, Arangino S, Alessandrini C, Volpe A. Influences of maternal weight on the secondary sex ratio of human offspring. *Hum Reprod.* 2004;19: 442–444.
22. Sjervold H, James J. Causes of variation in the sex ratio in dairy cattle. *Z Tiers Zuechtungsbiol.* 1979;95: 293–305.
23. Rosenfeld CS, Roberts RM. Maternal diet and other factors affecting offspring sex ratio: a review. *Biol Reprod.* 2004;71: 1063–1070.
24. Murray CM, Eberly LE, Pusey AE. Foraging strategies as a function of season and rank among wild female chimpanzees (*Pan troglodytes*). *Behav Ecol.* 2006;17: 1020–1028.
25. Archie EA, Altmann J, Alberts SC. Costs of reproduction in a long-lived female primate: injury risk and wound healing. *Behav Ecol Sociobiol.* 2014;68: 1183–1193.
26. Kruuk LE, Clutton-Brock TH, Rose KE, Guinness FE. Early determinants of lifetime reproductive success differ between the sexes in red deer. *Proc Biol Sci.* 1999;266: 1655–1661.
27. Simpson MJA, Simpson AE. Birth sex ratios and social rank in rhesus monkey mothers. *Nature.* 1982;300: 440–441.
28. Meikle DB, Tilford BL, Vessey SH. Dominance rank, secondary sex ratio, and reproduction of offspring in polygynous primates. *Am Nat.* 1984;124: 173–188.
29. Paul A, Kuester J. Dominance, kinship and reproductive value in female Barbary macaques (*Macaca sylvanus*) at Affenberg Salem. *Behav Ecol Sociobiol.* 1987;21: 323–331.
30. Vessey SH. Dominance among rhesus monkeys. *Polit Psychol.* 1984;5: 623.
31. Setchell JM, Jean Wickings E. Dominance, status signals and coloration in male mandrills (*Mandrillus sphinx*). *Ethology.* 2005;111: 25–50.
32. Rapaport T, Villaseñor FA, Altman RM, Nepomnaschy PA. Sex ratio and maternal age in a natural fertility, subsistence population: Daughters, sons, daughters. *Am J Phys Anthropol.* 2019;169: 368–376.
33. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature.* 2003;423: 825–837.
34. Jamain S, Quach H, Betancur C, Råstam M, Colineaux C, Gillberg IC, et al. Mutations of the X-linked genes encoding neurologins NLGN3 and NLGN4 are associated with autism. *Nat Genet.* 2003;34: 27–29.
35. Blanchard R, Bogaert AF. Homosexuality in men and number of older brothers. *Am J Psychiatry.* 1996;153: 27–31.
36. Bogaert AF. Biological versus nonbiological older brothers and men's sexual orientation. *Proc Natl Acad Sci U S A.* 2006;103: 10771–10774.
37. Nila S, Crochet P-A, Barthes J, Rianti P, Juliandi B, Suryobroto B, et al. Male homosexual preference: Femininity and the older brother effect in Indonesia. *Evol Psychol.* 2019;17: 1474704919880701.
38. Tio S, Restrepo-Ortiz CX, Roura-Torres B, Sauvadet L, Harté M, Charpentier MJE, et al. The Mandrillus Face Database: A portrait image database for individual and sex

- recognition, and age prediction in a non-human primate. *Data Brief*. 2023;47: 108939.
39. Peignot P, Charpentier MJE, Bout N, Bourry O, Massima U, Dosimont O, et al. Learning from the first release project of captive-bred mandrills *Mandrillus sphinx* in Gabon. *Oryx*. 2008;42. doi:10.1017/S0030605308000136
40. Charpentier M, Harté M, Poirotte C, De Bellefon JM, Laubi B, Kappeler PM, Renoult JP. Same father, same face: deep learning reveals selection for signaling kinship in a wild primate *Science Advances* 6:eaba3274. *Science Advances*. 2020;6.
41. Charpentier MJE, Poirotte C, Roura-Torres B, Amblard-Rambert P, Willaume E, Kappeler PM, et al. Mandrill mothers associate with infants who look like their own offspring using phenotype matching. *Elife*. 2022;11: e79417.
42. Altmann J. Observational study of behavior: Sampling methods. *Behaviour*. 1974;49: 227–266.
43. Galbany J, Romero A, Mayo-Alesón M, Itsoma F, Gamarra B, Pérez-Pérez A, et al. Age-related tooth wear differs between forest and savanna primates. *PLoS One*. 2014;9: e94938.
44. Poirotte C, Massol F, Herbert A, Willaume E, Bomo PM, Kappeler PM, et al. Mandrills use olfaction to socially avoid parasitized conspecifics. *Sci Adv*. 2017;3: e1601721.
45. Tiew S, Dezeure J, Cryer A, Lepou P, Charpentier MJE, Renoult JP. Social and sexual consequences of facial femininity in a non-human primate. *iScience*. 2023;26: 107901.
46. Bogaert AF, Skorska MN, Wang C, Gabriele J, MacNeil AJ, Hoffarth MR, et al. Male homosexuality and maternal immune responsivity to the Y-linked protein NLGN4Y. *Proc Natl Acad Sci U S A*. 2018;115: 302–306.
47. Gibson MA, Mace R. Strong mothers bear more sons in rural Ethiopia. *Proc Biol Sci*. 2003;270 Suppl 1: S108-9.
48. Marsál K, Persson PH, Larsen T, Lilja H, Selbing A, Sultan B. Intrauterine growth curves based on ultrasonically estimated foetal weights. *Acta Paediatr*. 1996;85: 843–848.
49. Setchell JM, Lee PC, Wickings EJ, Dixson AF. Growth and ontogeny of sexual size dimorphism in the mandrill (*Mandrillus sphinx*). *Am J Phys Anthropol*. 2001;115: 349–360.
50. Setchell JM. Sexual Selection and the differences between the sexes in Mandrills (*Mandrillus sphinx*). *Am J Phys Anthropol*. 2016;159: S105-29.
51. Setchell JM, Lee PC, Wickings EJ. Reproductive Parameters and Maternal Investment in Mandrills (*Mandrillus sphinx*). *International Journal of Primatology*. 2002;23: 51–68.
52. Eaton GC, Johnson DF, Glick BB, Worlein JM. Development in Japanese macaques (*Macaca fuscata*) social development: sexually dimorphic behavior in the first year of life. *Primates*. 1985;26: 238–248.
53. Eaton GG, Johnson DF, Glick BB, Worlein JM. Japanese macaques (*Macaca fuscata*) social development: Sex differences in Juvenile behavior. *Primates*. 1986;27: 141–150.
54. Roney JR, Maestripieri D. *Social Development and Affiliation. Primate Psychology*. Harvard University Press; 2005. pp. 171–204.
55. Kulik L, Amici F, Langos D, Widdig A. Sex differences in the development of social relationships in rhesus macaques (*Macaca mulatta*). *Int J Primatol*. 2015;36: 353–376.
56. Brown GR, Dixson AF. The development of behavioural sex differences in infant rhesus macaques (*Macaca mulatta*). *Primates*. 2000;41: 63–77.
57. Maestripieri D, Ross SR. Sex differences in play among western lowland gorilla (*Gorilla gorilla gorilla*) infants: implications for adult behavior and social structure. *Am J Phys Anthropol*. 2004;123: 52–61.
58. Lonsdorf EV. Sex differences in nonhuman primate behavioral development. *J Neurosci Res*. 2017;95: 213–221.
59. Charpentier M, Peignot P, Hossaert-Mckey M, Wickings EJ. Changes in social interactions during adolescence in male mandrills (*Mandrillus sphinx*). *Am J Primatol*. 2004;63: 63–73.
60. Phoenix CH, Goy RW, Gerall AA, Young WC. Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. *Endocrinology*. 1959;65: 369–382.
61. Dörner G, Götz F, Döcke W-D. Prevention of demasculinization and feminization of the brain in prenatally stressed male rats by perinatal androgen treatment. *Exp Clin Endocrinol Diabetes*. 2009;81: 88–90.
62. Thornton J, Zehr JL, Loose MD. Effects of prenatal androgens on rhesus monkeys: A model system to explore the organizational hypothesis in primates. *Horm Behav*. 2009;55: 633–64

---

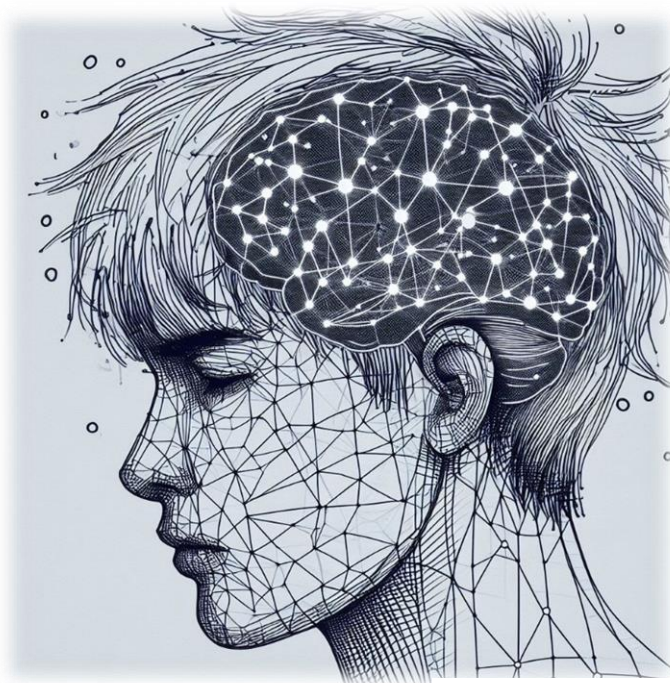
# CHAPITRE 4

---

---

## Comparaison de la sparsité et la typicalité dans les CNNs pour la prédiction de l'attractivité faciale

---



© Produit par DALL·E

### Manuscript 4



---

*Comparing activation sparsity and typicality  
in deep convolutional neural networks to  
predict facial attractiveness*

---

Sonia Tieo, Melvin Bardin, Roland Bertin-Johannet, Nicolas Dibot, Tamra Mendelson,  
William Puech\* Julien P. Renoult\*

\*Equal contribution

**(Article in preparation)**

## **ABSTRACT**

The principle of processing fluency, highlighting the ease with which sensory stimuli are perceived and processed, has garnered attention for its substantial explanatory power in the realm of aesthetic appreciation and beauty. Within this framework, two metrics of fluency emerge as pivotal: the sparsity of neuronal activation and statistical typicality. Utilizing Convolutional Neural Networks (CNNs) as a model for human perception, this study explores these metrics in the context of facial attractiveness, seeking to unravel the associated neural mechanisms. Our findings illustrate that the sparsity of neuronal activations, indicating efficient neural processing, serves as a more robust predictor of facial beauty than statistical typicality, which assesses alignment with an average representation. Despite efforts to enhance predictive accuracy by incorporating specific reference distributions based on attributes like ethnicity or gender, improvements in prediction capabilities were not realized. The outcomes highlight the significance of neural processing efficiency in aesthetic judgments of beauty, suggesting that attractiveness is more profoundly correlated with neural processing pathways than simple facial typicality. This investigation not only amplifies our understanding of aesthetic preferences but also opens avenues for future technological applications and theoretical advancements in varied fields.

## INTRODUCTION

Beauty holds significant influence across multiple aspects of human life. It shapes our perceptions, judgments (known as the "halo effect" [1], preferences [2], and thereby guides our decision-making across diverse arenas, such as personal relationships [3] and consumer choices [4]. In the realm of cognitive science, processing fluency—defined as the ease of interpreting sensory information—has gained significant attention as a plausible determinant of people's evaluation of beauty [5]. While fluency is currently subjectively measured through psychological experiments, few studies have attempted to model it, and none have compared the capabilities of existing models to effectively predict beauty.

As a concept in aesthetic psychology, processing fluency provides a powerful explanation for a wide range of aesthetic inclinations, for both simple and complex stimuli. A preference for basic features such as symmetrical visual patterns and high contrasts is explicable by their effortless processing [5–7]. More complex stimuli, like fractal patterns, would be appreciated for their smooth processing as well, since the self-similarity of fractals at different scales makes these patterns highly predictive [8,9]. Fluency would also explain the attractiveness of prototypical representations [10]. Prototypes are typified by familiar and easily discernible features, and thus prototype-like stimuli are processed with ease, enabling rapid and accurate categorization while enhancing memory retention. Regardless of their specific attributes, prototypes consistently earn preference across an array of stimuli, encompassing biological, inanimate, and abstract forms [10–12].

Fluency, therefore, has a high explanatory power, and using this concept to predict beauty would have numerous technological applications, but also fundamental implications, for instance, in allowing to study beauty in non-human animals [13]. However, the predictive capacity of fluency is limited by the scarcity and current limits of studies aiming to model this concept. Early research in modeling processing fluency primarily focused on objective measures of feature repetitions in stimuli, such as symmetry, contrast, and self-similarity, for visual stimuli [14]. While these metrics could provide valuable insights, they predominantly focus on the stimulus itself, assuming that the studied features ease information processing. Yet fluency is fundamentally rooted in the interaction between the stimulus and the perception of the

beholder, and thus more accurate metrics should target features as they are processed by the perceptual system.

A first step in modeling fluency is therefore to model the processing of features using a model of perception. The development of such models is uneven across the different sensory modalities, but is by far the most advanced for visual perception, on which we will focus in this study. [13,15] As demonstrated in numerous studies [16,17], Convolutional Neural Networks CNNs have recently emerged as powerful models of visual information processing, from low-level feature extraction to high-level semantic interpretation. Just like how our visual system processes information, CNNs start by extracting rudimentary features such as edges and contours in their initial layers. As the information flows through the network, the recognition of increasingly intricate features takes place, which parallels our visual system's ability to discern complex objects or scenes by combining simpler constituent elements. By studying neuronal activations within CNNs, which represent the response of different neurons of each layer to specific image inputs, we can thus gain valuable insights into how biological vision processes visual information.

The second step in modeling fluency is to choose a metric that characterizes the ease of information processing. Inspired by the information theory applied to biological systems [18], some authors have proposed to model fluency using the sparsity of the neuronal activation [14,19]. Sparsity measures the concentration of neuronal activity in a specific subset of features or patterns. A sparse stimulus thus activates only a few neurons simultaneously, leading to a low-cost, efficient processing of information [20]. Using the sparsity of neuronal activations to estimate fluency fits the prediction of Winkielman et al. [21], that “fluent patterns should be represented by more extreme values of activation”. Previous studies have provided empirical evidence supporting the link between sparsity and beauty. For instance, using a model of information processing in the primary visual cortex, [19] have shown a positive correlation between the sparsity of neuronal activations and the perceived attractiveness of female faces. Furthermore, sparsity has been identified as a robust predictor of face attractiveness compared to other factors such as body mass index, sexual dimorphism, averageness, and asymmetry [22]. Very recently, one study evaluated the ability of the sparsity of neuronal activations within a CNN to predict the

beauty of faces and artistic paintings [23]. The authors showed that sparsity alone could explain up to 28% of the variance in beauty scores.

Another metric of fluency proposed in the literature is the statistical typicality. Typicality describes the extent to which a stimulus aligns with an average representation. It is based on the underlying assumption that individuals form mental representations of averageness for various categories based on their past experiences and exposure to stimuli. Typicality is thus closely related to familiarity, which has been a well-studied factor influencing fluency in psychological studies [7]. Ryali and colleagues [24] demonstrated that the attractiveness of a face is partly explained by its statistical typicality, defined as the likelihood of the face image relative to an internal representation of the face distribution. The authors used the Active Appearance Model (AAM) as a model of face perception, which is built from features describing the shape and texture of faces. They then showed that the attractiveness of a given face can be predicted from its likelihood estimated from the distribution of faces of the same gender.

The contribution of this study is threefold. First, we propose a new model of fluency that applies the statistical typicality metric to convolutional neural networks (CNNs). More precisely, for each layer of the CNN the method estimates the likelihood of a stimulus encoding given a reference distribution of encoded features. The method thus extends previous applications of the statistical typicality ([24]; see also [25]) to a model of perception that describes visual information processing as it operates throughout the retina and the ventral stream of the visual system. Second, we aim to compare the ability of sparsity and statistical typicality computed in CNN layers to predict the attractiveness of human faces. Attractiveness is arguably the strongest determinant of facial beauty, to a point that both terms are generally used interchangeably in the psychological literature [2]. For this purpose, we used the publicly available Chicago Face Dataset [26], a comprehensive collection of face images with empirical scores of attractiveness. We encoded each face with a CNN, calculated its sparsity and likelihood at each layer, and trained two regression models (one with sparsity, the other with likelihood) to predict attractiveness. Third, we examine the extent to which the ability of statistical typicality to explain facial attractiveness is influenced by the choice of the reference distribution. Specifically, we investigate

whether specializing the reference distribution to include only faces of a single gender and/or a single ethnic group improves the ability of likelihood to predict attractiveness.

## **MATERIAL AND METHODS**

### **1. Materials**

We modeled the fluency of processing portrait images of Chicago Face Dataset (CFD hereafter; [26]). This dataset contains standardized photographic portraits of individuals aged 17 to 65, spanning a variety of ethnic backgrounds, including East Asian, Black, Hispanic, and White, with balanced representation across genders. We use a subset of 597 portraits of CFD, each depicting a neutral expression, frontal view, and standardized attire. This dataset also includes ratings of attractiveness, evaluated by independent judges. Each portrait is associated with one mean score of attractiveness.

### **2. Modeling face processing using CNNs**

To model the visual processing of faces, we compared two Convolutional Neural Networks (CNNs), VGG16 [27] and VGGFace [28], that have been pre-trained on the ImageNet and VGGFace datasets, respectively. VGG-16 includes 13 convolutions and two fully connected layers. ImageNet is a large dataset of 14 million images depicting about 20,000 categories including people, plants, animals and human-made objects. VGG16 trained on such a large and varied dataset allows modeling a visual cortex that is not specialized to one specific task [29,30]. In contrast, VGGFace is a variant of VGG16 specifically tuned for face recognition. The VGGFace dataset includes images capturing variations in facial expressions, angles, and lighting conditions. The fine-tuning of VGGFace to this dataset allows the entire network to adapt to face-specific features across all layers.

### **3. Metrics of fluency**

In a CNN, the output of a convolutional layer is a matrix of size  $H \times W \times C$  where each entry represents the activation of a “neuron”. The dimension  $C$  describes the number of channels, each embodying a unique feature, such as a distinct contrast or edge orientation.  $H \times W$  is termed a feature map, describing where in the image a feature is present;  $H$  and  $W$  are the height and the width of the feature map, respectively.

## Sparsity

The first metric of fluency is activation sparsity, which measures the concentration of neuronal activity in specific features. We quantified sparsity using the Gini index [31]. To do so, we flattened the activation matrices into one-dimensional vectors and sorted the vectors in ascending order. The Gini index was then computed using the formula:

$$Gini = (\sum(2 * i - n - 1) * x_i) / (n * \sum x_i)$$

where,  $n$  is the total number of activations in the layer (vector length) and  $x_i$  the activation value at index  $i$ . Higher Gini indices indicate a higher degree of sparsity, reflecting a more selective activation.

## Statistical typicality

The second metric of fluency investigated is statistical typicality, calculated in three steps. The first step reduces the dimensionality of the  $HxWxC$  matrices. As the images traverse through a pre-trained network, a multitude of feature maps are generated from the different layers. Some of these feature maps have an extensive number of activations (for instance, the first convolution layer of VGG16 outputs 3,211,264 activations, rendering the estimation of statistical typicality computationally intractable. We compared three strategies of dimensionality reduction. In the first strategy, we performed one Principal Component Analysis (PCA) per layer, thus considering all activations (ie, after flattening the  $HxWxC$  matrix). We kept the number of principal components allowing to account for 80% of the variance in activations [32]. This number varied across layers and datasets (between 28 and 622). In the second strategy, a PCA was applied individually to each feature map, preserving the components allowing to account for 80% of intra-map variance. The components from all feature maps within a specific layer were then concatenated and reduced further using a second PCA, again keeping the number of principal components allowing to account for 80% of the variance. In the third strategy, one first calculated the mean activation of each feature map and then concatenated all the means into a single vector for each layer. Despite their potential interest, in particular the low computation cost of the third strategy, neither the second nor the third strategy improved our results compared to the first one. We thus here only present results obtained with the first strategy, based on a layer-wise PCA.

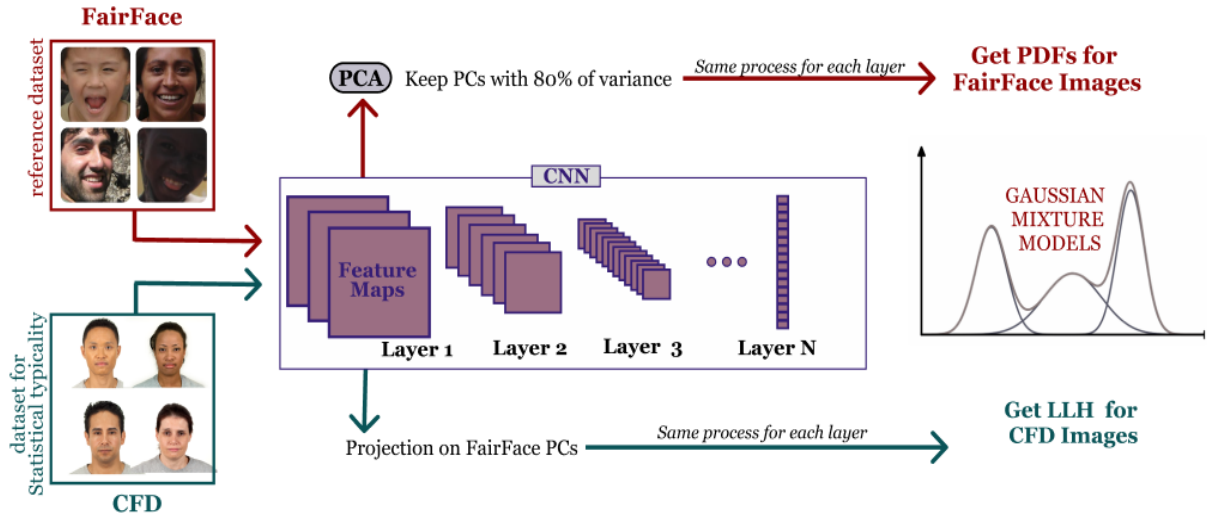
The second step for measuring statistical typicality was to establish a reference distribution of encoded features. Following the approach proposed by others [24,25], reference distributions were represented by Probability Density Functions (PDFs; one PDF per layer) from the reduced (by the layer-wise PCA) activations of all encoded portraits of a reference dataset. We used FairFace as a reference dataset [33], a collection of face portraits with a perfectly balanced representation of ethnic, gender, and age classes, enabling us to build an unbiased model of statistical typicality. PDFs were estimated from a balanced subset of 1,000 portrait images from FairFace. This number was chosen to allow a comparison of results when changing the reference dataset (see the section on the *Influence of the reference dataset*).

We estimated each PDF with Gaussian Mixture Models (GMMs), using the principal components as the variables in the Gaussian models. With their capacity to combine simple Gaussian density functions or "Gaussian components", GMMs adeptly capture the complexity of the underlying distribution of facial features. We determined the optimal number of Gaussian components with the Bayesian Information Criterion (BIC), setting an upper limit to 20 components.

In the third step, we used PDFs estimated from the reference datasets to estimate the log-likelihood (LHH) of each image of the Chicago Face Database (CFD), for every layer of the CNN. This was done after projecting the activation values of each image of CFD onto the principal components calculated using the FairFace dataset.

One limitation of GMMs is their sensitivity to singularities. Concretely, this means that one gaussian component can be fitted onto a single outlier, thus inflating the LLH of images located close to this outlier in the feature space. To mitigate this potential issue, the LLH of an image was calculated as the median value obtained after 100 replications of the analysis (including the second and third steps described previously, but not the first, dimensionality reduction step, because the layer-wise PCA is deterministic and thus performed only once). Using Bayesian Mixture Gaussian Models as an alternative to address the problem of singularities yielded similar results but at much higher computational cost (results not shown).





**Figure 1. Pipeline for estimating the statistical typicality of faces.** Images of the FairFace dataset are first encoded with a CNN (VGG16 or VGGFace). For each layer of the CNN, the dimensionality of the activation space describing the encodings is then reduced using a principal component analysis (PCA), keeping only the principal components (PCs) explaining 80% of the variance. One probability density function is estimated per layer using Gaussian Mixture Models (GMM) fitted onto the retained PCs. In a second step, images of the CFD dataset are encoded as in the first step, and their log-likelihood (LLH) calculated for each layer from the density probability functions (PDFs) estimated on FairFace. This entire process is repeated 100 times, and the statistical typicality corresponds to the median of all repetitions.

## Predicting attractiveness from sparsity and statistical typicality

Using the face portraits of CFD, we assessed the ability of layer-wise sparsity of activations and statistical typicality (LLHs) calculated from all convolutional and fully connected layers ('AllLayers' models, see below) to predict facial attractiveness. We conducted a regression analysis with attractiveness as the response variable, and the LLHs or sparsity values of every layer as independent variables. Our first model involving sparsity can be expressed as:

$$Attractiveness \sim \sum(b_l * Sparsity\_layer\_l),$$

with  $l$  varying from 1 to  $N$  (the number of layers for VGG or VGG16),  $Sparsity\_layer\_l$  indicating the sparsity of layer  $l$  and  $b_l$  the regression coefficient.

We similarly defined our model for statistical typicality as:

$$\text{Attractiveness} \sim \sum(a_l * LLH\_layer\_l),$$

with  $l$  varying from 1 to  $N$ ,  $LLH\_layer\_l$  representing each layer's  $LLH$  and  $a_l$  the regression coefficient.

We employed ridge regression models, rather than classical linear regression models, to address the inherent collinearity among layers, due to the fact that layer outputs are inputs of the following layers. Ridge regression effectively handles multicollinearity and overfitting by applying a regularization that balances model complexity and generalization [34]. We performed ridge regressions using the *glmnet* package in R, using a 10-fold cross-validation repeated 10 times. For ridge regression ( $\alpha=1$ ), we explored a range of penalty values ( $\lambda$ ), spanning from  $10^2$  to  $10^{-4}$ . Both the predictors and the response were centered and scaled prior to the analysis. The coefficient of determination,  $R^2$  is the explained variance of attractiveness.

In addition to our primary focus on “AllLayers”, we delved into the distinct subsets of the VGGFace neural network layers. These subsets included “LastConvLayer”, “PoolingLayers”, “ShallowLayers”, “MiddleLayers”, and “DeepLayers”. Each subset represents a specific arrangement of layers, with breakdowns detailed in the supplementary information Figure S1. By analyzing these subsets, we aimed to unravel how different layers of the neural network contribute to facial attractiveness predictions.

#### **4. Influence of the reference dataset**

Previous research suggested that the categorization of faces according to gender and ethnicity can impact perceived attractiveness [35,36]. We then investigated if a more specialized reference Probability Density Function (PDF) — built using a reference dataset tailored more specifically to a particular gender, ethnic group, or a

combination of the two — could enhance the ability of statistical typicality to predict attractiveness.

We used the gender and ethnic categories of FairFace to build 15 reference datasets, each containing exactly 1,000 images (see details in Table 1). These 15 datasets can be categorized into four types of reference datasets differing in the level of specialization: “all”, “ethnicity”, “gender” and “ethnicity x gender”. We consider that “ethnicity x gender” represents a more specialized reference dataset than “ethnicity” and “gender”, which are themselves more specialized than “all”. Ethnic and gender categories were kept balanced within the “all”, “ethnicity” and “gender” reference datasets. As previously, we estimated PDFs (one per layer) for each reference dataset. Then, for each image of one combination of gender and ethnic category of CFD, we calculated its LLHs with the different reference datasets. For Asian male individuals, for example, we calculated LLHs of all images of Asian males depicted in the CFD dataset considering PDFs estimated from a reference dataset including 1,000 portraits of FairFace depicting either i) individuals of both genders and all (Asian, Black, White, Latino) ethnic groups (“all” reference dataset), ii) all Asian males and females (“ethnicity”), iii) males of all ethnic groups (“gender”), and iv) Asian males only (“ethnicity x gender”) (Figure 2).

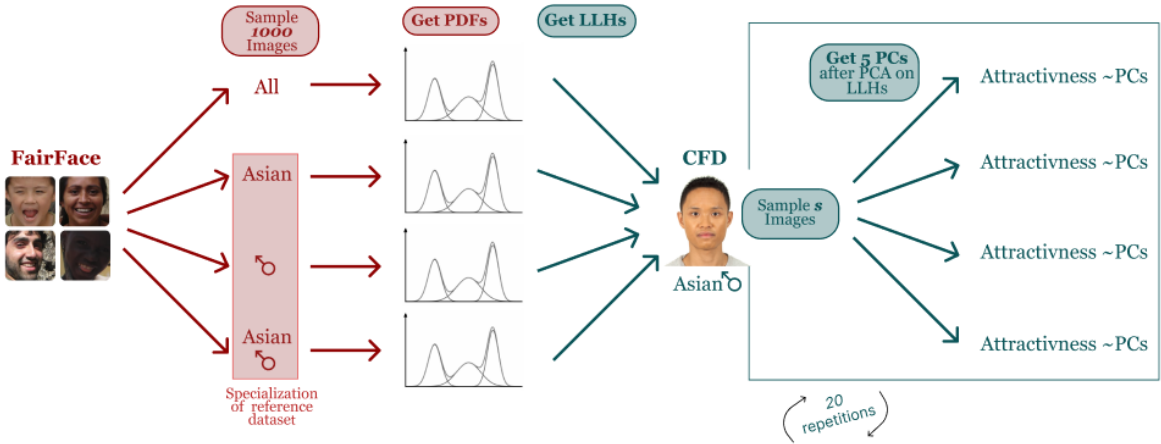
To mitigate the problem of having a different number of images between gender and ethnic categories in CFD (adjusted- $R^2$  cannot be used in ridge regressions and thus  $R^2$  is influenced by the number of observations), we randomly sampled 52 images from each category, equivalent to the minimum number of images present in any category. This sampling was repeated 20 times.

For each reference dataset, we thus obtained a set of LLHs (one per layer), which were as previously used in a ridge regression model to explain facial attractiveness. However, to address the challenge of having a large number of regressors relative to the number of observations (only 52 observations), rather than using raw LLHs values in the regression models, we reduced the number using a PCA. The first 5 principal components (PCs) cumulatively explained between 92% to 98% of variance. We calculated  $R^2$  for each ridge regressions, using cross-validation as above. To test the hypothesis that a more specialized reference dataset leads to higher  $R^2$  values, we then used a Generalized Linear Mixed Model (GLMM) with the  $R^2$  (derived from the mean of the 20 repetitions) as the response variable and the type of reference specialization

(categorical, four levels, see Table 1) as explanatory variable. The categories of “ethnicity x gender” within the CFD dataset (CFD\_Categories\_“ethnicity x gender”) were included as a random effect (categorical variable, eight levels, see Table 1) in the model. The model was thus expressed as:

$$R^2 \sim 5PCs(LLH) + ReferenceSpecializationType + 1|CFD\_Categories\_“ethnicity\ x\ gender”$$

The model was performed using lme4, and results were presented using sjPlot in R.



**Figure 2. Log-Likelihood (LLH) calculation using different reference datasets - An example using the Asian males category.** LLH values for Asian males in the Chicago Dataset were calculated using four types of reference dataset : 'all', 'Asian' (ethnicity-specific subset from FairFace), 'Males' (gender-specific subset from FairFace), and 'Asian Males' (ethnicity and gender-specific subset from FairFace). To address the variation in the number of images for each gender and ethnic category within the CFD dataset (see Table 1), we randomly selected a set of  $s=52$  images from each category, which corresponds to the smallest category size. The derived LLH values underwent dimensionality reduction using PCA, resulting in the first 5 principal components (PCs). These 5 PCs were then incorporated into a ridge regression model to predict facial attractiveness. This sampling process was carried out 20 times to ensure consistency. The process delineated in the figure is representative of the method applied to all 15 categories (Table 1).

**Table 1. Description of Reference datasets from FairFace dataset** Each reference dataset includes 1,000 images randomly sampled from the entire FairFace dataset (All), or from a subset of a single gender, a single ethnicity or a single gender of a single ethnicity. The column “Genders and ethnic groups” indicates the composition of the subset, the column “Reference specialization type” precise the type of specialization (all, ethnicity, gender or ethnicity x gender) and the last column “Reference dataset identifier” is the identifier of each reference dataset.

<b>Genders and ethnic groups</b>	<b>Reference specialization type</b>	<b>Reference dataset identifier</b>
all	all	FairFace_All
Asian, both genders	ethnicity	FairFace_A
Black, both genders	ethnicity	FairFace_B
Latino, both genders	ethnicity	FairFace_L
White, both genders	ethnicity	FairFace_W
Females, ethnicity	gender	FairFace_F
Males, all ethnicity	gender	FairFace_M
Asian females	ethnicity x gender	FairFace_AF
Black females	ethnicity x gender	FairFace_BF
Latino females	ethnicity x gender	FairFace_LF
White females	ethnicity x gender	FairFace_WF
Asian males	ethnicity x gender	FairFace_AM
Black males	ethnicity x gender	FairFace_BM
Latino males	ethnicity x gender	FairFace_LM
White males	ethnicity x gender	FairFace_WM

**Table 2. Description of the test datasets from Chicago Face Dataset dataset for the statistical typicality analysis** The subsets of images of CFD are divided by a subset of a single gender, a single ethnicity or a single gender of a single ethnicity (column “Genders and ethnic groups”) with their corresponding identifier given on the column “Test dataset identifier” and the corresponding number of pictures “Number of images”. Each subset of CFD is analyzed with each reference dataset indicated in the column “Reference datasets compared”.

<b>Genders and ethnic groups</b>	<b>Test dataset identifier</b>	<b>Number of images</b>	<b>Reference datasets compared</b>
Asian females	CFD_AF	57	FairFace_All/FairFace_A/FairFace-F/FairFace_AF
Black females	CFD_BF	104	FairFace_All/FairFace_B/FairFace-F/FairFace_BF
Latino females	CFD_LF	56	FairFace_All/FairFace_L/FairFace-F/FairFace_LF
White females	CFD_WF	90	FairFace_All/FairFace_W/FairFace-F/FairFace_WF
Asian males	CFD_AM	52	FairFace_All/FairFace_A/FairFace-M/FairFace_AM
Black males	CFD_BM	93	FairFace_All/FairFace_B/FairFace-M/FairFace_BM
Latino males	CFD_LM	52	FairFace_All/FairFace_L/FairFace-M/FairFace_LM
White males	CFD_WM	93	FairFace_All/FairFace_W/FairFace-M/FairFace_WM

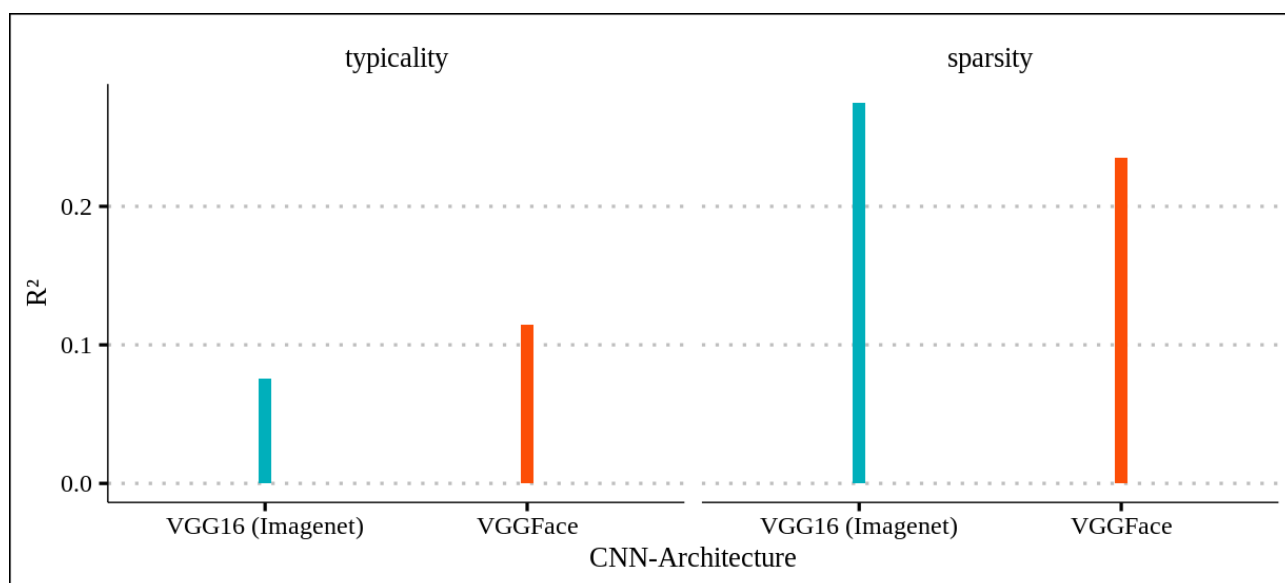
## RESULTS

### Comparing statistical typicality and sparsity

We first evaluated the ability of statistical typicality, measured from the LLH of face encodings at each layer of the CNN to explain variation in attractiveness. Using portrait images of FairFace as a reference dataset to calculate the LLHs of image of the CFD dataset, we found that statistical typicality explains 8% of variance in facial attractiveness ( $R^2$ ) with VGG16, and 11% of variance with VGGFace (Figure 3A).

In comparison, sparsity accounted for 27% and 23% of the variance in attractiveness with VGG16 and VGGFace, respectively (Figure 3B). These results suggest that variation in the sparsity of neuronal activations triggered by facial features explain variation in attractiveness more than the statistical typicality of these facial features does.

We then investigated whether we could increase the explanatory capacities of statistical typicality by including LLHs from a subset of VGGFace layers, the CNN that yielded the highest  $R^2$  for this fluency metric, rather than considering all layers. We considered four subsets of layers: regression models with “LastConvLayer” include the last convolutional layer of each block, “PoolingLayers” models includes all pooling layers, “ShallowLayers”, “MiddleLayers” and “DeepLayers” models include early, mid-tier and the deeper layers of the network, respectively. The specific layers included in each of these subsets are detailed in Figure S1. With none of these subsets the  $R^2$  score surpassed the one reached by sparsity. More precisely, we found that the fraction of explained variance was lower compared to when considering all layers, except with the “DeepLayers” subset that yielded similar performances (11% of explained variance of attractiveness; Figure S1).



**Figure 3.** Comparison of explained variance of attractiveness ( $R^2$ ) using sparsity and statistical typicality metrics derived from VGG16(imagenet) and VGGFace architectures.

In the previous analysis of statistical typicality, we found a slightly higher explanatory power when using VGGFace compared to VGG16 with the weights of Imagenet. This could be due to the specialization of VGGFace to process faces, leading to PDFs that are more tightly tuned to facial features and thus to more meaningful values of LLH. We thus wondered whether the statistical typicality metric would be more powerful when specializing the PDFs even further, such that LLHs are calculated

in reference to one gender or one ethnic group only, or even one gender of one ethnic group, rather than all faces considered together.

To investigate the role of specializing the reference datasets and PDFs further, we performed a regression model with the R<sup>2</sup> of the 15 models with different levels of specialization (Figure S2) as a response variable. When analyzing the different levels of the categorical variable “Reference specialization type”, we found that “ethnicity x gender” (estimate = -0.03, 95% CI [-0.06, 0], p = 0.027; Table 2) significantly but negatively influenced the explanatory power of statistical typicality. This confirms that specializing the reference dataset and associated PDFs does not increase the ability of statistical typicality to predict facial attractiveness. On the contrary, we obtained the best performance when using the All reference dataset -where images were sampled across all genders and ethnic groups of the FairFace dataset.



**Table 2. Influence of reference dataset specialization on facial attractiveness prediction.** Results from GLMM analysis exploring how different reference specialization types – “ethnicity”, “gender”, and “ethnicity x Gender” – influence the accuracy of predicting facial attractiveness. Fixed effects include the number of photos from CFD categories, while random effects encompass the 8 *Categories\_“ethnicity x gender”* from CFD images (as detailed in Table 1). The “All” reference specialization type, with diverse images across genders and ethnicities, demonstrates significantly better predictive accuracy than 'ethnicity' and 'ethnicity x gender' reference specialization types. The formula used is:  $R^2 \sim 5PCs(LLH) + ReferenceSpecializationType + 1|CFD\_Categories\_“ethnicity\ x\ gender”$ .

<b>Predictors</b>	<b>Estimates</b>	<b>CI</b>	<b>p</b>
(Intercept)	0.32	0.28 – 0.35	<0.001
Reference specialization type [ethnicity] (ref: All)	-0.02	-0.05 – 0.01	0.135
Reference specialization type [Gender] (ref: All)	-0.02	-0.05 – 0.01	0.112
Reference specialization type [ethnicity x gender] (ref:All)	-0.03	-0.06 – 0.00	<b>0.027</b>
<b>Random effects</b>			
N <i>CFD_Categories_“ethnicity x gender”</i>	8		
Observations	32		
Marginal R2 / Conditional R2	0.044 / 0.593		

## DISCUSSION

The objective of this study was to compare two fluency metrics – one of the main determinants of beauty – in predicting facial attractiveness. By comparing statistical models performed with the same dataset and including the same number of explanatory variables, we showed that the sparsity of neuronal activations is a better predictor than the statistical typicality of facial features.

This result sheds light on the neural mechanisms involved in the sensation of fluency studied in psychology. By analyzing the different types of preferences involved in fluency, Renoult and Mendelson [13] proposed that a stimulus is fluent when it is processed either with efficacy by the brain, meaning with high-quality processing and therefore minimal information loss, or with efficiency, meaning optimal resource utilization. The question of whether it is the efficiency or rather the efficacy of processing that explains the attractiveness of such stimuli has remained unaddressed.

For example, it has been shown that prototypes are attractive because they are easy for the mind [37], but ease can equally describe efficiency and efficacy. Prototypes are efficient stimuli because they are most quickly and precisely categorized and stored the longest in memory [37]. However, prototypes are also efficient because they only need to stimulate a few highly selective neurons to be recognized [38]. Similarly, statistical typicality measured through LLH can describe either efficacy or efficiency. Sparsity measured through the Gini index characterizes efficiency more directly, and only efficiency. Our results thus suggest that efficiency, rather than efficacy, describes fluency at the neuronal level.

Not all of the information about attractiveness conveyed by statistical typicality is included in neuronal sparsity. Dibot et al. [23] analyzed the contribution of the sparsity of individual CNN's layers and found that the first layers of VGG16-Imagenet explained most of the variance in facial beauty. This result is consistent with that of Renoult et al. [19], who showed that activation sparsity in a model of the primary visual cortex explains up to 17% of the variance. Indeed in VGG16 the first convolutional layers have been shown to model feature extraction as it operates in the primary visual cortex [17]. In contrast, with statistical typicality we found that the fraction of explained variance was the highest for the deeper layers (11%), and adding information from the shallower layers did not increase the R2 score further. In a CNN, the deeper layers encode complex patterns and their arrangement at the largest spatial level; for faces, this means the shape and coloration of facial elements (i.e. the nose, the mouth, the eyes) but also their relative position in the face [39]. At this configural level of perception, high statistical typicality thus likely increases the ability to recognize a face individually and to memorize it. Thus, while fluency could be driven by the efficiency of information processing in the early stages of the visual system, it may be more strongly influenced by efficacy in the later stages of the processing pathway. It is noteworthy that, if they affect different stages of the information processing pathway, efficacy and efficiency could nevertheless both influence the overall fluency. Indeed, previous studies in psychology have indeed shown that the ease of processing information at different stage of the visual system, as measured by detection time in the early stages and recognition performance in the later stages, triggers micro-experiences of fluency that aggregate into one global sensation of fluency [5,40]. To investigate further whether and how sparsity and typicality differently influence

fluency at different stages of visual perception, future studies should aim to correlate these metrics with empirical measures of fluency in detection and recognition tasks.

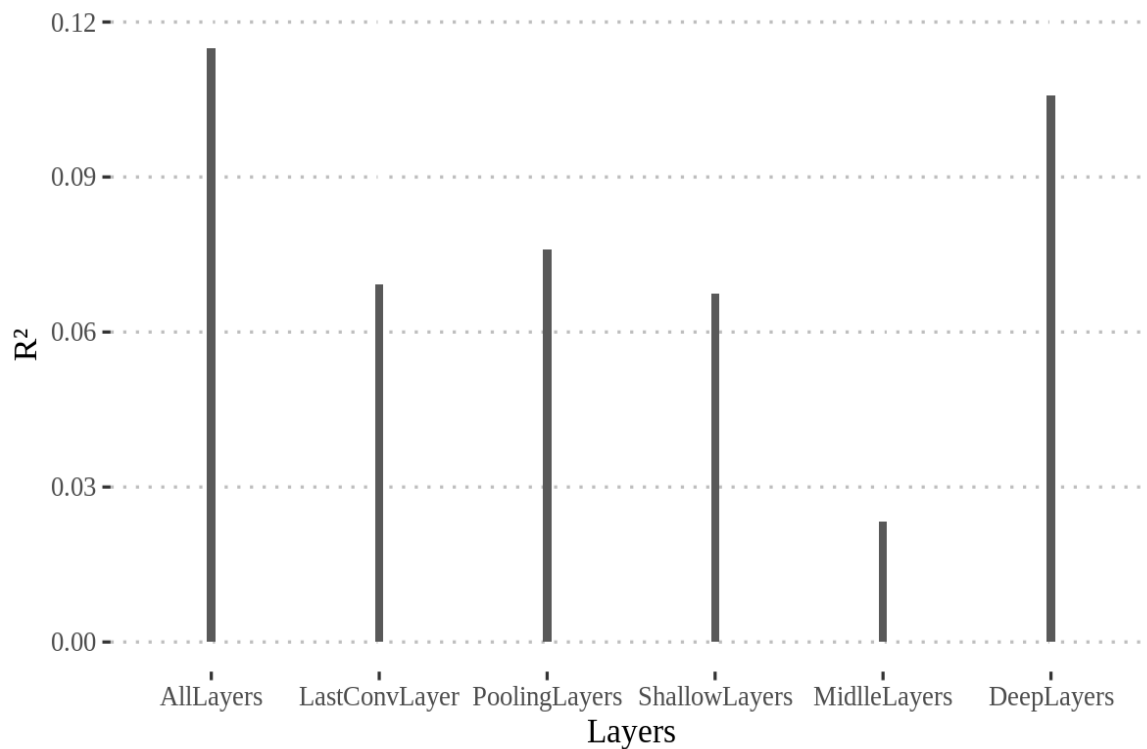
For a more parallel comparison, we can look towards the work of Ryali et al. [24]. They also defined a metric of statistical typicality centered on Log-Likelihood (LLH). However, while we employed CNNs to model perception, they opted for an alternative method (AAM). In their experiment examining facial attractiveness using also the Chicago Face Dataset, they found a strong correlation ( $r=0.386$ ) between attractiveness ratings and the LL of the stimuli. Their model integrates both low and high-level information, analogous to the deeper layers of our network. The alignment in our  $R^2$  values suggests that statistical typicality significantly influences attractiveness at the configural perception level. This concordance with Ryali et al. highlights the importance of configural perception and reaffirms the role of statistical typicality across different methods and datasets.

To give statistical typicality the maximum of chances, we explored the possibility of using a reference distribution refined to a specific type of face. We incorporated these refinements based on previous results showing that the brain could interpret face signals in ethnic or gender-related subgroups [24,41–43]. We then explored different reference distributions, focused on one ethnic group, one gender, or one gender of a given ethnic group. However, with none of the tested reference distributions, we found that statistical typicality surpassed sparsity. At first glance, this might appear contradictory to prior research indicating that facial attractiveness, when multiple groups are considered, correlates with how closely a face matches the prototype of its group. For instance, studies have found that computer-generated Caucasian and African-American faces gain attractiveness as they more closely resemble the average features of their racial group [35]. Our results may not necessarily be contradictory: Our perceptions of beauty are deeply woven with our exposure to various faces and the mental archetypes we form over time. Rhodes et al. (2003) emphasized that our engagements with a spectrum of faces shape our conception of the 'ideal' face [12]. Extending this notion, researchers like Lewis (2010) and Rhodes et al. (2005) posit that we might cultivate a 'composite' facial image that amalgamates average features from diverse groups, suggesting that our evaluations of attractiveness are not strictly conditioned by factors like ethnicity or gender. This could explain why focusing the

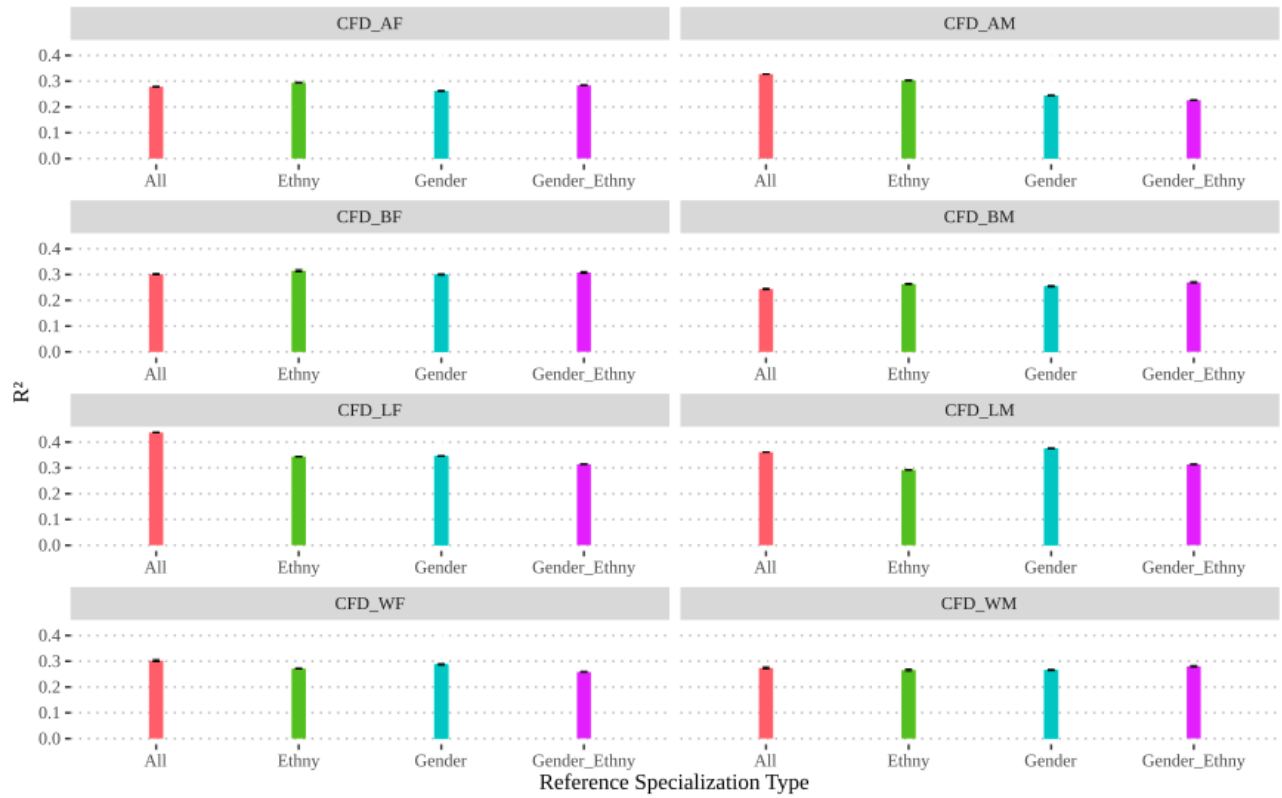
reference distribution on one specific category —be it gender, ethnicity, or their intersection— didn't enhance attractiveness predictions.

In conclusion, modeling fluency in perception involves capturing how a stimulus is processed within an observer's brain. In this research, we utilized the potential of Convolutional Neural Networks (CNNs) to mimic this processing within the visual system. Although past studies have employed CNNs to forecast beauty, they largely investigated the ability of features to directly predict attractiveness. In contrast, fluency characterizes how the visual system processes these features, thereby making it independent of the inherent meaning or nature of these features. Our use of CNNs allowed us to probe the underlying neural mechanisms of fluency. We discovered that the sparsity of neuronal activation, which portrays the efficiency of neural information processing, appears to be a more powerful determinant of beauty than statistical typicality. This finding underscores the importance of taking into consideration the processing cost when studying attractiveness.

## SUPPLEMENTARY DATA



**Figure S1. Comparison of explained variance of attractiveness ( $R^2$ ) using subset of layers of VGGFace. AllLayers:** all available layers of VGGFace. **LastConvLayer:** layers block1\_conv2, block2\_conv2, block3\_conv3, block4\_conv3, and block5\_conv3. **PoolingLayers:** layers block1\_pool, block2\_pool, block3\_pool, block4\_pool, and block5\_pool. **ShallowLayers:** layers input\_1, block1\_conv1, block1\_conv2, block1\_pool, block2\_conv1, block2\_conv2, and block2\_pool. **MiddleLayers:** Layers block3\_conv1, block3\_conv2, block3\_conv3, block3\_pool, block4\_conv1, block4\_conv2, block4\_conv3, and block4\_pool. **DeepLayers:** layers block5\_conv1, block5\_conv2, block5\_conv3, block5\_pool, fc6, fc7, fc8, and flatten.



**Figure S2. Explained variance of attractiveness ( $R^2$ ) for the 8 Categories “ethnicity x gender” from CFD images (CFD\_AF, CFD\_AM, CFD\_BF, CFD\_BM, CFD\_LF, CFD\_LM, CFD\_WF and CFD\_WM) by reference specialization type – ‘All’ (red), ‘ethnicity’(green) , ‘gender’ (blue), and ‘ethnicity x Gender’ (pink).**

## REFERENCES

1. Batres C, Shiramizu VKM. PSA001 Secondary Analysis: Examining the “attractiveness halo effect” across cultures. *PsyArXiv*. 2020. doi:10.31234/osf.io/c7hf3
2. Rhodes G. The evolutionary psychology of facial beauty. *Annu Rev Psychol*. 2006;57: 199–226.
3. Rhodes G, Simmons LW, Peters M. Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evol Hum Behav*. 2005;26: 186–201.
4. Lee AY, Labroo AA. The Effect of Conceptual and Perceptual Fluency on Brand Evaluation. *J Mark Res*. 2004 [cited 12 Sep 2023]. doi:10.1509/jmkr.41.2.151.28665
5. Reber R, Schwarz N, Winkielman P. Processing fluency and aesthetic pleasure: is beauty in the perceiver’s processing experience? *Pers Soc Psychol Rev*. 2004;8: 364–382.
6. Jacobsen T, Schubotz RI, Höfel L, Cramon DY. Brain correlates of aesthetic judgment of beauty. *Neuroimage*. 2006;29. doi:10.1016/j.neuroimage.2005.07.010
7. Reber R, Winkielman P, Schwarz N. Effects of perceptual fluency on affective judgments. *Psychol Sci*. 1998;9: 45–48.
8. Forsythe A, Nadal M, Sheehy N, Cela-Conde CJ, Sawey M. Predicting beauty: fractal dimension and visual complexity in art. *Br J Psychol*. 2011;102: 49–70.
9. Street N, Forsythe AM, Reilly R, Taylor R, Helmy MS. A Complex Story: Universal Preference vs. Individual Differences Shaping Aesthetic Response to Fractal Patterns. *Front Hum Neurosci*. 2016;10. doi:10.3389/fnhum.2016.00213
10. Winkielman P, Halberstadt J, Fazendeiro T, Catty S. Prototypes are attractive because they are easy on the mind. *Psychol Sci*. 2006;17: 799–806.
11. Winkielman P, Schwarz N, Fazendeiro TA, Reber R. The hedonic marking of processing fluency: Implications for evaluative judgment. 2003 [cited 12 Sep 2023]. Available: <https://www.semanticscholar.org/paper/The-hedonic-marking-of-processing-fluency%3A-for-Winkielman-Schwarz/750bf49044a127106a89bad9f90c01741f6adad>
12. Halberstadt J, Rhodes G. It’s not just average faces that are attractive: computer-manipulated averageness makes birds, fish, and automobiles attractive. *Psychon Bull Rev*. 2003;10: 149–156.
13. Renoult JP, Mendelson TC. Processing bias: extending sensory drive to include efficacy and efficiency in information processing. *Proc Biol Sci*. 2019;286: 20190165.
14. Redies C. A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spat Vis*. 2007;21: 97–117.
15. Mayer S, Landwehr JR. Quantifying visual aesthetics based on processing fluency theory: Four algorithmic measures for antecedents of aesthetic preferences. *Psychology of Aesthetics, Creativity, and the Arts*. 2018 [cited 12 Sep 2023]. doi:10.1037/ACA0000187
16. Kriegeskorte N. Deep neural networks: A new framework for modeling biological vision and brain information processing. *Annu Rev Vis Sci*. 2015;1: 417–446.
17. Lindsay GW. Convolutional neural networks as a model of the visual system: Past, present, and future. *J Cogn Neurosci*. 2021;33: 2017–2031.
18. Barlow HB. Possible principles underlying the transformations of sensory messages. *Sensory Communication*. The MIT Press; 2012. pp. 216–234.
19. Renoult JP, Bovet J, Raymond M. Beauty is in the efficient coding of the beholder. *R Soc Open Sci*. 2016;3: 160027.
20. Olshausen BA, Field DJ. Sparse coding of sensory inputs. *Curr Opin Neurobiol*. 2004;14: 481–487.
21. Winkielman P, Huber DE, Kavanagh L, Schwarz N. Fluency of consistency: When thoughts fit nicely and flow smoothly. *Cognitive Consistency: A Fundamental Principle in Social Cognition*. 2012; 89–111.
22. Holzleitner IJ, Lee AJ, Hahn AC, Kandrik M, Bovet J, Renoult JP, et al. Comparing theory-driven and data-driven attractiveness models using images of real women’s faces. *J Exp Psychol Hum Percept Perform*. 2019;45: 1589–1595.
23. Dibot N, Tio S Mendelson T, Puech W, Renoult J. Sparsity in an artificial neural network predicts beauty: towards a model of processing-based aesthetics. In prep.
24. Ryali CK, Goffin S, Winkielman P, Yu AJ. From likely to likable: The role of statistical typicality in human social assessment of faces. *Proc Natl Acad Sci U S A*. 2020;117: 29371–29380.
25. Briellmann AA, Dayan P. A computational model of aesthetic value. *Psychol Rev*. 2022;129: 1319–1337.
26. Ma DS, Correll J, Wittenbrink B. The Chicago face database: A free stimulus set of faces and norming data. *Behav Res Methods*. 2015;47: 1122–1135.
27. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. 2014. Available: <http://arxiv.org/abs/1409.1556>
28. Parkhi OM, Vedaldi A, Zisserman A. Deep Face Recognition. 2015 [cited 12 Sep 2023]. doi:10.5244/C.29.41
29. Güçlü U, van Gerven MAJ. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *J Neurosci*. 2015;35: 10005–10014.
30. Peterson JC, Abbott JT, Griffiths TL. Evaluating (and Improving) the Correspondence Between Deep Neural Networks and Human Representations. *Cogn Sci*. 2018;42: 2648–2669.
31. Hurley N, Rickard S. Comparing measures of sparsity. 2008 IEEE Workshop on Machine Learning for Signal Processing. IEEE; 2008. doi:10.1109/mlsp.2008.4685455
32. Iigaya K, Yi S, Wahle IA, Tanwisuth S, Cross L, O’Doherty JP. Neural mechanisms underlying the hierarchical construction of perceived aesthetic value. *Nat Commun*. 2023;14: 127.
33. Karkkainen K, Joo J. FairFace: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation. 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE; 2021. doi:10.1109/wacv48630.2021.00159
34. Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*. 2000;42: 80.
35. Potter T, Corneille O. Locating attractiveness in the face space: faces are more attractive when closer to their group prototype. *Psychon Bull Rev*. 2008;15: 615–622.
36. Ryali CK, Yu AJ. Beauty-in-averageness and its contextual modulations: A Bayesian statistical account. *bioRxiv*. bioRxiv; 2018. doi:10.1101/360651
37. Winkielman P, Halberstadt J, Fazendeiro T, Catty S. Prototypes are attractive because they are easy on the

mind. *Psychological Sciences*. 2006;17: 799–806.

38. Quiroga RQ, Kreiman G, Koch C, Fried I. Sparse but not 'grandmother-cell' coding in the medial temporal lobe. *Trends in Cognitive Sciences*. 2008;12: 87–91.

39. Khan K, Attique M, Khan RU, Syed I, Chung T-S. A multi-task framework for facial attributes classification through end-to-end face parsing and Deep Convolutional Neural Networks. *Sensors (Basel)*. 2020;20: 328.

40. Wurtz P, Reber R, Zimmermann TD. The feeling of fluent perception: a single experience from multiple asynchronous sources. *Conscious Cogn*. 2008;17: 171–184.

41. Levin DT. Classifying faces by race: The structure of face categories. *J Exp Psychol Learn Mem Cogn*. 1996;22: 1364–1382.

42. Kondo A, Takahashi K, Watanabe K. Influence of gender membership on sequential decisions of face attractiveness. *Atten Percept Psychophys*. 2013;75: 1347–1352.

43. Kramer RSS, Jones AL, Sharma D. Sequential effects in judgements of attractiveness: the influences of face race and sex. *PLoS One*. 2013;8: e82226.





---

# DISCUSSION GENERALE

---



© Produit par DALL·E

# **I. La fluence comme déterminant de la beauté : Modélisation à travers la prototypicalité et prédiction de l'attractivité**

## **A. Prototypicalité et Féminité**

### **1. Définitions et mesures**

Les premières approches visant à élaborer des « visages prototypiques » pour évaluer leur attractivité ont principalement reposé sur des modifications manuelles d'images. Des techniques informatiques de « morphing » ont été employées pour moyenner des visages ou moduler le degré de masculinité ou de féminité d'une image faciale spécifique, afin d'étudier l'influence de ces traits sur les préférences perçues. Les méthodologies employées sollicitaient des participants, souvent à travers des paradigmes de choix forcé ou des tâches d'évaluation, comme l'attribution de notes d'attractivité. Cependant, bien que ces études aient établi une corrélation entre la féminité/masculinité et l'attractivité, certaines ne définissent pas clairement les concepts de féminité/masculinité (problème 1). De plus, ces techniques ne s'appuient pas sur des mesures empiriques de prototypicalité, comme l'ont souligné plusieurs travaux (Jamin Halberstadt, 2006; Langlois & Roggman, 1990; Gillian Rhodes & Tremewan, 1996) (problème 2). Cela expliquerait pourquoi les liens empiriques entre la féminité (ou la masculinité) faciale phénotypique d'une part et la perception de la féminité (ou de la masculinité), d'autre part, et l'attractivité sont encore mal compris et présentent des incohérences dans la littérature (B. C. Jones et al., 2008; Little, Jones, et al., 2008; Peters et al., 2009; Puts, 2010; Scott et al., 2010) (problème 3).

Ces lacunes nous ont conduit, dans le cadre de cette thèse, à tenter de clarifier et de quantifier la notion de féminité. Pour cela, nous avons distingué la féminité en tant que dimorphisme par rapport aux visages masculins (« femaleness ») de la féminité moyenne des visages féminins (« averageness ») – en réponse au problème 1. Par ailleurs, nous avons pris en compte une troisième dimension, la féminité perçue, évaluée par des juges (« perceived femininity »). Les deux premières dimensions, « femaleness » et « averageness », ont été quantifiées grâce à notre méthode, et toutes deux ont été mises en relation avec des évaluations empiriques de la féminité perçue (perceived femininity) ainsi qu'avec l'attractivité perçue, afin de remédier au problème 1.

*La féminité définie par l'« averageness ».* En nous basant sur diverses études cognitives, nous avons noté que le prototype féminin est défini par la moyenne ou par le mode de la distribution des visages de femmes (A. L. Jones & Jaeger, 2019; Neumann, 1977). Dans le Chapitre 2 de la thèse, ce prototype est décrit comme le « centroïde ». Par conséquent, nous avons conceptualisé *cet* « averageness » à partir de la distance euclidienne par rapport au centroïde féminin.

*La féminité définie par « femaleness ».* Il est possible d'envisager une redéfinition du prototype traditionnellement évoqué en psychologie cognitive. Le concept de prototype pourrait être considéré comme une exagération ou une caricature d'une catégorie perceptive. Des traits intensifiés ou caricaturaux pourraient être plus distinctivement représentatifs d'une catégorie, et grâce à leurs caractères saillants, ils sont cognitivement traités avec fluence (Ameel & Storms, 2006). En ce sens, le prototype « femme » pourrait ne pas correspondre à une moyenne, mais plutôt à une version amplifiée de la féminité. Au-delà de la notion de prototype, certains travaux remettent en question la définition de la féminité comme proximité avec la moyenne des femmes, au contraire la féminité est davantage expliquée avec le dimorphisme sexuel. De plus, la moyenne d'un visage n'est pas l'unique critère déterminant son attirance (DeBruine et al., 2007; Perrett et al., 1994). L'attirance pourrait aussi être influencée par la distance et la direction par rapport au prototype opposé, comme le suggère l'hypothèse du contraste (Rosch & Mervis, 1975). Des études montrent que le degré de dimorphisme sexuel est un meilleur prédicteur de l'attractivité perçue (DeBruine et al., 2007; Perrett et al., 1998). De ce fait, dans le chapitre 2, nous avons intégré la mesure de « femaleness », définie par la distance au centroïde masculin.

Lorsque nous avons confronté nos deux mesures à des évaluations empiriques de féminité (Perceived Femininity) dans le Chapitre 2, la mesure « femaleness » a montré une meilleure corrélation avec les scores de féminité attribués par des juges, en expliquant **26** % de la variance de féminité perçue, tandis que la mesure « averageness » affichait une variance expliquée de **3**%. Ces résultats soutiennent donc l'hypothèse du contraste (Rosch & Mervis, 1975), suggérant que l'attractivité ne découle pas uniquement d'une moyenne. Bien que cette moyenne puisse contribuer à l'attractivité, les visages s'écartant du centroïde masculin semblent être perçus comme plus attractifs.

Mitteroecker et collaborateurs ont également soulevé ces problèmes (1 et 2) et ont entrepris une évaluation des diverses techniques disponibles pour définir un prototype « homme » et évaluer la masculinité (définie par la notion de « maleness »), en relation avec la masculinité perçue (Mitteroecker et al., 2015). Parmi ces méthodes, on a par exemple, le degré de typicalité du sexe masculin (Komori et al., 2011; Puts, 2010; Sanchez-Pages et al., 2014). Ce degré de typicalité peut se mesurer, par exemple, en évaluant l'écart entre la disposition et l'arrangement des traits spécifiques du visage d'un individu masculin (tels que les yeux, le nez, la bouche, et le menton) de ceux d'un visage féminin moyen. En revanche, Mitteroecker et al. (2015) ont proposé une méthode qui positionne chaque visage masculin sur un axe allant de la forme moyenne masculine à la forme moyenne féminine pour évaluer son degré de dimorphisme sexuel. Cette approche, basée sur l'utilisation de deux moyennes de groupe pour définir un axe de différences morphologiques, a également été employée par Komori et al. (2011) et Valenzano et al. (2006). Au-delà des mesures morphométriques traditionnelles, le Modèle Actif d'Apparence (AAM), fusionnant la forme et la texture faciale, permet une analyse plus précise des caractéristiques faciales. A travers cette littérature, ces mesures de masculinité expliquent entre **6** et **25%** de la variance dans les évaluations de masculinité empirique (perceived masculinity) (Koehler et al., 2004; Komori et al., 2011; Mitteroecker et al., 2015; Pound et al., 2009).

La mesure de féminité (« femaleness ») a permis d'expliquer jusqu'à 26% de la variance des évaluations de la féminité perçue, surpassant d'autres méthodes qui couvraient une variance de 6% à 25% pour la masculinité. Toutefois, la prédominance des études sur la masculinité impose une comparaison, malgré les possibles divergences inhérentes aux mesures de féminité et masculinité, nécessitant une interprétation prudente de ces résultats. Dans le corpus existant, une unique étude (Fiala et al., 2021) a distinctement examiné le concept de « femaleness » par rapport à celui de l'« averageness » en lien avec la féminité perçue. Cependant, cette étude, conduite sur cinq ethnies différentes, n'a pas fourni une mesure globale de féminité (« femaleness » ou « averageness ») pour tous ces groupes. Leurs résultats, partiels, ont révélé une différence inter-ethnique notable : la « femaleness » représente 11% de la féminité perçue chez les Colombiennes et 20% chez les Iraniennes. Pour les trois autres groupes, la relation entre « femaleness » et féminité perçue n'était pas significative. A l'inverse, chez les Tchèques, la notion d'« averageness » corrélait

d'avantage avec la féminité perçue, atteignant 22%, et cette même tendance s'est manifestée chez les Turques avec un taux de 12%. En formant des groupes basés sur l'origine ethnique avec les données utilisées dans le manuscrit du Chapitre 2, nous observons une similitude dans la magnitude des effets, lorsque nous mettons en relation la « femaleness » et la féminité perçue chez les femmes asiatiques, latines, africaines et caucasiennes. Ici, il est essentiel d'aborder ces comparaisons avec précaution ; divers facteurs, dont les méthodes utilisées et les spécificités des ensembles de données, peuvent influencer ces écarts.

*Limites de notre méthode.* Notre indicateur de féminité repose largement sur la perception subjective de la féminité (« perceived femininity »). Il est possible que les évaluateurs associent la féminité à des traits dimorphiques accentués, orientant ainsi notre mesure vers la « femaleness » plutôt que vers l'« averageness ». L'utilisation de ces données empiriques introduit un biais potentiel, puisque la féminité, pour certains évaluateurs (voir même la majorité), peut se traduire par une intensification de traits spécifiques plutôt que par une moyenne.

## **2. Corrélation entre féminité et attractivité**

Après avoir établi une définition claire et une méthodologie de quantification de la féminité, nous avons analysé la proportion de variance expliquée par la féminité dans l'attractivité des visages féminins et l'avons mise en parallèle avec les données existantes dans la littérature. Nos résultats sur un jeu de données de photographies standardisées (Chicago Face Dataset) indique que l'« averageness » expliquait seulement 2% de la variance de l'attractivité, tandis que la « femaleness » expliquait 18%. À titre de comparaison, une étude antérieure utilisant la même base de données, mais appliquant le modèle d'apparence actif (AAM), n'a identifié que 4% de variance de l'attractivité expliquée (Ryali et al., 2020). En mettant nos résultats en perspective avec d'autres études sur d'autres jeux de données, certaines ont noté une forte corrélation, atteignant un  $R^2$  de 0.77 (Hoss et al., 2005), entre l'attractivité et la féminité faciale. Ces études ont toutefois utilisé des techniques de morphing facial, modulant ainsi artificiellement la féminité ou la masculinité. Ces méthodes, bien qu'efficaces, éliminent des variations naturelles cruciales pour les impressions faciales. Bien que nos résultats puissent sembler modestes par rapport à certaines données de

la littérature, il est nécessaire de rappeler que nous avons privilégié l'utilisation de stimuli réels, non modifiés, afin de préserver les nuances intrinsèques des visages.

### **3. Sélection de caractéristiques faciales vs. approche automatisée basée sur les CNNs**

Comme je l'ai évoqué précédemment, de nombreuses études s'appuient sur des méthodes morphométriques traditionnelles. Ces méthodes impliquent généralement l'utilisation de points de repère prédéfinis sur le visage, qui sont ensuite mesurés et analysés pour déterminer certaines caractéristiques faciales. Par exemple, la distance entre les yeux ou la largeur du nez pourraient être mesurées à partir de ces points de repère. L'une des méthodes populaires est le modèle d'apparence active (AAM - Active Appearance Model). L'AAM est une technique utilisée pour modéliser l'apparence et la forme des objets, en particulier des visages dans ce contexte. Elle utilise un ensemble d'images d'apprentissage pour former un modèle statistique qui comprend à la fois la forme (les coordonnées des points de repère) et l'apparence (l'intensité des pixels) des objets. Une fois formé, l'AAM peut être ajusté pour s'adapter à de nouvelles images, permettant ainsi une identification et une mesure précises des caractéristiques des images (comme de visages).

Cependant, l'enjeu de ces méthodes réside dans la nécessité d'obtenir des visages rigoureusement standardisés afin de garantir une analyse précise. Tout écart, comme un angle de prise de vue différent ou une expression faciale variée, peut fausser les résultats. De plus, ces méthodes sont moins adaptables et transposables, en particulier lorsqu'il s'agit de les appliquer à des visages non standardisés ou à d'autres espèces où de nouveaux points de repère doivent être redéfinis (Damon et al., 2017). Face aux défis posés par les méthodes traditionnelles de morphologie faciale, nous avons opté pour l'utilisation des réseaux neuronaux convolutifs (CNN). Les CNN, contrairement aux méthodes conventionnelles, peuvent identifier et traiter de manière autonome les caractéristiques pertinentes d'une image sans nécessiter de points de repère prédéfinis. Dans le cadre de notre étude, nous avons utilisé les CNN pour extraire un ensemble de 128 caractéristiques pour chaque visage humain (ou de primates non-humains). Lorsqu'elles sont combinées, ces caractéristiques permettent de positionner chaque visage dans un espace multidimensionnel que nous appelons le « Face Space ». Dans cet espace, chaque dimension correspond à une des 128 caractéristiques et chaque visage est représenté par un vecteur unique à 128

dimensions. Le modèle « Face Space » n'est pas une nouveauté liée aux CNNs, mais il était auparavant basé sur des caractéristiques préétablies. Avec l'utilisation des CNN, nous ne nous reposons plus sur ces caractéristiques pré-établies, mais sur les 128 traits extraits pour représenter chaque visage dans cet espace multidimensionnel. Cette capacité confère aux CNN une flexibilité inégalée, adaptée à divers types de visages et potentiellement extensible à différentes espèces. Néanmoins, le transfert de cette méthode à d'autres espèces requiert un volume conséquent de photos.

Pour gérer la complexité dimensionnelle associée à cet espace, nous avons opté pour l'Analyse en Composantes Principales (ACP). À la suite de cette transformation, la première composante principale (PC<sub>1</sub>) s'est distinguée par sa capacité à discriminer les sexes. Nous l'avons donc utilisée pour distinguer les photos prises de visages masculins de celles prises de visages féminins. L'évaluation de la féminité a été réalisée sur cette PC<sub>1</sub>.

Afin de valider l'efficacité de notre méthode, nous avons analysé des photos de portraits non standardisées avec les CNN. Il est à noter qu'à notre connaissance, il n'existe pas de base de données publique associant portraits non standardisés à des évaluations empiriques de féminité perçue. Par conséquent, notre étude s'est focalisée sur l'évaluation de l'attractivité. Pour ce faire, nous nous sommes appuyés sur SCUT-FBP5500, une base de données de portraits trouvée en ligne, représentant des individus caucasiens et asiatiques dans diverses poses, caractéristiques (comme la coiffure), expressions, et conditions d'éclairage et de fonds. Ces portraits étaient tous étiquetés selon leur attractivité perçue. Nos résultats ont montré des corrélations distinctes :  $R^2 = 0,11$  pour les individus asiatiques et  $R^2 = 0,24$  pour les caucasiens. Ces résultats sont du même ordre de grandeur que ceux obtenus avec des photos standardisées (CFD).

Toutefois, lors de l'analyse globale de la base SCUT-FBP5500, en combinant les données des individus asiatiques et caucasiens, la corrélation est notablement plus faible ( $R = 0,004$ ). Dans le cadre de l'étude des données de visages standardisées (CFD), qui présentent aussi une diversité ethnique, la première composante principale (PC<sub>1</sub>) a conduit à une distinction robuste des visages en fonction du genre. Ainsi, cette dimension a été privilégiée pour la caractérisation de la féminité. Toutefois, la dynamique de discrimination se modifie en étudiant les données de SCUT-FBP5500.



Une exploration distincte des échantillons caucasiens et asiatiques a confirmé que la PC1 préservait son rôle discriminant par genre. Cependant, lors de la fusion des deux sous-groupes ethniques, la PC1 discriminait par l'origine ethnique, plutôt que par le genre. De ce fait, sur la PC1, les centroïdes associés aux catégories Homme et Femme se sont retrouvés superposés pour ce jeu de données, ce qui résulte en une corrélation réduite entre la « femaleness » et l'attractivité. En élargissant l'analyse à la deuxième composante principale (PC2), qui s'est révélée être une dimension discriminante en fonction du genre, la « femaleness » représente 6% de la variance totale de la féminité perçue. Bien que ce chiffre soit inférieur aux 26% enregistrés avec le CFD, cette variation est prévisible compte tenu de la nature non standardisée des images intégrées dans SCUT-FBP5500. Les résultats obtenus démontrent que, selon la composition de l'échantillon, une dimension (telle que PC1) peut perdre sa pertinence discriminative au profit d'une autre caractéristique, ici l'ethnie. Une analyse et une connaissance minutieuse des données est donc nécessaire avant de définir la féminité.

*Pourquoi ne pas adopter l'entièreté des 128 caractéristiques au lieu de la PC qui discrimine le mieux par le sexe ?* L'utilisation complète de ces dimensions nous donne des corrélations de  $R=0.42$  pour le CFD et de  $R=0.23$  pour le SCUT, proches des résultats obtenus sur la dimension la plus discriminante en fonction du genre (soit PC1 pour CFD et PC2 pour SCUT). Toutefois, une telle granularité peut s'avérer excessive dans notre contexte, où l'objectif est de mesurer la féminité au sein d'un réseau dont la tâche est de classer le genre et de ré-identifier une centaine d'individus, engendrant potentiellement une redondance d'informations entre les caractéristiques. De plus, comme le cas de SCUT l'a illustré, certaines caractéristiques qui ne sont pas directement liées à la tâche peuvent émerger spontanément, comme celles associées à l'origine ethnique. L'adoption d'une approche monodimensionnelle permet d'éviter ces complications inhérentes à la multidimensionnalité et réduit également les redondances.

#### **4. Féminité : disparités intra-groupes**

Avec le 2<sup>ème</sup> chapitre de cette thèse, nous avons étudié en profondeur la perception de la féminité, en prenant en compte une vue d'ensemble et en ciblant également des catégories ethniques spécifiques : femmes asiatiques, africaines, caucasiennes et latino-américaines. Contrairement aux conclusions précédemment

établies (Fiala et al., 2021), nos résultats dévoilent des corrélations significatives entre l'attractivité et notre indicateur de féminité, « Femaleness », pour chaque groupe ethnique. Bien que ces corrélations varient, elles démontrent toutes une tendance similaire, suggérant une « universalité » pour l'attractivité de la féminité des femmes.

Afin de comprendre les variations des coefficients de corrélation au sein des différentes ethnies (intra-groupes), nous avons formulé l'hypothèse que le force de corrélation entre féminité et attractivité pourrait être moins forte au sein des groupes ethniques présentant un dimorphisme sexuel moins accentué. Pourtant, après une analyse du dimorphisme facial pour chaque catégorie - en utilisant la distance des centroïdes masculins et féminins - aucune corrélation n'a été trouvée entre le dimorphisme intra-groupe et la force de corrélation (non montré dans le Chapitre 2). La question des fondements biologiques de ces variations demeure non résolue. Face à ces constatations, une approche inter-groupes semble indispensable. Les perceptions et les stéréotypes sont fortement influencés par les comparaisons avec d'autres groupes. Prenons l'exemple des femmes asiatiques. Au sein de leur propre groupe ethnique, les femmes asiatiques pourraient ne pas être perçues comme particulièrement féminines par les observateurs asiatiques, car elles sont évaluées par rapport à une norme interne de ce groupe. Cependant, lorsqu'elles sont comparées à des femmes d'autres groupes ethniques, en raison des stéréotypes associant les Asiatiques à des traits plus féminins (Schug et al., 2015), elles pourraient être perçues comme plus féminines. Dans ce contexte, l'étude de Schug, Alt et Klauer (2015) a démontré que les perceptions au sein d'un groupe pouvaient être fortement influencées par des stéréotypes basés à la fois sur le genre et l'ethnie. Par exemple, à cause de stéréotypes associant les Asiatiques à des traits féminins, les hommes asiatiques sont souvent perçus comme moins représentatifs de leur groupe ethnique. A l'inverse, des traits plus masculins associés aux Africains pourraient conduire à percevoir une femme africaine comme moins typique de son groupe. En se limitant à étudier les perceptions à l'intérieur d'un groupe ethnique spécifique, on pourrait omettre des influences cruciales découlant des stéréotypes inter-groupes. Adopter une telle approche inter-groupes nous permet non seulement de comprendre comment un groupe est perçu en contraste avec d'autres, mais aussi comment ces dynamiques inter-groupes peuvent nuancer ou influencer les perceptions au sein d'un même groupe. Une telle démarche pourrait impliquer d'évaluer le dimorphisme facial inter-ethnie (distance entre le

centroïde Femme de chaque groupe ethnique en rapport avec un centroïde masculin global), puis de corrélérer ces mesures de dimorphisme à la féminité perçue.

Il convient de souligner une limitation notable de notre travail : nous ne disposons pas d'informations sur les caractéristiques des évaluateurs, comme leur âge, leur sexe, leur orientation sexuelle ou leur origine ethnique, éléments qui peuvent influencer la perception et l'évaluation de l'attractivité. Les recherches ultérieures pourraient gagner en précision en intégrant ces variables, affinant ainsi notre compréhension des mécanismes sous-jacents.

## B. De la Féminité comme Prototype à la Typicalité Statistique : Vers une Perspective Élargie

### 1. Typicalité statistique à tous les niveaux de perception

L'utilisation des réseaux neuronaux convolutifs (CNN) a apporté une nouvelle dimension à la compréhension de la prototypicalité faciale. En analysant spécifiquement la féminité obtenue à l'avant-dernière couche du CNN, nous avons émis l'hypothèse que c'est précisément à ce niveau de perception que les catégories sémantiques (et prototypiques) se forment, la couche finale incarnant une structure évoluée de perception sémantique de haut niveau (Long et al., 2015; Rocco et al., 2018; Simonyan & Zisserman, 2015). Néanmoins, une interrogation subsiste : la typicalité faciale ne s'établit-elle pas dès les premiers stades de la perception visuelle ? Pour illustrer cela, prenons l'exemple simplifié d'un neurone qui détecte une orientation spécifique au centre de l'image. Si la plupart des visages possèdent un trait orienté verticalement, ceux qui présentent une orientation horizontale ou diagonale seront considérés comme atypiques. Malgré la présence potentielle de bruit dans ces premiers encodages, ce simple neurone pourrait déjà évoquer la notion de typicalité. Aux niveaux intermédiaires du CNN, les différentes orientations détectées par divers neurones peuvent se combiner pour former des figures plus complexes, telles qu'un triangle. Si la majorité des visages affiche un nez qui ressemble grossièrement à un triangle isocèle, ceux qui présentent un triangle distordu seront jugés atypiques. Ainsi, à mesure que l'on progresse dans les couches du réseau, la prototypicalité s'intensifie. Il est donc intéressant de vérifier cette hypothèse au sein d'un CNN, d'autant plus qu'une telle étude n'a jamais été réalisée. Notre hypothèse initiale propose que la

corrélation entre la typicalité statique des couches finales du CNN et l'attractivité pourrait être plus prononcée que dans les couches antérieures. Il convient d'analyser également si une telle corrélation, bien que potentiellement atténuée, est observable au sein des couches intermédiaires ou initiales du réseau.

Dans le cadre du Chapitre 4 de cette thèse, nous avons évalué la typicalité statistique des portraits de la base de données CFD à travers différentes couches. Ces données n'ayant pas été discutées précédemment, il est opportun de les présenter ici. Nous avons donc pour chaque visage, une valeur d'attractivité d'une part, et une mesure typicalité statistique (LLH) pour à chaque couche, d'autre part. Nous ne trouvons aucune corrélation positive entre les LLH de chacune de ces couches et l'attractivité. Contre toute attente, nous avons constaté une corrélation négative, particulièrement prononcée pour les couches avancées. En utilisant une régression multiple avec les LLH de chaque couche pour estimer l'attractivité, nous avons constaté que ce modèle n'expliquait que 11% de la variance. En segmentant les couches en trois catégories distinctes (bas niveau, niveau intermédiaire, haut niveau), il s'avère que les couches de haut niveau expliquaient 10% de la variance, équivalent au modèle initial avec toutes les couches. Les couches de bas niveau et les couches intermédiaires expliquaient peu l'attractivité (<6%). Notre analyse a révélé que la typicalité, telle qu'exprimée dans les dernières couches du CNN, contribue davantage à expliquer l'attractivité. Il se pourrait que la typicalité ne soit pas discernable ou appréciée aux stades antérieurs, du moins dans le contexte de notre étude.

Notre approche de la typicalité statistique se distingue par l'intégration complète des informations des couches de bas et haut niveaux dans notre régression linéaire. L'étude de Ryali (Ryali et al., 2020), centrée sur les visages et la typicalité statistique, offre une comparaison pertinente. Leur modèle pourrait être assimilé aux couches profondes d'un réseau de neurones, car il englobe à la fois des informations de bas niveau (traits primaires du visage) et de haut niveau (configurations et relations entre ces traits). Nos  $R^2$  comparables (14% pour Ryali et al. 2020 ; 11% pour notre mesure) renforcent cette idée, suggérant que la typicalité statistique des caractéristiques joue probablement un rôle au niveau de la perception des visages et dans la détermination de l'attractivité. Toutefois, la corrélation demeure relativement faible, indiquant que d'autres facteurs peuvent également contribuer de manière significative à la perception de l'attractivité faciale. Dans cette étude, nous avons centré

notre mesure sur la typicalité statistique des visages. Cependant, la méthodologie employée est généralisable et peut être aisément adaptée à d'autres types de prototypes, qu'il s'agisse de catégories d'œuvres d'art ou de chiens.

## **2. Modulation du prototype : adaptabilité basée sur l'expérience**

En matière de perception des visages, le prototype sert de modèle de référence à partir duquel nous évaluons et catégorisons les visages que nous rencontrons. Ainsi, ce que nous considérons comme un visage prototypique n'est pas statique : il peut évoluer en fonction de notre environnement et de notre expérience. Par exemple, le visage prototypique que nous avons en tête peut changer selon les visages que nous voyons régulièrement. Si dans une culture, un certain type de visage est considéré comme prototypique et donc attractif, ce standard peut changer si nous sommes immergés dans une autre culture avec des visages différents et des normes esthétiques différentes. Ainsi, la notion de prototypicalité est étroitement liée à celle de familiarité. L'étude de Principe et Langlois (2012) a mis en évidence ce phénomène. Dans leurs recherches, ils ont familiarisé les participants à des visages hybrides, composés de traits de chimpanzés et d'humains. Après cette familiarisation, ils ont constaté que ces participants avaient tendance à évaluer les visages ayant une certaine proportion de traits de chimpanzé comme étant plus attractifs. Cette observation suggère donc que notre perception de l'attractivité peut être influencée par ce à quoi nous sommes exposés.

## **3. Retour à la fluence : interprétation sous les notions d'efficacité et d'efficience**

Comme nous l'avons évoqué, la fluence perceptive se réfère à la facilité avec laquelle un stimulus est traité et est souvent liée à des évaluations esthétiques positives. Cette facilité sous-tend deux notions : l'efficacité et l'efficience (J. P. Renoult & Mendelson, 2019). Rappelons que l'efficacité désigne un traitement de haute précision avec une minimisation de la perte d'information, tandis que l'efficience se rapporte à une utilisation judicieuse des ressources pour obtenir un résultat optimal. La typicalité statistique, mesurée par le LLH, offre une perspective intéressante à ce propos. Elle peut être envisagée sous les deux angles : un visage typique qui s'aligne étroitement

avec les visages couramment rencontrés, permet un traitement précis et sans erreur (efficace) et en mobilisant peu de neurones (efficient). Cependant, nous avons montré que la typicalité statistique était un prédicteur faible de l'attractivité. Cela nous conduit à examiner d'autres facteurs qui pourraient influencer la fluence perceptive. Dans le Chapitre 4, nous avons montré que la sparsité (concentration de l'activité) est un meilleur prédicteur de l'attractivité des visages. Cela suggère que des traitements plus parcimonieux, symbolisés par une sparsité élevée, peuvent être plus représentatifs de la fluence perceptive.

Pour des perspectives futures, il faudrait examiner de manière approfondie la relation entre la sparsité et le LLH à travers les différentes couches du réseau neuronal. Dans un réseau de neurones, un visage typique, indiqué par un LLH élevé, peut être traité de manière dense dans les couches initiales. Ici, « dense » signifie qu'une grande proportion de neurones est activée (contraire de « sparse »), car ces couches traitent des caractéristiques élémentaires omniprésentes, comme les contours ou les textures. Toutefois, en progressant vers les couches profondes du réseau, ce même visage typique est susceptible d'être traité de manière plus parcimonieuse. De ce fait, bien que la sparsité et la typicalité puissent fluctuer selon la couche en question, il est probable qu'elles soient intrinsèquement connectées, en particulier dans les couches profondes où la précision de la reconnaissance est importante. Nous anticipons donc que les deux mesures soient fortement corrélées positivement dans ces couches finales. Si nous intégrons ces deux mesures dans un modèle prédictif d'attractivité, il est plausible que la sparsité capte une portion significative de la variance expliquée par la typicalité statistique. En d'autres termes, une fois la sparsité prise en compte, la contribution de la typicalité statistique à la prédiction pourrait être atténuée. Si cette hypothèse se vérifie, cela mettrait en évidence le rôle prédominant de la sparsité dans le cadre de la perception de l'attractivité faciale.

## II. Féminité : Conséquences socio-sexuelles et origine biologique et évolutive

### A. Impact de la féminité sur les comportements de primates non-humains

Dans le chapitre 2, nous avons examiné l'impact de la féminité sur les comportements de mandrills issus d'une grande population naturelle (Gabon). L'une des caractéristiques distinctives de notre étude réside dans son approche longitudinale, qui couvre une période allant de la naissance des individus à l'âge adulte. Cette perspective longitudinale diffère des expériences en laboratoire, car elle nous permet d'observer les comportements des mandrills en milieu naturel, dans des contextes sociaux réalistes.

L'analyse des interactions sociales des femelles adultes mandrills (Chapitre 2) révèle une influence significative de la féminité faciale sur divers comportements, indiquant des préférences sociales et sexuelles pour certains types de femelles. En effet, celles présentant des caractéristiques faciales moins féminines ont plus souvent approchées et subissent un taux d'agression plus élevé de la part des mâles mais aussi des femelles de leur groupe et indépendamment de variables confondantes comme l'âge, le rang social, et la saison. De plus, ces femelles moins féminines reçoivent plus de comportements de copulation de la part des mâles, ils ne montent que les femelles plus féminines, contrairement à nos prédictions initiales.

***Coercition et choix des mâles ?*** Ces résultats sont à interpréter à la lumière de nos connaissances acquises sur notre espèce d'étude. Le fait que les femelles plus masculines soient plus agressées suggèrent un choix reproducteur de la part des mâles. En effet, les agressions des mâles envers les femelles mandrills indiquent un mécanisme de coercition sexuelle, où les mâles usent d'agression pour augmenter leur succès reproducteur, souvent en infligeant des blessures ou du stress (Baniel et al., 2017; Muller et al., 2007; Smit et al., 2022). Chez les mandrills étudiés, les femelles en estrus (autour de leur période ovulatoire) reçoivent plus d'agression de la part des mâles que les femelles non fertiles, comme les femelles gestantes ou en lactation. Une agression accrue se traduit par une plus forte probabilité d'accouplement avec les

agresseurs, surtout si les agressions ont été répétées sur des périodes prolongées (incluant des périodes en dehors de la saison de reproduction), suggérant ainsi un processus d'intimidation sur le long-terme, une forme de coercition sexuelle relativement peu reportées (Smit et al., 2022). Ainsi, une agression accrue indique clairement un intérêt pour les femelles et non une forme de « répulsion sociale ». Le fait que les femelles « masculinisées » soient l'objet de plus d'agression de la part des mâles tout au long de l'année, suggèrent que ces femelles constituent des partenaires de reproduction probablement privilégiés par les mâles. Ces résultats soulèvent des questions quant aux qualités potentielles de ces femelles peu féminines, comme des meilleurs soins maternels ou plus généralement une meilleure valeur sélective de leurs progénitures, qui conduisent les mâles à les choisir. Les relations entre masculinité des femelles et valeur sélective vont au delà des questions abordées lors de cette thèse mais constituent des pistes intéressantes de recherches. Alternativement, les femelles moins féminines sont aussi celles qui sont plus agressives envers leurs partenaires socio-sexuels. Ainsi, les mâles coercitifs pourraient devoir agresser plus ces femelles en réponse à leur propre agressivité afin de les contraindre à la reproduction. De la même façon, des analyses comportementales plus détaillées, comme l'analyse de séquences comportementales, où chaque réaction est étudiée en fonction d'une action initiale, sont maintenant nécessaires pour répondre à ces questions.

***Compétition intra-femelles ?*** Les femelles moins féminines sont également plus agressives et reçoivent en retour plus d'agressions de la part de leurs partenaires femelles, qui pourrait être la traduction comportementale de compétition intra-sexuelle. Bien que la compétition entre femelles ait été moins étudiée que celle observée entre mâles pour l'accès à la reproduction, des recherches ont montré que les femelles pouvaient aussi rivaliser pour l'accès à la nourriture, aux partenaires sexuels ou aux soins parentaux (Stockley & Bro-Jørgensen, 2011). Ainsi, si les mâles préfèrent copuler avec des femelles moins féminines, cela pourrait engendrer une agressivité accrue envers elles de la part des femelles. Des études complémentaires sont nécessaires pour déterminer si les femelles moins féminines sont avantagées dans la compétition pour les ressources, quelles qu'elles soient, par rapport aux femelles plus féminines. Il est intéressant de noter que dans notre étude, nous n'avons trouvé aucune corrélation entre féminité des visages femelles et rang de dominance. Pourtant, chez le mandrill, les femelles de haut rang sont préférées par les mâles comme partenaire de



reproduction (J. M. Setchell & Jean Wickings, 2006; Smit et al., 2022) et ont un accès privilégié à la nourriture et une meilleure fitness (Charpentier pers. Com). Ainsi, féminité et rang social pourraient être deux paramètres indépendants fortement liés à la valeur sélective des femelles. Des analyses de long-terme sur les succès reproducteurs de ces femelles permettront de répondre à ces questions.

### ***Impact de la masculinité des mâles sur les choix de la part des femelles ?***

Tandis que nos recherches ont principalement mis l'accent sur les femelles mandrills, il serait également pertinent d'explorer l'impact d'une variation de masculinité des mâles adultes sur les comportements de leurs partenaires socio-sexuels. Les femelles mandrills préfèrent les mâles dominants présentant des colorations faciales vives, ce qui est illustré par divers comportements, notamment : une proximité et un toilettage accrus envers ces mâles, ainsi qu'une coopération et une acceptation élevées des interactions et présentations sexuelles de la part de ces mâles. (J. M. Setchell, 2005). La testostérone joue un rôle essentiel dans la masculinisation du visage des primates (Johnston, 2006; Ian S. Penton-Voak & Chen, 2004). Chez les mâles mandrills, la production de testostérone se manifeste par des variations dans la coloration de la face et des parties génitales (J. M. Setchell, 2005). En effet, le mâle dominant est à la fois plus coloré et présente également les niveaux les plus élevés de testostérone, tandis que les autres mâles, avec des niveaux hormonaux plus modérés, arborent une coloration moins saturée et moins brillante. Ainsi, la préférence des femelles pour les mâles aux visages vivement colorés peut être interprétée comme une préférence pour la masculinité, en raison des liens entre taux de testostérone, coloration faciale, et manifestation de traits masculins proéminents et sélectivement avantageux pour la reproduction (le mâle dominant engendrant en 60 et 70% des descendants chaque année : Charpentier et al. 2005 ; 2020). Par conséquent, il est plausible que les mâles avec des caractéristiques faciales plus masculines soient privilégiés comme partenaires sexuels. Toutefois, cette préférence des femelles pourrait être nuancée par des compromis évolutifs, particulièrement en termes d'investissement parental. En effet, les mâles très masculins pourraient être préférés car ils fournissent un avantage génétique via une progéniture avec de meilleurs gènes, ou parce qu'ils produisent des mâles descendants présentant des caractéristiques préférées par les femelles (sexy-son hypothesis : (Weatherhead & Robertson, 1979)), mais ces mêmes mâles pourraient également être associés à un investissement parental moindre (David M. Buss &

Schmitt, 1993; S. W. Gangestad & Simpson, 2000; Steven W. Gangestad et al., 2007; Honekopp et al., 2007). En effet, chez les humains, les femmes montrent une préférence pour les hommes plus masculins dans le cadre de relations à court terme, tandis que les hommes présentant des traits plus féminins sont souvent choisis pour des relations à long terme (David M. Buss & Shackelford, 2008), grâce à de meilleures qualités paternelles (David M. Buss, 1991). De plus, le choix des femmes est également influencé par la phase du cycle menstruel dans laquelle elles sont. Notamment, au cours de la phase fertile de leur cycle, les femmes montrent une préférence marquée pour les visages plus masculins, mais cette tendance est particulièrement prononcée dans le contexte de relations éphémères à court terme (B. C. Jones et al., 2008; I. S. Penton-Voak et al., 1999; Puts, 2010; Stephen et al., 2012). Enfin, des traits masculins exacerbés sont souvent corrélés à une capacité compétitive élevée (Pound et al., 2009; Windhager et al., 2011), qui se traduit par une compétition intersexuelle forte, une défense territoriale supérieure et un meilleur accès aux ressources (D. M. Buss, 1989; Symons, 1995). Bien que ces traits puissent théoriquement favoriser la préférence des femmes pour ces partenaires, ils pourraient aussi constituer une menace pour ces femmes et leur progéniture en raison d'une agressivité accrue, et ainsi dissuader la sélection de tels mâles (Mehta & Beer, 2010). Ces différents paramètres devront être considérés, et des suivis fins et longitudinaux des choix socio-sexuels et de leurs conséquences sur la fitness des partenaires et de leurs descendants effectués, pour quantifier de façon adéquate l'impact de la masculinité sur les comportements et le choix des femmes.

***Différences entre choix et préférences :*** Chez l'homme, de nombreuses études ont documenté un impact de la féminité des visages sur les préférences déclarées. Cependant, il est important de noter que ces préférences sont souvent exprimées dans des contextes contrôlés de laboratoire et ne reflètent pas nécessairement des comportements observés dans des situations réelles. Par exemple, si les hommes peuvent initialement préférer des visages très féminins, ils pourraient se montrer réticents à les choisir comme partenaires à long terme. Cette divergence découle d'une association couramment perçue entre une féminité élevée et un risque accru d'infidélité (Little et al., 2014). Dans ce contexte, un homme pourrait craindre d'élever un enfant qui n'est pas le sien. De plus, dans des contextes sociaux, comme lors de processus d'embauche, des traits moins féminins pourraient être privilégiés. Ceci est

dû à des préjugés et biais cognitifs tels que l'effet de halo, où la féminité est souvent associée à une compétence moindre (Banchefsky et al., 2016; Sczesny et al., 2006). Ces exemples mettent en évidence la nécessité de distinguer préférences déclarées et comportements réels. Les études longitudinales, malgré leur rareté due à la complexité de suivre des personnes sur de longues périodes, pourraient nous fournir des informations plus détaillées sur les véritables choix effectués dans des contextes naturels et leurs conséquences, comme par exemple sur le succès reproducteur. Par exemple, l'attractivité est un prédicteur du succès reproducteur des femmes mais cette relation n'est pas linéaire (Jokela, 2009). Les données montrent que des femmes jugées attractives (mais pas extrêmement) ont 16 % d'enfants de plus que celles jugées moins attractives. Cependant, les femmes qui sont perçues comme étant parmi les plus attractives (quartile supérieur) ont seulement 6 % d'enfants de plus que celles du quartile inférieur (Jokela, 2009). Cette étude de Jokela met en évidence la nuance et la complexité des interactions entre attractivité, féminité et succès reproducteur. Cette observation n'est pas propre aux femmes. Les hommes très masculins ne sont pas systématiquement préférés ou sélectionnés comme partenaires. Les raisons sont multiples : des hommes à masculinité élevée peuvent être perçus comme étant moins chaleureux, moins fiables et moins investis dans leur rôle parental (Boothroyd et al., 2011; Gray et al., 2002; Little, Burriss, et al., 2008). Par ailleurs, une masculinité accentuée peut être associée à des traits telles que l'agressivité et la dominance et les femmes sont exposées à un risque d'agression en les choisissant (Li et al., 2014). En somme, tout comme pour la féminité, la masculinité, en tant que facteur d'attractivité, est multifactorielle et dépend largement du contexte socioculturel ainsi que des préférences individuelles. L'un des avantages notables de travailler avec des espèces animales autres que l'Homme réside dans l'élimination de l'aspect déclaratif, intrinsèque aux études humaines qui s'appuient sur les réponses aux questionnaires et les nombreux biais associés qui en découlent, rendant ainsi les termes de « préférence » et « choix » synonymes lors de l'étude des comportements socio-sexuels de primates non-humains.

## B. Préférences des Mandrills pour les Femelles Moins Féminines : Origine évolutive

L'introduction de cette thèse a abordé les origines des préférences esthétiques chez les animaux, explorant deux paradigmes majeurs : la théorie des indicateurs de qualité et la théorie de la fluence.

Dans le cas des mandrills, nos observations ne tendent pas à valider la théorie de la fluence qui propose que des individus « prototypiques » seraient naturellement privilégiés en raison de leur traitement perceptif fluent. Ainsi, on s'attendrait à ce que les femelles mandrills qui correspondent le mieux au prototype féminin bénéficient d'avantages sociaux. Pourtant, nos observations montrent que les femelles mandrills moins féminines (considérées comme atypiques) sont en réalité l'objet de comportements biaisés de la part de leurs conspécifiques : elles sont plus choisies comme partenaires sexuels et sont plus souvent associées spatialement à leurs partenaires mâles et femelles, mais elles subissent également plus d'agressions.

La théorie des indicateurs de qualité postule que les préférences en matière de sélection de partenaires sont guidées par des signaux reflétant la qualité intrinsèque d'un individu. Chez les humains, la féminité est associée à des indicateurs tels que le taux d'estrogènes, qui est corrélé à la fertilité ou aux compétences maternelles (Law Smith et al., 2012; Little et al., 2011; Smith et al., 2006). Ainsi, chez les mandrills, les femelles les plus dimorphiques, avec des traits distinctement féminins, devraient être favorisées. Cependant, nos observations montrent que les mâles mandrills manifestent une préférence pour les femelles affichant des traits masculinisés. Ces traits moins féminins pourraient être corrélés à des attributs potentiellement avantageux, tels que la capacité à rivaliser pour des ressources ou une défense optimisée contre les prédateurs, élevant ainsi la question de savoir si, chez les mandrills par exemple, les femelles exhibant des caractéristiques plus masculines jouissent d'une valeur sélective supérieure par rapport à leurs consœurs moins masculines. De façon similaire chez les humains, dans des environnements plus difficiles ou hostiles, des femmes présentant des traits plus masculins sont valorisées pour leur possible résilience accrue et leur capacité à gérer les stress environnementaux (Little et al., 2006). Cette tendance à privilégier des traits alternatifs, parfois en contradiction avec les prédictions, peut être conceptualisée comme un « renversement de la valence », illustrant la dynamique et

l'adaptabilité des stratégies de sélection de partenaires face aux variations environnementales et écologiques. A titre illustratif, certaines populations de poissons du genre *Xiphophorus* ont vu leurs préférences pour des traits particuliers évoluer. (Höbel & Rodríguez, 2022; Rosenthal, 2018) La préférence ancestrale pour des épées – extensions pigmentées de la nageoire caudale – a été affaiblie, perdue (Rosenthal et al., 2002) ou renversée (Wong & Rosenthal, 2006) dans différentes lignées phylogénétiques. Un second exemple provenant des mouches Diopsidae, caractérisées par des yeux situés aux extrémités de pédoncules oculaires, illustre comment un renversement rapide en réponse à la sélection pourrait expliquer la diversité des préférences à une échelle phylogénétique (Wilkinson & Reillo, 1994) . Les auteurs ont sélectionné la longueur des pédoncules oculaires sexuellement dimorphiques afin d'examiner l'évolution des préférences. Bien que les pédoncules plus longs soient, a priori, plus visibles et donc favorisés, les femelles des lignées de contrôle et celles sélectionnées pour la longueur des pédoncules ont préféré les mâles avec des pédoncules longs, alors que les femelles des lignées sélectionnées pour des pédoncules plus courts ont préféré les mâles avec des pédoncules plus courts. On peut donc supposer une inversion dans la valeur accordée à la féminité observée chez les mandrills. Cela ouvre de nouvelles perspectives de recherche. Il serait intéressant d'examiner si ces traits masculins confèrent des avantages dans la compétition pour les ressources et impactent le succès reproducteur des femelles ainsi que leur survie et celle de leurs descendants. Il serait également pertinent d'explorer les mécanismes sous-jacents tels que les facteurs génétiques ou environnementaux, qui favorisent l'apparition de ces traits masculins chez les femelles.

Les résultats de cette thèse ne permettent pas de confirmer de manière concluante la théorie de la fluence. Bien que nos données semblent incliner en faveur de la théorie des indicateurs de qualité (si les femelles masculines présentent des avantages sélectifs), il est à noter que les mécanismes décrits par ces deux théories peuvent opérer conjointement. Des recherches en esthétique empirique montrent que la fluence intensifie l'attractivité d'un stimulus déjà évalué positivement, mais cet effet s'atténue ou renforce l'aversion lorsque la valence est renversée (Albrecht & Carbon, 2014). En effet, si une catégorie perceptuelle est associée à des attributs négatifs – prenons l'exemple des hooligans associés à l'agressivité – un individu prototypique de cette catégorie aura tendance à se voir attribuer davantage d'attributs négatifs et

conduire les autres personnes à l'éviter. Pour le formuler autrement, un prototype très représentatif d'une catégorie, "hérite" davantage des propriétés (qu'elles soient positives ou négatives) de cette catégorie que des exemplaires atypiques ou moins représentatifs. Relions cela au cas des mandrills : si, dans leur contexte social, le prototype « Femelle » est, par défaut, évalué négativement en raison d'une valence négative intrinsèque à la catégorie (par exemple, du point de vue des mâles), ce mécanisme de fluence et d'héritage des attributs de la catégorie pourrait expliquer ou renforcer des comportements ou des perceptions négatives envers les femelles typiques de la catégorie. Toutefois, ce qui ressort clairement de nos recherches, c'est que chez certains primates non-humains, la perception de la féminité influence l'attractivité, avec une possible valorisation des attributs masculins.

Dans la continuité de ces observations, une approche comparative inter-espèces pourrait nous éclairer, en offrant un moyen de distinguer entre les préférences universelles liées à des mécanismes perceptifs communs (entre des espèces phylogénétiquement proches) de celles qui sont le fruit de pressions sélectives spécifiques à chaque espèce. Une proposition expérimentale, par exemple, consisterait à exposer des mandrills mâles à des visages humains présentant divers niveaux de féminité. Si ces primates manifestent une préférence marquée pour des visages humains très féminins, cela indiquerait un mécanisme perceptif général, non spécifiquement ancré dans des enjeux reproductifs. Dans un contexte inverse, si des humains sont confrontés à des images de mandrills de divers niveaux de féminité et manifestent une préférence pour des visages plus féminins, cela renforcerait la théorie selon laquelle la préférence pour la féminité est basée sur une origine perceptive et non uniquement sur des pressions évolutives liées à la sélection sexuelle. Cette idée est inspirée par les travaux de Damon et collaborateurs (Damon et al., 2017). Ces chercheurs ont étudié si les évaluations d'attractivité par des juges humains pouvaient anticiper les préférences visuelles des humains et des macaques rhésus. En analysant le temps de regard (mesure de préférence visuelle), ils cherchaient à savoir si les participants, humains ou macaques, étaient attirés par les visages jugés attractifs et si cette attirance était spécifique à chaque espèce. Les résultats ont montré que les participants humains ont regardé plus longtemps les visages humains attractifs, mais pas les visages de singes attractifs. À l'inverse, les macaques ont regardé plus longtemps les visages de macaques attractifs, mais pas les visages humains attractifs.

Cependant, les jugements humains sur l'attractivité des visages de singes prédisaient les préférences visuelles des macaques, suggérant une concordance partielle entre les caractéristiques évaluées par les humains et celles qui attirent l'attention visuelle chez les macaques. Appliquée aux mandrills et aux humains, une telle étude pourrait permettre de rendre compte de l'existence d'un mécanisme de traitement commun à toutes les espèces, conduisant à des préférences pour la prototypicalité chez les primates.

### C. Origine de la variation de la féminité des visages chez les mandrills juvéniles

Dans le chapitre 3, nous avons étudié les déterminants de la variabilité de la féminité chez les mâles mandrills juvéniles, en se concentrant sur deux hypothèses principales : l'influence du nombre de grands frères et l'impact de la condition maternelle en relation avec la théorie de Trivers-Willard (Trivers & Willard, 1973).

L'effet du nombre de grands frères sur la féminité faciale est soutenu par une étude unique menée chez l'Homme (Nila et al., 2019). En effet, la plupart des études ont examiné l'influence du nombre de grands frères sur l'homosexualité masculine, un proxy plutôt « grossier » de la féminité. Dans l'étude de Nila et collègues (2019), l'association entre féminité et nombre de grand frère est hypothétiquement attribuée à la production d'antigènes anti-mâles lors de gestations successives. Toutefois, chez les mandrills étudiés, ni le rang de naissance ni le nombre de grands frères n'ont montré d'effet sur la féminité faciale des mâles juvéniles.

Au contraire, nous trouvons que le nombre de grandes sœurs explique en partie la féminisation des visages de jeunes mâles mandrills. La théorie de Trivers-Willard (Trivers & Willard, 1973) postule que la condition physiologique de la mère peut influencer le sexe de son descendant. Cette hypothèse prédit que dans des conditions favorables, les mères produisent des descendants du sexe qui profitera le plus de ces conditions. Dans le contexte des mandrills, cela suggère que les mères en meilleure condition pourraient être plus à même de supporter le coût énergétique supposé plus élevé de production de mâles, comme démontré chez de nombreuses espèces de primates (Jasienska et al., 2017; Muehlenbein & Bribiescas, 2005; Tamimi et al.,

2003). Notre étude a révélé que les fils nés de mères plus jeunes étaient plus féminins, ce qui appuie l'idée que la production d'un fils pourrait être énergétiquement plus coûteuse. Cependant, nous n'avons pas observé d'effet du rang de dominance des mères sur la féminisation faciale de leurs fils, bien que des études antérieures aient indiqué que les femelles de haut rang soient souvent en meilleure condition.

De plus, étant donné le rôle connu des hormones dans la détermination du sexe et dans les variations de traits masculins ou féminins (Clemens, 1978; Hines, 2011; Thornton et al., 2009; van de Beek et al., 2009), une analyse des niveaux hormonaux chez les mères pendant leur gestation serait pertinente. Considérer d'autres indicateurs de la condition maternelle, tels que le poids et l'exposition au parasitisme, pourrait également nous éclairer davantage sur la propension des mères mandrills en conditions moins favorable à privilégier la naissance de femelles ou de mâles féminisés, comme prédit par la théorie de Trivers-Willard.

Dans le chapitre 3, nous avons également examiné si la féminité faciale des mâles juvéniles influençait leurs propres comportements jugés féminins ou masculins. Nos résultats n'ont montré aucune corrélation entre féminité et comportements sociaux. Toutefois, il serait maintenant intéressant d'étudier si ces mâles féminisés sont socialement moins bien (ou mieux) intégrés et les comportements reçus de la part de leurs conspécifiques.



## CONCLUSION GENERALE

Cette thèse s'est principalement intéressée à la complexité inhérente au concept de beauté, le scrutant à travers le prisme de la théorie de la fluence et, plus précisément, de la prototypicalité, perçue comme une composante intégrale de celle-ci. Tout au long de cette thèse, j'ai essayé de démontrer que la beauté s'étendait au-delà de la vision humaine. Considérer la beauté comme une construction strictement humaine érige des barrières dans l'explication de l'évolution et de la diversité des ornements sexuels attractifs à travers les espèces. Alors que la théorie des indicateurs de qualité avance une perspective selon laquelle l'évolution de signaux esthétiques, circonscrits dans leur attrait au sein de leur propre espèce, est le fruit de la sélection sexuelle, la théorie de la fluence émet l'hypothèse d'une universalité de la perception de la beauté. L'exploration des mécanismes neuronaux qui sous-tendent notre perception de la beauté et leur possible partage avec d'autres espèces demeure un domaine nouveau de recherches à explorer. Cette thèse postule que la fluence, potentiellement un attribut commun parmi les espèces dotées de systèmes visuels, sculpte notre perception esthétique en favorisant ce qui est traité de manière fluente. Utilisant les CNNs pour disséquer la beauté et examiner les métriques de la fluence, en prenant la féminité comme un cas spécifique de la prototypicalité et intégrant également la typicalité statistique et la sparsité, cette recherche propose des pistes de réflexion afin de comprendre la manière dont la beauté est perçue et est neurologiquement traitée. La question se pose désormais de savoir comment les modèles de fluence et les considérations évolutionnistes peuvent être intégrés plus globalement avec les facteurs socioculturels et individuels. Les recherches futures sont appelées à explorer plus en profondeur les mécanismes de perception sous-jacents et à étendre ces concepts à un éventail plus large d'espèces et de contextes environnementaux.

Par ailleurs, une problématique s'est dégagée entre les préférences mesurées en laboratoire et les véritables choix effectués en contexte naturel. Même si les choix peuvent être guidés par les préférences, les signaux préférés ne conduisent pas toujours à des choix comportementaux manifestes. Cette divergence nous amène à la distinction entre « aimer » et « choisir », deux facettes d'un même système de récompense mais médiées par des chemins neurochimiques différents. Le cas des signaux esthétiques offre une application pratique de cette distinction : bien qu'un signal puisse être « aimé » pour son traitement neuronal fluide, et donc considéré comme beau, cela ne se

traduit pas nécessairement par un acte d'accouplement qui serait également « aimé » ou « choisi ». Cette dynamique entre le fait « d'aimer » et de « choisir » façonne non seulement les interactions intra-espèces mais pose également des questions sur les mécanismes évolutifs à l'œuvre dans le choix de partenaires. Grâce à l'étude de primates non-humains vivant en conditions naturelles, nous avons contourné la distinction entre choisir et préférer. Toutefois, il serait bénéfique de valider nos résultats empiriques en menant des études en laboratoire portant sur le temps de regard ou sur des mesures physiologiques telles que les réponses zygomatiques, en fonction de visages de femelles mandrills masculinisées ou féminisées artificiellement à l'aide de réseaux génératifs antagonistes (GAN).

En somme, ce travail a tracé un premier chemin vers une exploration plus profonde des liens entre esthétique, cognition et évolution, établissant ainsi un pont entre la perception sensorielle et les dynamiques évolutives. Les investigations futures, à l'intersection des neurosciences et de la biologie évolutive, pourront s'appuyer sur cette base pour continuer à déchiffrer les mystères de la beauté dans le règne animal.

---

# BIBLIOGRAPHIE

---

- Abernethy, K. A., White, L. J. T., & Wickings, E. J. (2002). Hordes of mandrills ( *Mandrillus sphinx* ): extreme group size and seasonal male presence. *Journal of Zoology (London, England: 1987)*, 258(1), 131–137.
- Albrecht, S., & Carbon, C.-C. (2014). The Fluency Amplification Model: fluent stimuli show more intense but not evidently more positive evaluations. *Acta Psychologica*, 148, 195–203.
- Altmann, J. (1974). *Observational Study of Behavior: Sampling Methods*.
- Ameel, E., & Storms, G. (2006). From prototypes to caricatures: Geometrical models for concept typicality. *Journal of Memory and Language*, 55(3), 402–421.
- Andersson, M. B. (1994). *Sexual selection*. Princeton University Press.
- Banchefsky, S., Westfall, J., Park, B., & Judd, C. M. (2016). But you don't look like A scientist!: Women scientists with feminine appearance are deemed less likely to be scientists. *Sex Roles*, 75(3–4), 95–109.
- Baniel, A., Cowlshaw, G., & Huchard, E. (2017). Male violence and sexual intimidation in a wild primate society. *Current Biology: CB*, 27(14), 2163-2168.e3.
- Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In *Sensory Communication* (pp. 216–234). The MIT Press.
- Beardsley, M. C. (1981). *Aesthetics*. Hackett Publishing.
- Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language--a quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia*, 50(11), 2669–2683.
- Boothroyd, L. G., Cross, C. P., Gray, A. W., Coombes, C., & Gregson-Curtis, K. (2011). Perceiving the facial correlates of sociosexuality: Further evidence. *Personality and Individual Differences*, 50(3), 422–425.

- Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology*, 63(4), 545–552.
- Brockmeyer, T., Kappeler, P. M., Willaume, E., Benoit, L., Mboumba, S., & Charpentier, M. J. E. (2015). Social organization and space use of a wild mandrill (*Mandrillus sphinx*) group. *American Journal of Primatology*, 77(10), 1036–1048.
- Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses testing in 37 cultures. *Behavioral and Brain Sciences*, 12, 1–49.
- Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *The American Psychologist*, 53(5), 533–548.
- Buss, David M. (1991). Mate selection for good parenting skills. *The Behavioral and Brain Sciences*, 14(3), 520–521.
- Buss, David M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology*, 50(3), 559–570.
- Buss, David M., & Schmitt, D. P. (1993). Sexual Strategies Theory: An evolutionary perspective on human mating. *Psychological Review*, 100(2), 204–232.
- Buss, David M., & Shackelford, T. K. (2008). Attractive women want it all: Good genes, economic investment, parenting proclivities, and emotional commitment. *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior*, 6(1), 147470490800600.
- Cadena, S. A., Denfield, G. H., Walker, E. Y., Gatys, L. A., Tolia, A. S., Bethge, M., & Ecker, A. S. (2019). Deep convolutional models improve predictions of macaque V1 responses to natural images. *PLoS Computational Biology*, 15(4), e1006897.
- Cadieu, C., Kouh, M., Pasupathy, A., Connor, C. E., Riesenhuber, M., & Poggio, T. (2007). A model of V4 shape selectivity and invariance. *Journal of Neurophysiology*, 98(3), 1733–1750.

- Charpentier, Marie, Peignot, P., Hossaert-McKey, M., Gimenez, O., Setchell, J. M., & Wickings, E. J. (2005). Constraints on control: factors influencing reproductive success in male mandrills (*Mandrillus sphinx*). *Behavioral Ecology: Official Journal of the International Society for Behavioral Ecology*, *16*(3), 614–623.
- Charpentier, Mje, Harté M Poirotte C De Bellefon, J. M., Laubi, B., Kappeler, P. M., & Renoult, J. P. (2020). Same father, same face: deep learning reveals selection for signaling kinship in a wild primate *Science Advances* *6*:eaba3274. *Science Advances*, *6*.
- Chen, M. K., & Risen, J. L. (2010). How choice affects and reflects preferences: revisiting the free-choice paradigm. *Journal of Personality and Social Psychology*, *99*(4), 573–594.
- Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., & Oliva, A. (2016). Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Scientific Reports*, *6*, 27755.
- Clemens, L. (1978). Prenatal endogenous androgenic influences on masculine sexual behavior and genital morphology in male and female rats\*1. *Hormones and Behavior*, *10*(1), 40–53.
- Damon, F., Méary, D., Quinn, P. C., Lee, K., Simpson, E. A., Paukner, A., Suomi, S. J., & Pascalis, O. (2017). Preference for facial averageness: Evidence for a common mechanism in human and macaque infants. *Scientific Reports*, *7*, 46303.
- DeBruine, L. M., Jones, B. C., Unger, L., Little, A. C., & Feinberg, D. R. (2007). Dissociating averageness and attractiveness: attractive faces are not always average. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(6), 1420–1430.
- Dewey, J. (1934). *Art As Experience*.
- Dezeure, J., Charpentier, M. J. E., & Huchard, E. (2022). Fitness effects of seasonal birth timing in a long-lived social primate living in the equatorial forest. In *Animal Behaviour* (Vol. 185, pp. 113–126). <https://doi.org/10.1016/j.anbehav.2022.01.002>

- Dirks, W., Lemmers, S. A. M., Ngoubangoye, B., Herbert, A., & Setchell, J. M. (2020). Odontochronologies in male and female mandrills (*Mandrillus sphinx*) and the development of dental sexual dimorphism. *American Journal of Physical Anthropology*, *172*(4), 528–544.
- Dragojevic, M., & Giles, H. (2016). I don't like you because you're hard to understand: The role of processing fluency in the language attitudes process. *Human Communication Research*, *42*(3), 396–420.
- Dubey, A., Jayadeva, & Agarwal, S. (2016). Examining representational similarity in ConvNets and the primate visual cortex. In *arXiv [cs.CV]*. arXiv. <http://arxiv.org/abs/1609.03529>
- Dubey, R., Peterson, J., Khosla, A., Yang, M.-H., & Ghanem, B. (2015, December). What makes an object memorable? *2015 IEEE International Conference on Computer Vision (ICCV)*. 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile. <https://doi.org/10.1109/iccv.2015.130>
- Dubuc, C., Allen, W. L., Maestriperi, D., & Higham, J. P. (2014). Is male rhesus macaque red color ornamentation attractive to females? *Behavioral Ecology and Sociobiology*, *68*(7), 1215–1224.
- Fantz, R. L. (1958). Pattern vision in young infants. *The Psychological Record*, *8*(2), 43–47.
- Fechner, G. T. (1876). *Vorschule der Ästhetik*.
- Fechner, G. T. (2016). *Vorschule Der Aesthetik*. Hansebooks.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *1*(1), 1–47.
- Fiala, V., Třebický, V., Pazhoohi, F., Leongómez, J. D., Tureček, P., Saribay, S. A., Akoko, R. M., & Kleisner, K. (2021). Facial attractiveness and preference of sexual dimorphism: A comparison across five populations. *Evolutionary Human Sciences*, *3*(e38), 1–37.

- Forster, M., Gerger, G., & Leder, H. (2015). Everything's relative? Relative differences in processing fluency and the effects on liking. *PloS One*, *10*(8), e0135944.
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, *36*(4), 193–202.
- Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: trade-offs and strategic pluralism. *The Behavioral and Brain Sciences*, *23*(4), 573–587; discussion 587-644.
- Gangestad, Steven W., Garver-Apgar, C. E., Simpson, J. A., & Cousins, A. J. (2007). Changes in women's mate preferences across the ovulatory cycle. *Journal of Personality and Social Psychology*, *92*(1), 151–163.
- Garcia, C. M., & Ramirez, E. (2005). Evidence that sensory traps can evolve into honest signals. *Nature*, *434*(7032), 501–505.
- Ghirlanda, S., Jansson, L., & Enquist, M. (2002). Chickens prefer beautiful humans. *Human Nature*, *13*(3), 383–389.
- Gomez, P., & Torelli, C. J. (2015). It's not just numbers: Cultural identities influence how nutrition information influences the valuation of foods. *Journal of Consumer Psychology: The Official Journal of the Society for Consumer Psychology*, *25*(3), 404–415.
- Graf, L. K. M., Mayer, S., & Landwehr, J. R. (2018). Measuring processing fluency: One versus five items. *Journal of Consumer Psychology: The Official Journal of the Society for Consumer Psychology*, *28*(3), 393–411.
- Gray, P. B., Kahlenberg, S. M., Barrett, E. S., Lipson, S. F., & Ellison, P. T. (2002). Marriage and fatherhood are associated with lower testosterone in males. *Evolution and Human*



- Behavior: Official Journal of the Human Behavior and Evolution Society*, 23(3), 193–201.
- Güçlü, U., & van Gerven, M. A. J. (2015). Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 35(27), 10005–10014.
- Halberstadt, J., & Rhodes, G. (2000). The attractiveness of nonface averages: implications for an evolutionary explanation of the attractiveness of average faces. *Psychological Science*, 11(4), 285–289.
- Halberstadt, Jamin. (2006). The generality and ultimate origins of the attractiveness of prototypes. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, 10(2), 166–183.
- Halberstadt, Jamin, & Rhodes, G. (2003). It's not just average faces that are attractive: computer-manipulated averageness makes birds, fish, and automobiles attractive. *Psychonomic Bulletin & Review*, 10(1), 149–156.
- Harmon-Jones, E., & Allen, J. J. B. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual differences approaches. *Personality & Social Psychology Bulletin*, 27(7), 889–898.
- Hines, M. (2011). Prenatal endocrine influences on sexual orientation and on sexually differentiated childhood behavior. *Frontiers in Neuroendocrinology*, 32(2), 170–182.
- Höbel, G., & Rodríguez, R. L. (2022). Positive-to-negative behavioural responses suggest hedonic evaluation in treefrog mate choice. *Proceedings. Biological Sciences*, 289(1967), 20211822.
- Holzleitner I.J., Lee A.J., Hahn A.C., Kandrik M., Bovet J., Renoult J., Simmons D., Garrod O., DeBruine L.M., Jones B.C. (2018). A data-driven model of women's facial

- attractiveness reliably outperforms theory-driven models. *PsyArXiv Preprints*.  
<https://doi.org/10.31234/osf.io/vhc5k>
- Honekopp, J., Rudolph, U., Beier, L., Liebert, A., & Muller, C. (2007). Physical attractiveness of face and body as indicators of physical fitness in men. *Evolution and Human Behavior: Official Journal of the Human Behavior and Evolution Society*, 28(2), 106–111.
- Hoss, R. A., Ramsey, J. L., Griffin, A. M., & Langlois, J. H. (2005). The role of facial attractiveness and facial masculinity/femininity in sex classification of faces. *Perception*, 34(12), 1459–1474.
- Hubel, D. H., & Wiesel, T. N. (1960). Receptive fields of optic nerve fibres in the spider monkey. *The Journal of Physiology*, 154(3), 572–580.
- Hulse, S. V., Renoult, J. P., & Mendelson, T. C. (2020). Sexual signaling pattern correlates with habitat pattern in visually ornamented fishes. *Nature Communications*, 11(1), 2561.
- Hyvarinen, A. (2009). *Natural image statistics: A probabilistic approach to early computational vision*. Springer.
- ImageNet large scale visual recognition competition 2012 (ILSVRC2012)*. (2012).  
<https://image-net.org/challenges/LSVRC/2012/results.html>
- Jasienska, G., Bribiescas, R. G., Furberg, A.-S., Helle, S., & Núñez-de la Mora, A. (2017). Human reproduction and health: an evolutionary perspective. *Lancet*, 390(10093), 510–520.
- Johnston, V. S. (2006). Mate choice decisions: the role of facial beauty. *Trends in Cognitive Sciences*, 10(1), 9–13.
- Jokela, M. (2009). Physical attractiveness and reproductive success in humans: Evidence from the late 20 century United States. *Evolution and Human Behavior: Official Journal of*

*the Human Behavior and Evolution Society*, 30(5).

<https://doi.org/10.1016/j.evolhumbehav.2009.03.006>

- Jones, A. L., & Jaeger, B. (2019). Biological Bases of Beauty Revisited: The Effect of Symmetry, Averageness, and Sexual Dimorphism on Female Facial Attractiveness. *Symmetry*, 11(2), 279.
- Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R., & Law Smith, M. J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior*, 37(1), 78–84.
- Kingdon, J., Happold, D., Butynski, T., Hoffmann, M., Happold, M., & Kalina, J. (2013). *Mammals of Africa*. Bloomsbury Natural History.
- Koehler, N., Simmons, L. W., Rhodes, G., & Peters, M. (2004). The relationship between sexual dimorphism in human faces and fluctuating asymmetry. *Proceedings. Biological Sciences*, 271 Suppl 4(suppl\_4), S233-6.
- Komori, M., Kawamura, S., & Ishihara, S. (2011). Multiple mechanisms in the perception of face gender: Effect of sex-irrelevant features. *Journal of Experimental Psychology. Human Perception and Performance*, 37(3), 626–633.
- Kranz, F., & Ishai, A. (2006). Face perception is modulated by sexual preference. *Current Biology: CB*, 16(1), 63–68.
- Kriegeskorte, N. (2015). Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. *Annual Review of Vision Science*, 1, 417–446.
- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, 1(2), 115–121.
- Law Smith, M. J., Dedy, D. K., Moore, F. R., Jones, B. C., Cornwell, R. E., Stirrat, M., Lawson, J. F., Feinberg, D. R., & Perrett, D. I. (2012). Maternal tendencies in women are

- associated with estrogen levels and facial femininity. *Hormones and Behavior*, 61(1), 12–16.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4), 541–551.
- Li, Y., Bailey, D. H., Winegard, B., Puts, D. A., Welling, L. L. M., & Geary, D. C. (2014). Women's preference for masculine traits is disrupted by images of male-on-female aggression. *PloS One*, 9(10), e110497.
- Lindsay, G. W. (2021). Convolutional Neural Networks as a Model of the Visual System: Past, Present, and Future. *Journal of Cognitive Neuroscience*, 33(10), 2017–2031.
- Little, A. C., Burriss, R. P., Jones, B. C., DeBruine, L. M., & Caldwell, C. A. (2008). Social influence in human face preference: men and women are influenced more for long-term than short-term attractiveness decisions. *Evolution and Human Behavior: Official Journal of the Human Behavior and Evolution Society*, 29(2), 140–146.
- Little, A. C., Cohen, D. L., Jones, B. C., & Belsky, J. (2006). Human preferences for facial masculinity change with relationship type and environmental harshness. *Behavioral Ecology and Sociobiology*, 61(6), 967–973.
- Little, A. C., Jones, B. C., & DeBruine, L. M. (2008). Preferences for variation in masculinity in real male faces change across the menstrual cycle: Women prefer more masculine faces when they are more fertile. *Personality and Individual Differences*, 45(6), 478–482.
- Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. <https://doi.org/10.1098/rstb.2010.0404>

- Little, A. C., Jones, B. C., Feinberg, D. R., & Perrett, D. I. (2014). Men's strategic preferences for femininity in female faces. *British Journal of Psychology (London, England: 1953)*, *105*(3), 364–381.
- Long, J., Shelhamer, E., & Darrell, T. (2015, June). Fully convolutional networks for semantic segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA. <https://doi.org/10.1109/cvpr.2015.7298965>
- Mehta, P. H., & Beer, J. (2010). Neural mechanisms of the testosterone-aggression relation: the role of orbitofrontal cortex. *Journal of Cognitive Neuroscience*, *22*(10), 2357–2368.
- Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*(1), 89–115.
- Metz-Lutz, M.-N., Bressan, Y., Heider, N., & Otzenberger, H. (2010). What physiological changes and cerebral traces tell us about adhesion to fiction during theater-watching". *Frontiers in Human Neuroscience*, *4*.
- Mitteroecker, P., Windhager, S., Müller, G. B., & Schaefer, K. (2015). The morphometrics of “masculinity” in human faces. *PLoS One*, *10*(2), e0118374.
- Muehlenbein, M. P., & Bribiescas, R. G. (2005). Testosterone-mediated immune functions and male life histories. *American Journal of Human Biology: The Official Journal of the Human Biology Council*, *17*(5), 527–558.
- Muller, M. N., Kahlenberg, S. M., Emery Thompson, M., & Wrangham, R. W. (2007). Male coercion and the costs of promiscuous mating for female chimpanzees. *Proceedings. Biological Sciences*, *274*(1612), 1009–1014.
- Mushtaq, F., Bland, A. R., & Schaefer, A. (2011). Uncertainty and cognitive control. *Frontiers in Psychology*, *2*, 249.

- Naug, D., & Arathi, H. S. (2007). Receiver bias for exaggerated signals in honeybees and its implications for the evolution of floral displays. *Biology Letters*.  
<https://doi.org/10.1098/rsbl.2007.0436>
- Neumann, P. G. (1977). Visual prototype formation with discontinuous representation of dimensions of variability. *Memory & Cognition*, 5(2), 187–197.
- Nila, S., Crochet, P.-A., Barthes, J., Rianti, P., Juliandi, B., Suryobroto, B., & Raymond, M. (2019). Male homosexual preference: Femininity and the older brother effect in Indonesia. *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior*, 17(4), 1474704919880701.
- Nsi Akoue, G., Mbading-Mbading, W., Willaume, E., Souza, A., Mbatchi, B., & Charpentier, M. J. E. (2017). Seasonal and individual predictors of diet in a free-ranging population of mandrills. *Ethology: Formerly Zeitschrift Für Tierpsychologie*, 123(9), 600–613.
- Oh, K. P., & Badyaev, A. V. (2010). Structure of social networks in a passerine bird: consequences for sexual selection and the evolution of mating strategies. *The American Naturalist*, 176(3), E80-9.
- Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4), 481–487.
- Owsley, C. (2013). Visual processing speed. *Vision Research*, 90, 52–56.
- Peignot, P., Charpentier, M. J. E., Bout, N., Bourry, O., Massima, U., Dosimont, O., Terramorsi, R., & Jean Wickings, E. (2008). Learning from the first release project of captive-bred mandrills *Mandrillus sphinx* in Gabon. *Oryx: The Journal of the Fauna Preservation Society*, 42(1), 122–131.
- Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., & Minamisawa, R. (1999). Menstrual cycle alters face preference. *Nature*, 399(6738), 741–742.

- Penton-Voak, Ian S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior: Official Journal of the Human Behavior and Evolution Society*, 25(4), 229–241.
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., Henzi, S. P., Castles, D. L., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, 394(6696), 884–887.
- Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgments of female attractiveness. *Nature*, 368, 239–242.
- Peters, M., Simmons, L. W., & Rhodes, G. (2009). Preferences across the menstrual cycle for masculinity and symmetry in photographs of male faces and bodies. *PloS One*, 4(1), e4138.
- Peterson, J. C., Abbott, J. T., & Griffiths, T. L. (2016). Adapting deep network features to capture psychological representations. In *arXiv [cs.CV]*. arXiv. <http://arxiv.org/abs/1608.02164>
- Posner, M. I. (1970). Abstraction and the process of recognition. In *Psychology of Learning and Motivation* (pp. 43–100). Elsevier.
- Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77(3, Pt.1), 353–363.
- Pound, N., Penton-Voak, I. S., & SurrIDGE, A. K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proceedings. Biological Sciences*, 276(1654), 153–159.
- Prabhu. (2018, March 4). *Understanding of Convolutional neural network (CNN) — deep learning*. Medium. <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>
- Prum, R. O. (2018). *Evolution of beauty*. Random House.

- Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evolution and Human Behavior: Official Journal of the Human Behavior and Evolution Society*, 31(3), 157–175.
- Quinn, P. C., Conforto, A., Lee, K., O’Toole, A. J., Pascalis, O., & Slater, A. M. (2010). Infant preference for individual women’s faces extends to girl prototype faces. *Infant Behavior & Development*, 33(3), 357–360.
- Reber, P. J., Stark, C. E., & Squire, L. R. (1998). Contrasting cortical activity associated with category memory and recognition memory. *Learning & Memory (Cold Spring Harbor, N.Y.)*, 5(6), 420–428.
- Reber, R., & Schwarz, N. (2006). Perceptual fluency, preference, and evolution. *Polish Psychological Bulletin*, 37(1), 16–22.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver’s processing experience? *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, 8(4), 364–382.
- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, 9(1), 45–48.
- Redies, C., Hänisch, J., Blickhan, M., & Denzler, J. (2007). Artists portray human faces with the Fourier statistics of complex natural scenes. *Network (Bristol, England)*, 18(3), 235–248.
- Renoult, J. P., & Mendelson, T. C. (2019). Processing bias: extending sensory drive to include efficacy and efficiency in information processing. *Proceedings. Biological Sciences / The Royal Society*, 286(1900). <https://doi.org/10.1098/rspb.2019.0165>
- Renoult, Julien P., Bovet, J., & Raymond, M. (2016). Beauty is in the efficient coding of the beholder. *Royal Society Open Science*, 3(3), 160027.



- Rhodes, G., Yoshikawa, S., Clark, A., Lee, K., McKay, R., & Akamatsu, S. (2001). Attractiveness of facial averageness and symmetry in non-western cultures: in search of biologically based standards of beauty. *Perception, 30*(5), 611–625.
- Rhodes, Gillian. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.
- Rhodes, Gillian, Sumich, A., & Byatt, G. (1999). Are average facial configurations attractive only because of their symmetry? *Psychological Science, 10*(1), 52–58.
- Rhodes, Gillian, & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science, 7*(2), 105–110.
- Rhodes, Gillian, Yoshikawa, S., Palermo, R., Simmons, L. W., Peters, M., Lee, K., Halberstadt, J., & Crawford, J. R. (2007). Perceived health contributes to the attractiveness of facial symmetry, averageness, and sexual dimorphism. *Perception, 36*(8), 1244–1252.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*(11), 1019–1025.
- Rocco, I., Arandjelovic, R., & Sivic, J. (2018, June). End-to-end weakly-supervised semantic alignment. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT. <https://doi.org/10.1109/cvpr.2018.00723>
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. In *Cognitive Psychology* (Vol. 7, Issue 4, pp. 573–605). [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9)
- Rosenthal, G. G. (2018). Evaluation and hedonic value in mate choice. *Current Zoology, 64*(4), 485–492.

- Rosenthal, G. G., Wagner, W. E., Jr, & Ryan, M. J. (2002). Secondary reduction of preference for the sword ornament in the pygmy swordtail *Xiphophorus nigrens* (Pisces: Poeciliidae). *Animal Behaviour*, *63*(1), 37–45.
- Ryali, C. K., Goffin, S., Winkielman, P., & Yu, A. J. (2020). From likely to likable: The role of statistical typicality in human social assessment of faces. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(47), 29371–29380.
- Ryan, M. J., & Rand, A. S. (1990). The sensory basis of sexual selection for complex calls in the tungara frog, *Physalaemus pustulosus* (sexual selection for sensory exploitation). *Evolution; International Journal of Organic Evolution*, *44*(2), 305.
- Sanchez-Pages, S., Rodriguez-Ruiz, C., & Turiegano, E. (2014). Facial masculinity: how the choice of measurement method enables to detect its influence on behaviour. *PloS One*, *9*(11), e112157.
- Schaeffer, J.-M. (2015). *L'expérience esthétique*. Gaillmard.
- Schug, J., Alt, N. P., & Klauer, K. C. (2015). Gendered race prototypes: Evidence for the non-prototypicality of Asian men and Black women. *Journal of Experimental Social Psychology*, *56*, 121–125.
- Scott, I. M. L., Pound, N., Stephen, I. D., Clark, A. P., & Penton-Voak, I. S. (2010). Does masculinity matter? The contribution of masculine face shape to male attractiveness in humans. *PloS One*, *5*(10), e13585.
- Sczesny, S., Spreemann, S., & Stahlberg, D. (2006). Masculine = Competent? Physical Appearance and Sex as Sources of Gender-Stereotypic Attributions. In *Swiss Journal of Psychology* (Vol. 65, Issue 1, pp. 15–23). <https://doi.org/10.1024/1421-0185.65.1.15>
- Seeliger, K., Fritsche, M., Güçlü, U., Schoenmakers, S., Schoffelen, J.-M., Bosch, S. E., & van Gerven, M. A. J. (2018). Convolutional neural network-based encoding and decoding of visual object recognition in space and time. *NeuroImage*, *180*, 253–266.

- Seibert, D., Yamins, D. L., Ardila, D., Hong, H., DiCarlo, J. J., & Gardner, J. L. (2016). A performance-optimized model of neural responses across the ventral visual stream. In *bioRxiv*. <https://doi.org/10.1101/036475>
- Serre, T., Kreiman, G., Kouh, M., Cadieu, C., Knoblich, U., & Poggio, T. (2007). A quantitative theory of immediate visual recognition. *Progress in Brain Research*, *165*, 33–56.
- Setchell, J., Lee, P. C., & Wickings, E. J. (2002). Reproductive Parameters and Maternal Investment in Mandrills (*Mandrillus sphinx*). *International Journal of Primatology*, *23*, 51–68.
- Setchell, J. M. (2005). Do female mandrills prefer brightly colored males? *International Journal of Primatology*, *26*(4), 715–735.
- Setchell, J. M., Charpentier, M., & Wickings, E. J. (2005). Sexual selection and reproductive careers in mandrills (*Mandrillus sphinx*). *Behavioral Ecology and Sociobiology*, *58*(5), 474–485.
- Setchell, J. M., & Jean Wickings, E. (2006). Mate Choice in Male Mandrills (*Mandrillus sphinx*). *Ethology: Formerly Zeitschrift Für Tierpsychologie*, *112*(1), 91–99.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Technic. J*, *27*, 379–423.
- Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. <https://arxiv.org/pdf/1409.1556.pdf>
- Smit, N., Baniel, A., Roura-Torres, B., Amblard-Rambert, P., Charpentier, M. J. E., & Huchard, E. (2022). Sexual coercion in a natural mandrill population. *Peer Community Journal*, *2*(e36). <https://doi.org/10.24072/pcjournal.134>
- Smith, M. J. L., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D. R., Boothroyd, L. G., Durrani, S. J., Stirrat, M. R., Whiten, S., Pitman, R. M., & Hillier, S.

- G. (2006). Facial appearance is a cue to oestrogen levels in women. *Proceedings. Biological Sciences / The Royal Society*, 273(1583), 135–140.
- Stephen, I. D., Oldham, F. H., Perrett, D. I., & Barton, R. A. (2012). Redness enhances perceived aggression, dominance and attractiveness in men's faces. *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior*, 10(3), 562–572.
- Stockley, P., & Bro-Jørgensen, J. (2011). Female competition and its evolutionary consequences in mammals. *Biological Reviews of the Cambridge Philosophical Society*, 86(2), 341–366.
- Symons, D. (1995). Beauty is in the adaptations of the beholder: The evolutionary psychology of human female sexual attractiveness. In P. R. D. Abramson S (Ed.), *Sexual nature, sexual culture* (pp. 80–119). The University of Chicago Press.
- Tamimi, R. M., Lagiou, P., Mucci, L. A., Hsieh, C.-C., Adami, H.-O., & Trichopoulos, D. (2003). Average energy intake among pregnant women carrying a boy compared with a girl. *BMJ (Clinical Research Ed.)*, 326(7401), 1245–1246.
- Thornhill, R., & Møller, A. P. (2007). Developmental stability, disease and medicine. *Biological Reviews of the Cambridge Philosophical Society*, 72(4), 497–548.
- Thornton, J., Zehr, J. L., & Loose, M. D. (2009). Effects of prenatal androgens on rhesus monkeys: A model system to explore the organizational hypothesis in primates. *Hormones and Behavior*, 55(5), 633–644.
- Trivers, R. L., & Willard, D. E. (1973). Natural selection of parental ability to vary the sex ratio of offspring. *Science (New York, N.Y.)*, 179(4068), 90–92.
- Valentine, T., Darling, S., & Donnelly, M. (2004). Why are average faces attractive? The effect of view and averageness on the attractiveness of female faces. *Psychonomic Bulletin & Review*, 11(3), 482–487.

- van de Beek, C., van Goozen, S. H. M., Buitelaar, J. K., & Cohen-Kettenis, P. T. (2009). Prenatal sex hormones (maternal and amniotic fluid) and gender-related play behavior in 13-month-old Infants. *Archives of Sexual Behavior*, 38(1), 6–15.
- Volokitin, A., Roig, G., & Poggio, T. (2017). Do Deep Neural Networks suffer from crowding? In *arXiv [cs.CV]*. arXiv. <http://arxiv.org/abs/1706.08616>
- Vrana, S. R., & Van den Bergh, O. (1995). Perceptual and perceptual-motor fluency as a basis for affective judgements: Individual differences in motor memory activation. *Cognition & Emotion*, 9(6), 529–547.
- Waitt, C., Little, A. C., Wolfensohn, S., Honess, P., Brown, A. P., Buchanan-Smith, H. M., & Perrett, D. I. (2003). Evidence from rhesus macaques suggests that male coloration plays a role in female primate mate choice. *Proceedings. Biological Sciences*, 270 Suppl 2(suppl\_2), S144-6.
- Walton, G. E., & Bower, T. G. R. (1993). Newborns form “prototypes” in less than 1 minute. *Psychological Science*, 4(3), 203–205.
- Weatherhead, P. J., & Robertson, R. J. (1979). Offspring quality and the polygyny threshold: “the sexy son hypothesis.” *The American Naturalist*, 113(2), 201–208.
- Westerman, D. L., Klin, C. M., & Lanska, M. (2015). On the (elusive) role of oral motor-movements in fluency-based memory illusions. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 41(4), 1003–1013.
- Wilkinson, G. S., & Reillo, P. R. (1994). Female choice response to artificial selection on an exaggerated male trait in a stalk-eyed fly. *Proceedings. Biological Sciences*, 255(1342), 1–6.
- Willmore, B., & Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Network (Bristol, England)*, 12(3), 255–270.

- Wilson, V. A. D., Kade, C., & Fischer, J. (2021). Testing the relationship between looking time and choice preference in long-tailed macaques. *Animal Behavior and Cognition*, 8(3), 351–375.
- Windhager, S., Schaefer, K., & Fink, B. (2011). Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *American Journal of Human Biology: The Official Journal of the Human Biology Council*, 23(6), 805–814.
- Winkielman, P., Berntson, G. G., & Cacioppo, J. T. (2007). The psychophysiological perspective on the social mind. In *Blackwell Handbook of Social Psychology: Intraindividual Processes* (pp. 89–108). Blackwell Publishers Inc.
- Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science*, 17(9), 799–806.
- Winters, S., Dubuc, C., & Higham, J. P. (2015). Perspectives: The looking time experimental paradigm in studies of animal visual perception and cognition. *Ethology: Formerly Zeitschrift Für Tierpsychologie*, 121(7), 625–640.
- Wong, B. B. M., & Rosenthal, G. G. (2006). Female disdain for swords in a swordtail fish. *The American Naturalist*, 167(1), 136–140.
- Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23), 8619–8624.
- Zahavi, A. (1997). *Handicap principle, the: A missing piece of Darwin's puzzle*. Oxford University Press.

Zebrowitz, L. A., & Rhodes, G. (2004). Sensitivity to “bad genes” and the anomalous face overgeneralization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior*, 28(3), 167–185.







– RÉSUMÉ –

L'étude de l'esthétique, influencée par le courant philosophique interactionniste, considère la beauté comme le résultat de l'interaction entre l'objet observé et l'observateur. La théorie de la fluence perceptive, issue des sciences cognitives, suggère que l'attractivité d'un objet repose sur un traitement rapide, précis et économique en terme de coût cognitif. Cependant, quantifier la fluence et l'appliquer au-delà du domaine humain représente un défi. Cette thèse explore ce phénomène chez une autre espèce de primates, le mandrill (*Mandrillus sphinx*), présenté dans le Chapitre 1. Dans les Chapitres 2 et 3, une dimension de la fluence perceptive est explorée, à savoir la prototypicalité. La théorie de la fluence suggère que les prototypes sont préférés en raison de leur traitement fluent. En particulier, la féminité, envisagée comme une forme particulière de cette prototypicalité, a été quantifiée à l'aide des réseaux neuronaux convolutifs (CNN) utilisés comme modèles de perception visuelle. La mesure de féminité développée explique jusqu'à 26% de la féminité perçue et 18% de l'attractivité évaluée chez les humains. Dans le Chapitre 2, j'examine l'impact de la féminité des femelles mandrills adultes sur les comportements socio-sexuels de leurs conspécifiques. Les résultats indiquent que les femelles moins féminines suscitent davantage d'attention et d'intérêt sexuel de la part de leurs partenaires sociaux, remettant en question l'application de la théorie de la fluence dans ce contexte. Le Chapitre 3 explore les causes et les conséquences de la variation de la féminité faciale chez les mandrills mâles juvéniles. Les résultats indiquent que les mâles les plus féminins naissent de mères plus jeunes et ont, en moyenne, moins de grandes sœurs. Cependant, une féminité accrue chez certains mâles ne semble pas influencer de manière significative leurs comportements sociaux. Le Chapitre 4, à visée méthodologique, élargit la perspective en développant une mesure de prototypicalité applicable à l'espèce humaine mais généralisable à d'autres. Ainsi, une nouvelle métrique, la "typicalité statistique," est élaborée en utilisant les CNN et comparée à une autre mesure de fluence, la sparsité. Cette typicalité statistique n'explique, toutefois, que 11% de la variance dans l'attractivité chez les humains, tandis que la sparsité en explique 27%. Dans l'ensemble, cette thèse cherche à établir des liens entre beauté, fluence, modélisation de la perception avec des outils tirées de l'intelligence artificielle et comportements sociaux.

Mots-clés: Beauté, Féminité, Écologie comportementale, Mandrills, Théorie de la fluence, Traitement de l'information, IA

– ABTRACT –

Contemporary aesthetics, influenced by the interactionist philosophical movement, posits that beauty results from the interaction between an observed object and the observer. Perceptual fluency theory, stemming from cognitive sciences, suggests that an object's attractiveness relies on fast, precise, and cognitively economical information processing. However, quantifying fluency and extending its application beyond humans is challenging. This PhD thesis explores this phenomenon in another primate species, the mandrill (*Mandrillus sphinx*), introduced in Chapter 1. In Chapters 2 and 3, a focus is done on one dimension of perceptual fluency, the prototypicality. The fluency theory proposes that prototypes are preferred due to their fluent processing. More specifically, femininity, considered as a particular form of prototypicality, is quantified using convolutional neural networks (CNNs) as models of visual perception. The femininity metric developed explains up to 26% of perceived femininity and 18% of perceived attractiveness in humans. In Chapter 2, the impact of adult female mandrills' femininity on their socio-sexual behaviors with their groupmates is examined. Results indicate that less feminine females attract more attention and sexual interest from their conspecifics, challenging the application of the fluency theory in this context. In Chapter 3, I study the causes and consequences of facial femininity variation among juvenile male mandrills. Results indicate that the most feminine males are born from young mothers and have fewer older sisters than more masculine males. However, this increased femininity does not change behaviors of these juvenile males. Chapter 4, a methodological study, broadens the perspective by developing a measure of prototypicality applicable to humans and generalizable to other species. A new metric, the "statistical typicality," is generated using CNNs and compared to another fluency measure, the sparsity. Statistical typicality only explains, however, 11% of the variance in attractiveness in humans, while sparsity accounts for 27% of this variance. Overall, this PhD thesis seeks to establish interactions among beauty, fluency, perception modeling using artificial intelligence tools, and social behaviors.

Keywords: Beauty, Femininity, Fluency Theory, Behavioral Ecology, Mandrills, Information Processing, AI